THE UNIVERSITY OF CHICAGO

ESSAYS ON ENERGY ECONOMICS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE IRVING B. HARRIS

GRADUATE SCHOOL OF PUBLIC POLICY STUDIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

MICHELE ROSSELOT DAVIES

CHICAGO, ILLINOIS

JUNE 2017

For my husband, whose love and support made this possible. And for my parents, who raised me to be curious about the world and gave me a firm foundation.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the guidance and patience of my committee. Jim Sallee, thank you for always being interested and for asking the questions that led me to answers. Dan Black and Bob Rosner, thank you for keeping me on track and shepherding me through. I also acknowledge Thom Covert who has been generous with his time and expertise. Thank you to Cynthia Cook-Conley for always being there and to Fay Booker for saving me and my data. Thank you to Lydia Veliko for helping me with scheduling and, more importantly, for helping me keep perspective.

Thank you to my companions in the journey over these sometimes difficult years when I was not always at my best. Becka, Jenn, NaYoung and Pallavi, you've been understanding, encouraging and smiling. You've made it possible for me to keep coming to my cube every day. To my friends and family outside of Harris - Mom, Alex, AJ, Bonnie, Amy, Pang – thank you for understanding when I've disappeared for months at a time and for being there when I re-emerge. Your belief in me has given me courage. And finally, to my husband, George, thank you for your patience and understanding, for never mentioning "trics", and for your love and support.

Contents

List of Figures

List of Tables

Chapter 1

Discrimination and Skill in Bargaining Outcomes: Evidence from Mineral Rights Leases

1.1 Introduction & Literature Review

Since the 1980s, there have been major technological advances in oil and gas extraction technologies that have allowed resources to be extracted from under densely populated areas. The combination of massive hydraulic fracturing and horizontal drilling have made natural gas production from shale formations economically viable. The first shale formation to be explored using this new combination of techniques was the Barnett Shale in north central Texas, underlying the city of Fort Worth and the surrounding areas to the north, west and south. Mineral rights are generally owned by the surface land-owner and so this new technology has placed thousand of urban and suburban land-owners in the position to negotiate mineral rights leases for the natural gas under their houses. This is a major event in the financial lives of the affected households, one for which they probably do not have previous experience or much prior information.

Natural gas extraction companies working in shale formations are negotiating with nu-

merous small land owners and assembling leases quickly in these densely populated areas. The literature in personal and household finance suggests that in the U.S. racial and ethnic minorities negotiate or are offered less favorable terms than white households for mortgages, cars and other forms of credit.

In this paper I ask two main questions about mineral rights lease negotiations. First, are leasing companies prejudiced against Black or Hispanic households in their leasing practice leading to less advantageous leases for these households? And is this prejudice mitigated by competition? Second, do some households have greater skill in negotiating leases? In other words, are households with higher levels of education able to negotiate for better lease terms? The answers to these questions shed light, not only on the relationships between households and natural resource extraction companies, but also on broader questions of personal and household finance.

I find that when I examine two main components of a mineral rights lease contract, the royalty rate, which is arguably the most important aspect, and the length of the lease that minority status and education have different effects. Royalties are lower in minority neighborhoods, especially in Hispanic neighborhoods with lower levels of competition. This effect is largely explained by lower levels of education, which have a strong correlation with negotiated royalty rates. Shorter lease lengths, which are more favorable to the land owner, are also strongly associated with higher levels of education. Both Black and Hispanic neighborhoods have longer, less favorable, lease lengths, which are not explained by educational differences. Again, areas with higher levels of competition show lease lengths more favorable to land owners. In order to understand why the relationship between race, ethnicity and education are different for royalties than for lease length, I explore a potential mechanism. Records of internet searches during this time period show that the term "royalty" has a high number of searches, where as "lease length" or other close alternatives were not searched enough to even show up in the data. The much greater salience of royalties suggests that all demographics were more educated about royalties than lease lengths.

1.1.1 Literature Review

Broadly speaking, the literature distinguishes between two types of discrimination, taste-based, first formalized by Becker (1957), and statistical discrimination, where a group average is attributed to an individual. Becker (1957) suggests that with taste-based discrimination an employer will pay an employee from an non-preferred group a lower wage, such that the full cost to the employer, the wage plus the cost of distaste for that worker, is not greater than the cost of employing someone from a preferred group. This line of thinking applies to gender-based discrimination, as well as race-based discrimination. In a model of gender wage discrimination, Flabbi (2010) shows that not all employers need to be prejudiced in order for wage differences to appear between men and women even in firms where the employer is not prejudiced. Because of the existence of some prejudiced employers, women have a lower outside option and so even non-prejudiced firms will offer women lower wages. Becker suggests that increasing competition will decrease this discriminatory behavior. Indeed, Black and Brainerd (2004) finds that the male-female wage gap in the US decreases as imports to the US increase.

A literature in personal finance suggests that lower levels of education lead to disadvantageous decision-making about credit and borrowing. Woodward (2008) finds that education, measured as percentage of people in a census tract with a college education, is an important indicator of the fees they will pay when originating a mortgage. A mortgage in a census tract with 100% of the adult population college educated will cost $1100 less in fees than in a census tract with no college graduates, holding all else equal. In magnitude this is three times greater than the discount received by non-minorities relative to African-American borrowers. Woodward also finds that race and education differences disappear when the mortgage is of the simpler, no-cost form. She concludes that lenders and brokers present more favorable terms to borrowers they expect to be better informed. Similarly, Ayres and Siegelman (1995) finds that in negotiating a new car price black men, black women and white women were all given higher initial prices by car dealers than white men, and in addition, they were less able to negotiate down from that initial higher offer. Dealers, on average, made triple the

profit on black men as on white men. This paper suggests that the discrimination may not be taste-based, but rather statistical discrimination, reflecting assumptions made by sellers about the search costs buyers face and their level of information about prices and their negotiability. In a closely related setting to this study, Vissing (2015) finds that lease language which offers environmental protections are less common in minority neighborhoods.

1.2 Institutional Context

In the US landowners own the minerals on or under their land, unlike most other countries in the world where minerals are owned by the state. This means that oil, gas or mining companies must negotiate with private landowners for access to their minerals. In the late 1990s, the first commercially viable shale gas well was developed in the Barnett Shale, which underlies the city of Fort Worth, Texas and the areas to the north, west and south. In conventional oil and gas production where large quantities of oil or gas can be tapped from one drilling location, but those locations are scarce and difficult to find. In contrast, when producing oil or gas from shale, much smaller amounts of hydrocarbons are produced from any one well, but over a shale formation production is almost certain. The shale production business, then, becomes a race to acquire as vast a portfolio of mineral rights as quickly as possible. Mineral rights leasing and production began in the rural areas, where population density is low and the average lease size ran to hundreds of acres. Starting in the mid-2000s, firms began to negotiate leases in the more densely populated suburban and urban areas of Fort Worth, where they began to encounter a much broader demographic range of counter-parties.

The primary component of a mineral rights leases is the royalty payment, a percentage of each month's production for as long as the well is producing, and sometimes an up-front signing bonus. The State of Texas sets a floor for the royalty on minerals at 1/8 or 12.5%. For most people, this will be a once-in-a-lifetime transaction, where they are unlikely to have had previous exposure to mineral rights leases. Another important component of the lease is the length of the term. Generally, a longer lease term is advantageous for the company

Year	Leases Number	Median Royalty %	Median Area Acres	Median Lease Length Months
2006	28,848	20	0.36	36
2007	90,516	22	0.22	60
2008	62,819	25	0.25	36
2009	10,739	25	0.25	36
2010	13,584	25	0.25	36
2011	12,848	25	0.25	36
2012	5,836	22	0.23	36
2013	846	25	0.36	36
Total	226,036	25	0.24	36

Table 1.1: Lease summary statistics

because it gives them more flexibility to develop the well, whereas a shorter lease term is advantageous to the mineral rights owner because they will start to receive royalties more quickly.

1.3 Data

The Barnett Shale, seen in Figure 1.1, spans urban, suburban and rural areas, which allows us to examine the bargaining outcomes for a wide variety of populations. I use geo-coded lease data from 2006 to 2013 for the Barnett Shale area that includes the date of the lease, acreage, royalty rate, and length of the lease for more than 240,000 leases, summarized in Table 1.1. Bonus payments are not required by law to be reported to the state and so the data includes bonus payments for only 718 of the leases. While bonuses can be a meaningful part of the conpensation for mineral rights leases, I do not use them for my analysis because of their sparse representation in my data. Leasing activity peaks in 2007, as shown in Figure 1.2. The lease data comes from DrillingInfo, a commercial supplier of data to the oil and gas industry.

I spatially match the leases to Census block-group level demographic data from the 2013 five-year American Community Survey for the years 2009 to 2013 and demographic data at the Census tract level from the 2010 five-year ACS for the years 2006 to 2008. A block group contains between 600 and 3,000 individuals, and tracts have an average population of 4000 people. Census tracts are areas roughly equivalent to neighborhoods. Several block groups

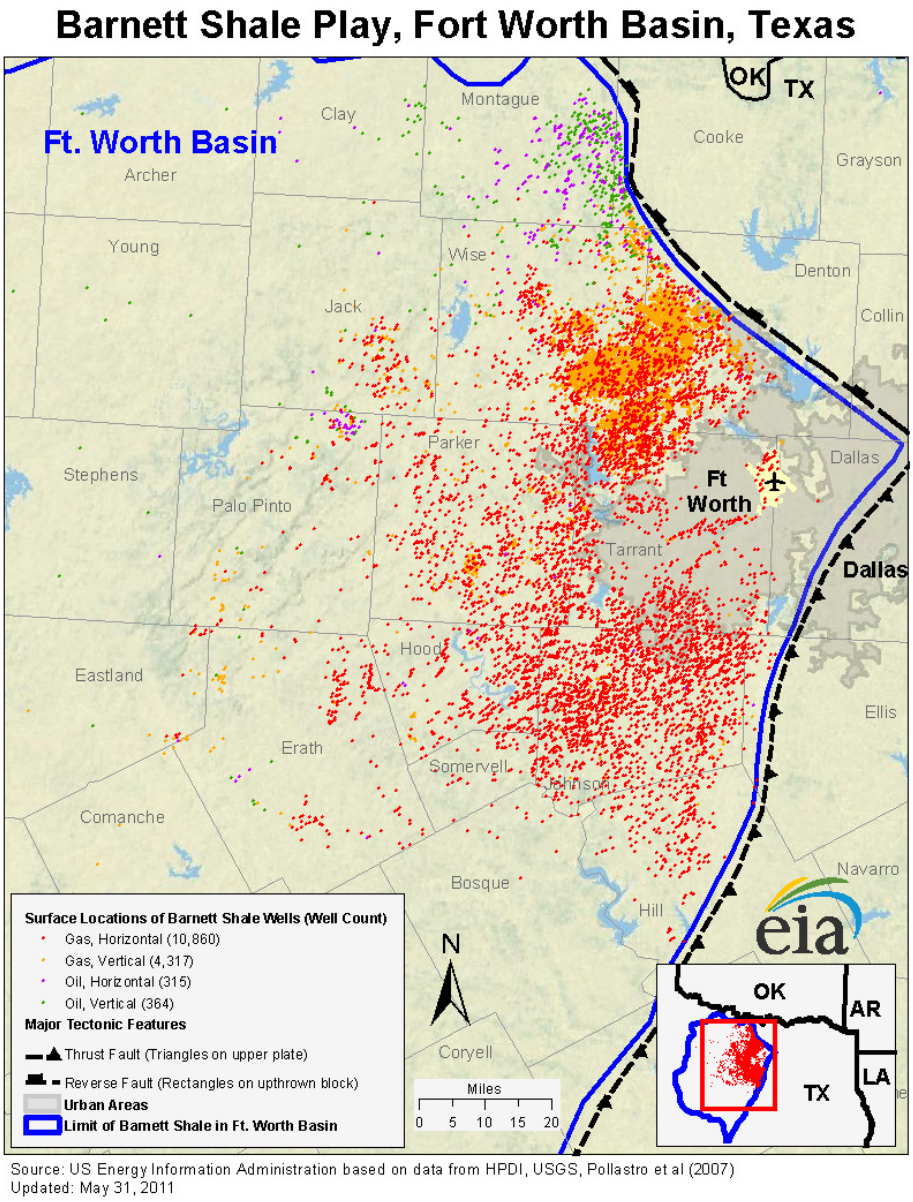

Figure 1.1: Barnett Shale map with producing wells, 2011

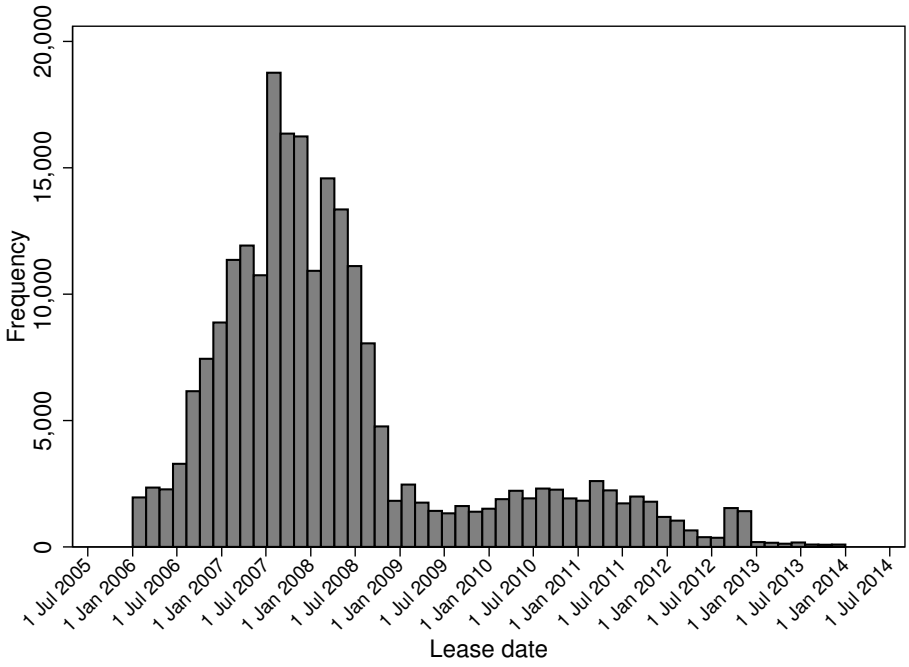

Figure 1.2: Leasing activity, 2006 to 2013

Table 1.2: Educational attainment by census block group for 2009-2013

Highest Educational Attainment	Obs	Mean	Std. Dev.	Min	Max
Percent 6th Grade or Less	1,113	2.80%	4.30%	0.00%	40.30%
Percent 7th to 12th Grade	1,113	6.30%	5.70%	0.00%	29.40%
Percent High School Diploma	1,113	13.20%	6.30%	0.00%	36.10%
Percent Associates Degree	1,113	4.40%	2.90%	0.00%	20.80%
Percent Bachelors Degree	1,113	13.20%	8.90%	0.00%	49.90%
Percent Masters Degree	1,113	4.60%	4.10%	0.00%	21.70%
Percent Professional Degree	1,113	0.90%	1.50%	0.00%	10.50%
Percent Doctorate Degree	1,113	0.60%	1.20%	0.00%	12.70%

make up each tract. Summary statistics for educational attainment by census block group are found in Table 1.2. It shows that block groups have an average of 13.2% of the population holding high school diplomas as their final degree. 13.2% of the block group populations also hold bachelors degrees. Summary statistics for the racial and ethnic makeup of the census block groups can be found in Table 1.3, where on average, the largest groups are White, at 77.9%, and Black, at 10%. 21% are Hispanic. The Census designates White and Black as races, while Hispanic is an ethnicity, where people may be racially Black, White or some other race. In this study, if people identify as Hispanic, regardless of race, I code them as Hispanic. This study will focus on these three groups, Blacks, Whites and Hispanics.

Table 1.3: Racial and ethnic identification at the census block group level for 2009-2013

Race or Ethnicity	Obs	Mean	Std. Dev.	Min	Max
Percent White	1,113	77.90%	19.50%	0.00%	100.00%
Percent Black	1,113	10.00%	14.50%	0.00%	86.40%
Percent Asian	1,113	3.70%	6.20%	0.00%	59.00%
Percent Native American	1,113	0.50%	1.40%	0.00%	16.90%
Percent Pacific Islander	1,113	0.20%	1.90%	0.00%	47.30%
Percent Other	1,113	5.10%	9.10%	0.00%	71.50%
Percent Two or More	1,113	2.50%	3.50%	0.00%	38.00%
Percent Hispanic	1,113	22.10%	20.30%	0.00%	100.00%

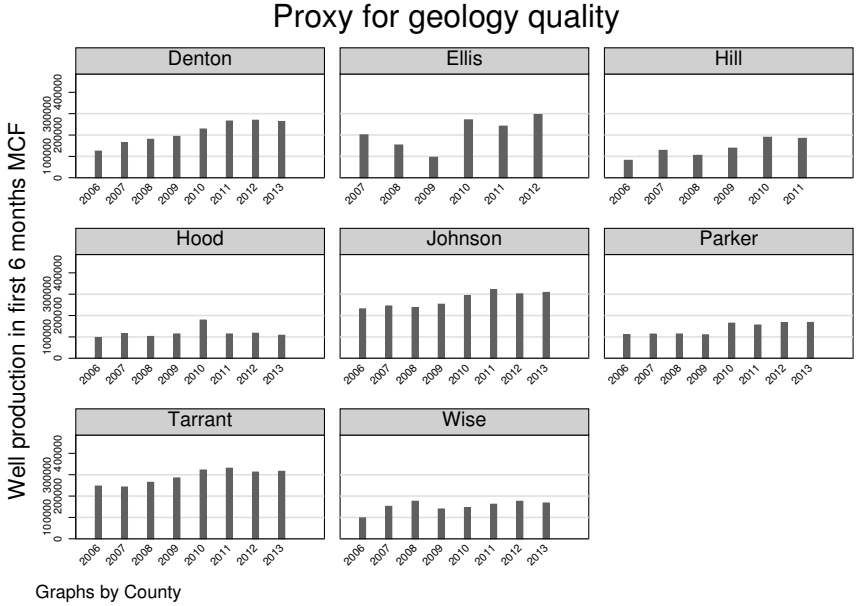

Figure 1.3: Geology quality measure

Also from DrillingInfo, I have geo-coded well production data. I use this to construct a measure of geology quality. Because shale formations vary slowly over space, any given well will produce very similarly to a nearby well, all else equal. I use as a control information that leasing companies would also have publicly available to them. My measure averages the early production from wells starting production in a county in the calendar quarter preceding the signing date of each lease. [1]The yearly averages for this measure can be seen in Figure 1.3.

In order to understand whether competition has an effect on negotiated outcomes, I

[1] DrillingInfo has also constructed a measure of geology quality measure based on thermal maturity, depth and thickness of the shale. I do not use this measure because using production from nearby wells has more explanatory power.

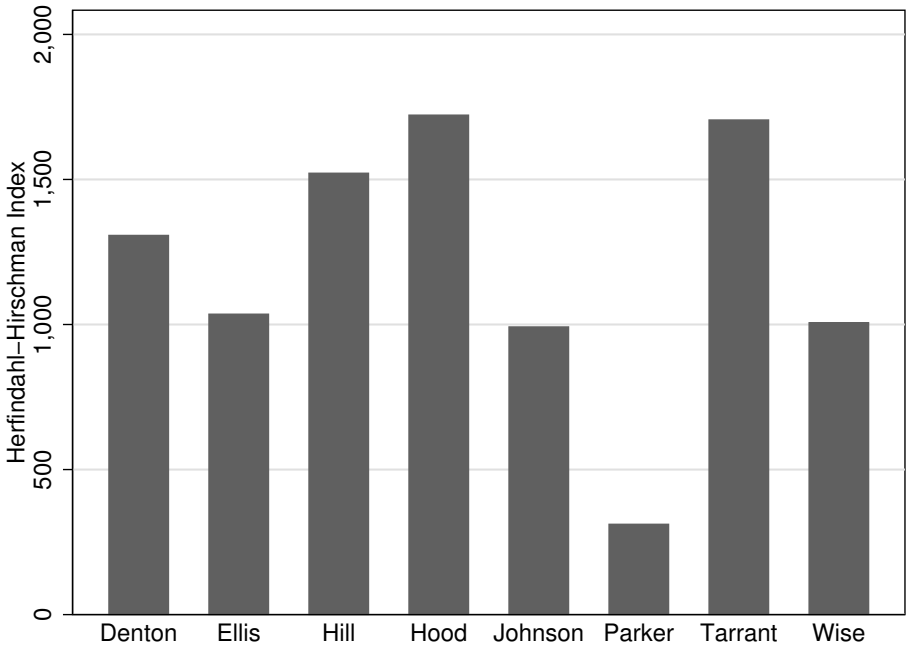

Figure 1.4: Herfindahl-Hirschman Index, reflecting market concentration for each county

calculate a Herfindahl-Hirschman Index (HHI), which is a commonly used measure of market concentration. It is constructed by squaring the market shares of all firms in an market and then adding them up. The HHI can range in values from close to zero up to 10,000 if, for example, one firm had 100% market share the HHI would be 100^2, which is 10,000. I measure market share as the percentage of a county's land area leased by each firm measured at the end of the period in 2013. My measure most likely understates industry concentration, as not all of each county's land is available for leasing given municipal-level restrictions, among other things. The HHI for each county can be seen in Figure 1.4. The U.S. Justice Department considers markets with an HHI of between 1,500 and 2,500 to be "moderately concentrated" and above 2,500 to be "highly concentrated".

1.4 Model

The firm's objective is to maximize profit at each well site.

$$Profit = (1 - Royalty) * Production - Costs \tag{1.1}$$

Taking drilling costs and well production as given, this translates into the firm wishing to minimize royalty payments, and in order to maximize flexibility, the firm also wishes to maximize the term length of the lease. On the other hand, the household's objective is to maximize their royalty and minimize the term length of the lease to bring those royalty payments forward in time. While my data does not speak to the negotiation process, it does allow me to see the results of successful negotiations.

I use an ordinary least squares regression to estimate the effect of education and race on bargaining outcomes. I take the following model to the data:

$$Y_{in} = X_n\beta + Z_n\alpha + \gamma S_n + \varepsilon_i \tag{1.2}$$

where Y_i is the negotiated outcome variable, either the royalty rate or lease length, for lease i in neighborhood n. The vector X represents the geological characteristics of the area. Z is a vector of neighborhood demographic characteristics, including race and ethnicity, and S is a neighborhood measure of negotiation skill, which I proxy by average years of education in the neighborhood. The errors are clustered at the county level.

1.5 Royalties

1.5.1 Education

While I cannot directly observe individual education levels in the data, I take the education levels of the census block group or tract, depending on the year, as a proxy. In the case of negotiating mineral rights lease agreements in the Barnett Shale, Table 1.4 shows that a neighborhood's average years of schooling has a statistically significant effect on the negotiated royalty. This holds even after controlling for household income, which can be seen in columns (4) - (6) and industry concentration in columns (2)-(4). The negative coefficient on Lease length, in column (3) suggests that there is no trade-off between royalties and lease length, but rather that households that negotiate good royalty rates also negotiate shorter leases. In other words, their leases are better on both dimensions.

Table 1.4: Education and royalty terms

VARIABLES	(1) Royalty %	(2) Royalty %	(3) Royalty %	(4) Royalty %	(5) Royalty %	(6) Royalty %
Ave years of schooling	0.158***	0.192***	0.154**	0.287***	0.289***	0.289***
	(0.0292)	(0.0501)	(0.0601)	(0.0160)	(0.0274)	(0.0273)
Lease length, months			-0.0187**			
			(0.00558)			
HHI industry concentration		0.00340	0.00361		0.00336	0.00336
		(0.00201)	(0.00200)		(0.00203)	(0.00202)
Est. median household inc.				-1.29e-05**	-9.71e-06	-9.71e-06
				(3.77e-06)	(6.45e-06)	(6.45e-06)
Acreage per lease						0.000156
						(0.000366)
Constant	14.25***	10.71**	11.74***	13.43***	10.13**	10.12**
	(1.009)	(3.095)	(3.147)	(0.932)	(2.968)	(2.968)
Observations	242,252	242,252	242,252	242,252	242,252	242,252
R-squared	0.112	0.167	0.170	0.116	0.169	0.169
Geology controls	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the county level.

1.5.2 Race and Ethnicity

Table 1.5 presents the regression of royalty rates on the racial/ethnic makeup of neighborhoods, with the percentage of Whites as the omitted category. Column (2) suggests the same pattern as in Table 1.4, that those with higher royalties also have shorter leases, although in this specification it is not statistically significant. Column (3) shows that the greater the Hispanic population, the lower the royalty rate is likely to be controlling for industry concentration. This goes away, and in fact reverses, in column (4) where there is also a control for education. In other words, given a level of education, the proportion of minorities in a neighborhood does not lower the royalty paid.

In order to see if the effects might be more pronounced for neighborhoods where one race or ethnicity dominates, I constructed a dummy for any neighborhood in which the majority of people are of one race or ethnicity. Summary statistics for these neighborhoods are in Table 1.6. We see that out of 1,152 neighborhoods, 908 are majority White. These have

Table 1.5: Race and Ethnicity

VARIABLES	(1) Royalty %	(2) Royalty %	(3) Royalty %	(4) Royalty %	(5) Royalty %
% Hispanic	0.154	0.359	-0.767***	1.717*	-0.540
	(0.962)	(0.697)	(0.164)	(0.890)	(0.768)
% Black	2.107	2.308	0.978	1.817*	1.132
	(1.870)	(1.626)	(0.552)	(0.779)	(1.018)
Lease length, months		-0.0120			
		(0.0166)			
HHI industry concentration			0.00332	0.00297	0.00330
			(0.00203)	(0.00192)	(0.00198)
Ave years of schooling				0.385**	
				(0.139)	
Median HH income					3.21e-06
					(1.06e-05)
Acreage per lease					0.000185
					(0.000361)
Constant	16.22***	16.65***	13.33***	8.363**	13.10***
	(1.008)	(1.490)	(3.147)	(2.780)	(3.125)
Observations	242,252	242,252	242,252	242,252	242,252
R-squared	0.116	0.117	0.164	0.173	0.164
Geology controls	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the county.

Table 1.6: Summary statistics for neighborhoods with a majority race or ethnicity

Majority Race	N Tracts	Mean Education, years	Mean Median HH income, $	Mean Royalty, %	Mean Term length, months	Mean Acreage	Mean HHI (Industry concentration)
Black	47	11.9	33,821	23.2	48.6	4	1,724
Hispanic	115	10.5	35,555	22.5	45.7	14	1,671
White	908	13.5	69,744	21.5	38.9	38	1,492
None	82	12.6	55,736	22.8	43.4	12	1,752
Total	1,152	13.1	63,893	21.8	40.3	32	1,569

higher average levels of education, household income and lot size. Forty-seven neighborhoods are majority Black and 115 are majority Hispanic. There are 82 neighborhoods which have no race or ethnicity in the majority. Black and Hispanic neighborhoods have, on average, lower levels of education and household income than White or no-majority neighborhoods.

Table 1.7 shows results for a specification with dummies for majority Hispanic and majority Black neighborhoods. We can see that once industry concentration is controlled for Hispanic neighborhoods have lower royalty rates than Black or White neighborhoods. Controlling for household income (column (5)) does not change the estimate for Hispanic neighborhoods very much. In column (6), however, where education is introduced as a control, the negative coefficient for Hispanic neighborhoods goes away, suggesting that the lower royalty rates received by Hispanics may be explained by lower levels of education.

1.6 Lease Lengths

Another important term negotiated is the length of the lease. On average, lease lengths in this sample are 44 months. In other words, a company has 44 months to start producing gas from a well on the land, at which point the lease will be held by production (HBP) for as long as the well is producing. If a company does not produce gas within those 44 months, the mineral rights revert to the owner, at which point he or she may lease the mineral rights to another firm. A shorter term pushes a company to develop the well and begin producing gas for themselves and producing royalties for the owner sooner. In Table 1.8, we see that average years of education are an important factor in the lease negotiation. Higher levels of education lead to shorter lease lengths, as do higher household incomes. Not surprisingly, higher levels of industry concentration lead to more favorable terms, longer lease lengths, for

Table 1.7: Royalties in neighborhoods with a majority race or ethnicity

VARIABLES	(1) Royalty %	(2) Royalty %	(3) Royalty %	(4) Royalty %	(5) Royalty %	(6) Royalty %
Majority Hispanic	-0.220	-0.616***	-0.185	-0.205	-0.590**	0.121
	(0.469)	(0.0751)	(0.273)	(0.687)	(0.239)	(0.325)
Majority Black	0.445	0.0339	0.485	0.459	0.0582	0.243
	(0.620)	(0.134)	(0.400)	(0.866)	(0.347)	(0.369)
HHI industry concentration		0.00342			0.00342	0.00333
		(0.00205)			(0.00205)	(0.00201)
Lease length, months			-0.00385			
			(0.0223)			
Median HH income				6.61e-07	1.14e-06	-9.28e-06
				(7.88e-06)	(7.31e-06)	(6.79e-06)
Ave years of schooling						0.305***
						(0.0520)
Constant	16.27***	13.17***	16.42***	16.23***	13.10***	9.918***
	(0.940)	(3.186)	(1.656)	(1.236)	(3.210)	(2.790)
Observations	242,386	242,386	242,386	242,252	242,252	242,252
R-squared	0.109	0.163	0.109	0.109	0.163	0.169
Geology controls	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the county.

Table 1.8: Education and lease length terms

VARIABLES	(1) Lease length months	(2) Lease length months	(3) Lease length months	(4) Lease length months	(5) Lease length months
Ave years of schooling	-2.179***	-2.184***	-2.065***	-1.497***	-1.494***
	(0.0739)	(0.102)	(0.0517)	(0.0674)	(0.0679)
Royalty, %		0.0314			
		(0.261)			
HHI industry concentration			0.0115***	0.0113***	0.0112***
			(0.00212)	(0.00210)	(0.00205)
Median HH income				-5.70e-05***	-5.70e-05***
				(4.98e-06)	(5.02e-06)
Acreage per lease					-0.00391**
					(0.00121)
Constant	66.92***	66.48***	54.92***	51.52***	51.89***
	(6.183)	(8.973)	(2.544)	(2.631)	(2.613)
Observations	242,252	242,252	242,252	242,252	242,252
R-squared	0.092	0.092	0.155	0.161	0.162
Geology controls	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the county.

the firms. In this case, household income of the neighborhood has an additional effect, over and above the effect of education.

I also look at whether race and ethnicity are correlated with negotiated lease lengths. Table 1.9 shows strongly significant effects on the average lease length for a neighborhood correlated with the racial and ethnic makeup of the neighborhood. Unlike the results in Table 1.5 looking at royalties, here we see in columns (4) and (5) that minority status still has explanatory power when education and household income are taken into account. So, with lease lengths, unlike royalties, it seems there is some discrimination, and here it affects both Black and Hispanic households.

In order to understand, as with royalties, if the differences by race and ethnicity can be explained, in part, by education, Table 1.10 shows results for regressing lease lengths for race-majority neighborhoods on average educational attainment and industry concentration. These results show the same general pattern as we saw in Table 1.9, with Whites (the omitted category) negotiating shorter lease lengths, while Blacks and Hispanics have longer

Table 1.9: Race, ethnicity and lease lengths

VARIABLES	(1) Lease length months	(2) Lease length months	(3) Lease length months	(4) Lease length months	(5) Lease length months
% Hispanic	17.12***	14.66***	14.86***	7.031***	6.228**
	(1.366)	(0.634)	(0.563)	(1.624)	(1.845)
% Black	16.79***	14.27***	14.02***	11.37***	10.34***
	(2.447)	(0.832)	(0.814)	(1.245)	(1.455)
HHI industry concentration		0.00903***	0.00816***	0.00928***	0.00929***
		(0.00207)	(0.00165)	(0.00198)	(0.00200)
Ave years of schooling				-1.214***	-0.993***
				(0.156)	(0.121)
Median HH income					-3.12e-05***
					(6.99e-06)
Acreage per lease					-0.00363**
					(0.00121)
Royalty, %		-0.264			
		(0.186)			
Constant	36.33***	32.77***	29.25***	44.92***	44.52***
	(5.416)	(3.582)	(2.354)	(3.269)	(3.310)
Observations	242,252	242,252	242,252	242,252	242,252
R-squared	0.137	0.172	0.166	0.175	0.178
Geology controls	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the county.

Table 1.10: Lease lengths in neighborhoods with a majority race or ethnicity

VARIABLES	(1) Lease length months	(2) Lease length months	(3) Lease length months	(4) Lease length months	(5) Lease length months	(6) Lease length months
Majority Hispanic	9.092***		4.358**	7.889***	5.510***	2.917***
	(0.830)		(1.330)	(0.373)	(0.458)	(0.554)
Majority Black	10.21***		7.881***	8.962***	6.329***	5.654***
	(1.223)		(1.508)	(0.447)	(0.524)	(0.521)
Est. median household inc.					-8.52e-05***	-4.72e-05***
					(4.84e-06)	(5.13e-06)
Acreage per lease					-0.00380**	-0.00378**
					(0.00127)	(0.00120)
HHI industry concentration		0.0123***		0.0104***	0.0102***	0.0105***
		(0.00230)		(0.00194)	(0.00191)	(0.00197)
Ave years of schooling			-1.465***			-1.110***
			(0.135)			(0.0886)
Constant	38.46***	28.03***	57.42***	29.07***	35.12***	46.71***
	(6.731)	(3.186)	(7.666)	(2.670)	(2.436)	(2.777)
Observations	242,386	242,386	242,252	242,386	242,252	242,252
R-squared	0.092	0.079	0.114	0.142	0.165	0.173
Geology controls	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the county.

lease lengths. Unlike the results for royalties, differences in average education do not fully explain the neighborhood differences. Industry concentration also seems to exacerbate this discriminatory effect.

1.7 Salience

The previous sections show that educational differences explain the lower royalties negotiated in minority neighborhoods, but they do not explain the longer lease lengths. It is puzzling that discrimination should be apparent in some provisions of the mineral rights leases but not others. One potential explanation is that royalty rates are a more salient feature of the lease, and that households sought out and therefore had more information about royalties than term lengths.

Figure 1.5 shows relative Google search frequencies for the Dallas-Fort Worth metropolitan area from 2006 to 2013 using Google Trends[2]. This figure reflects the number of searches for the terms "royalty", "mineral rights", "lease term" and "Barnett Shale" over the time period of concern. The Dallas-Fort Worth area, as defined by Google is larger than our area of interest, but there were no other oil or gas producing formations active in this area during this time. Searches for the term "royalty" even exceed searches for the term "Barnett Shale" for all but a six month period. Searches for "lease term" do not even register in the count. Alternatives to "lease term", such as "lease length" and "term length" produced the same results. This suggests that while royalties were the subject of significant interest, households did not educate themselves on lease lengths to the same degree. This lack of knowledge about term lengths allowed firms to extract larger concessions in minority neighborhoods, especially where there was less competition.

1.8 Conclusion

This study demonstrates a relationship between education, race, competition and contract

[2]https://trends.google.com/trends/explore?date=2006-01-01%202013-12-31&geo=US-TX-623&q=royalty,mineral%20rights,lease%20term,barne
Accessed 3/29/2017.

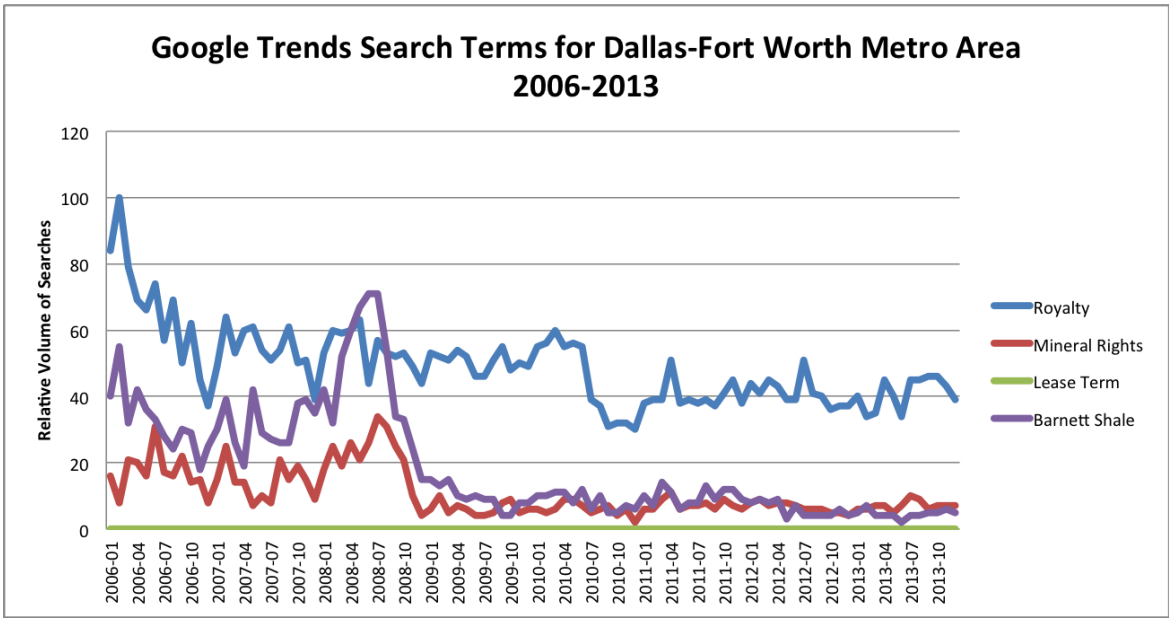

Figure 1.5: Internet searches

negotiations, and that this relationship changes with the salience of the negotiated term. In the case of royalties, I find that lower rates in minority neighborhoods are explained by lower educational attainment. In other words, there does not seem to be discrimination at work, but rather a lower level of education suggests a lower level of negotiation skill or information. In addition, higher levels of industry concentration do lead to greater disparities for minority neighborhoods.

With respect to lease length, which is an aspect of the lease with much lower salience, which induced almost no Google searches, less favorable terms in minority neighborhoods are not explained by education and may be the result of discrimination. Here as well, industry concentration exacerbates the less favorable terms. In the case of mineral right leases, the top-line feature, the royalties where people are informed, the playing field is level, given educational attainment.

These results may be relevant to other personal finance and negotiation settings. For example, Woodward (2008) finds that while minorities negotiate less favorable mortgages, this largely disappears when mortgages are of a simple, no-cost form. In other words, when the secondary, less salient features of a mortgage are eliminated, the negotiation on pri-

mary features proceed without apparent discrimination. This suggests the simplification and standardization of financial contracts may help mitigate discriminatory practices.

Chapter 2

The Tragedy of the Anticommons or Holdout? Evidence from Mineral Rights Leases

2.1 Introduction

In the Barnett Shale of north, central Texas regulation requires a minimum acreage for a natural gas well. In the western portion of the Barnett Shale, where the surface is occupied by ranches and other large tracts of land, this poses no constraint. The eastern part of the Barnett Shale, on the other hand, lies underneath the city and suburbs of Fort Worth, Texas, which has much higher population density and much smaller lot sizes. Pooling these small lots together to meet the acreage requirement can mean signing tens or hundreds of leases. Pooling leased land into a unit large enough to drill a well on is voluntary in Texas and the terms are written into each individual lease. This provides a setting for a holdout problem, or a tragedy of the anticommons, for every well drilled. The tragedy of the anticommons, first formulated in Heller (1998), occurs when "multiple owners are each endowed with the right to exclude others from a scarce resource, and no one has an effective privilege of use." The paper postulates that the prices for all sub-parts will be higher. In holdout, the last

participant, feeling they have monopoly power over the successful completion of a pool will seek to extract all the surplus.

In conventional oil and gas production where large quantities of oil or gas can be tapped from one drilling location, but those locations are scarce and difficult to find. In contrast, when producing oil or gas from shale, smaller amounts of hydrocarbons are produced from any one well, but over a shale formation production is almost certain. The shale production business, then, becomes a race to acquire as vast a portfolio of land as possible. The difference has been described as being between big game hunting and manufacturing.

In this paper, which is one of the few quantitive empirical paper on the anticommons, I ask two questions. First, do we see evidence of the tragedy of the anticommons in this context? Is there a relationship between the number of leases required to make up a pool and the royalty paid for each of them? Second, do we see evidence of holdout? In other words, do the last leases included in a pooled unit show evidence of greater bargaining power on the part of the mineral rights owner through higher negotiated royalties?

After proposing a model of mineral rights assembly, I find that, consistent with the theory of the anticommons, a pooled unit made up of a greater number of leases will have a higher average royalty rate. However, I find scant evidence for holdout. The last lease to enter a pool commands only a slightly higher royalty rate than the first. I take advantage of changes in the regulation of the Barnett Shale to support both of these conclusions.

2.1.1 Literature Review

The tragedy of the commons occurs when too many parties have the right of use of a resource and no one has the right to exclude others. This leads to overuse of this common resource, such as excessive grazing or over-fished fisheries. The paper illustrates the idea of the spatial anticommons, which Heller (1998) introduced with the example of large apartments in Moscow that had been broken up into living space for multiple families during the communist era. After the fall of the Soviet Union, both residents and developers who wanted to sell the apartments found they could not. There was no market for just a living room, for

example; the small pieces individually, each controlled by a separate family, were worth less than the sum of their parts. Any one family declining to sell their portion, however, could prevent the sale of the whole apartment.

The problem of (above ground) land assembly is understood as an impediment to growth in urban areas, where assembling small parcels into larger ones with new uses can increase or change economic activity. In Hornbeck and Keniston (2014), the authors find that areas with more extensive damage from the Great Boston Fire of 1872 had higher levels of economic activity in the long run because the pre-existing infrastrucutre and buildings on small parcels were more easily assembled into larger parcels with higher value use. Brooks and Lutz (2016) finds that land assembly in Los Angeles is at a sub-optimal level because of market frictions such as holdout. Fragmented ownership has also been studied in the case of canal-building in Provence before and after the French Revolution (Rosenthal (1990)) and in patenting in biomedical research (Heller and Eisenberg (1998)).

2.2 Texas Regulation

The Texas Railroad Commission (RRC) regulates all oil and gas production in the state of Texas. In order to fulfill their role of "prevent[ing] waste of the state's natural resources... [and protecting] the correlative rights of different interest owners," (TX Railroad Commission 2017), the RRC establishes rules for spacing and density. All oil and gas wells require certain minimum spacing, acreage and distance from lease lines. When a new oil or gas field starts to be tapped for production, the RRC statutorily defines the field and the rules for producing from it. Companies operating in the field may petition the RRC for changes to the field rules. Because the Barnett Shale was the first commercially exploited shale formation, the initially applied rules were adopted from conventional fields and through numerous petitions were made more appropriate for shale formations. The relevant regulations, or special field rules, determined by the RRC for the Barnett Shale are summarized in Table 2.1.

For the purposes of this paper, the most important component is the optional spacing. Practically speaking, this means that a firm must assemble, at minimum, the number of acres

Table 2.1: Special field rules for natural gas wells in the Barnett Shale

Date	Optional spacing, acres	Between-well spacing, feet	Spacing to lease line, feet
April 1994	80	1200	467
November 1998	40	660	330
August 2005	20	N/A	330
January 2008	20	0	330
September 2009	20	0	330

required to fulfill the optional spacing requirement in order to apply for a permit to drill a well. This decreases from 80 acres to 20 acres from 2005 onward. No minimum between-well spacing requirement exists for the Barnett, which means that multiple wells can be drilled from the same well pad. This does not become common until 2014, which is after our period of study, 1997 to 2013.

Multiple plots are joined together into a unit that satisfies the regulatory requirement through pooling. This is especially important in urban and suburban areas where lot sizes are small, as they are in the eastern portion of the Barnett shale, where it extends underneath the city and suburbs of Fort Worth, Texas. Texas is unique among US states with oil and gas production in that it does not use forced pooling, which requires landowners to join the pool regardless of consent. Pooling, then, in Texas is voluntary, with two mechanisms for overcoming holdout. The first is spacing exceptions, governed under RRC Rule 37. Figure 2.1 shows a schematic of how spacing exceptions under Rule 37 work. If a mineral rights owner does not want to participate in the pool, the firm may still drill, but in such a way that they do not extract resources from that plot. Natural gas producing companies seek to limit these kinds of exceptions as it reduces their potential production and their flexibility in planning the well. These spacing exceptions are the most common way to avoid holdout in Texas. There is also a rarely used forced pooling regulation, called the Mineral Interest Pooling Act. It requires firms to prove that all other options in contacting or negotiating have been exhausted, which is costly.

These spacing exceptions can be granted for any land that has not come under lease, whether the owner could not be found or whether the parties could not agree on terms. The spacing exceptions may hamper the firm's ability to produce as much as they otherwise

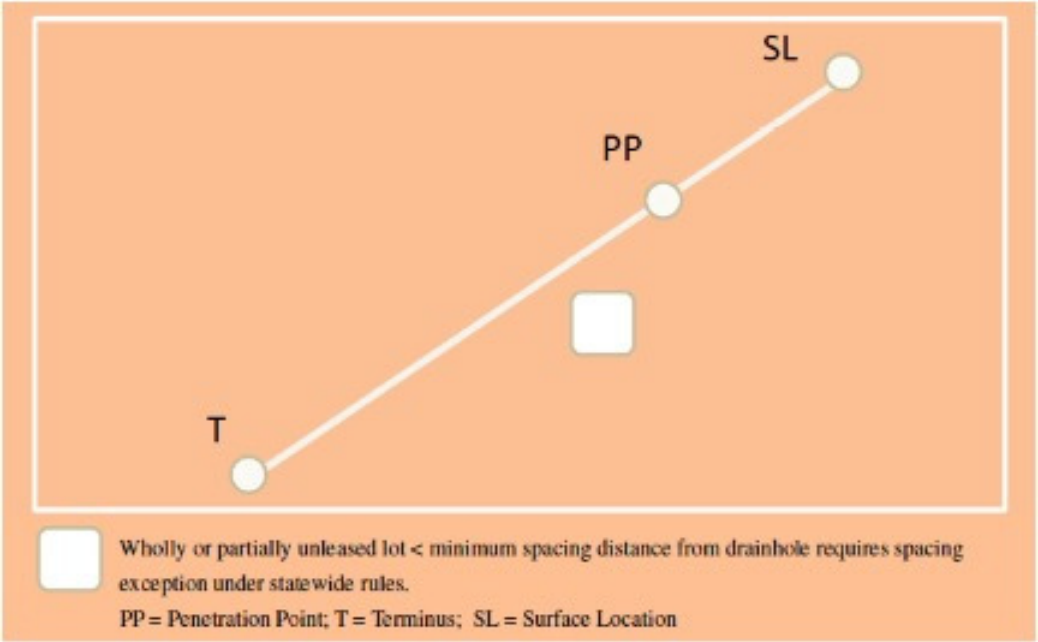

Figure 2.1: Spacing exceptions

would, so they are incentivized to come to terms with the land owners, but not subject them to hold-up scenarios that would extract all rents.

2.3 Data

The Barnett Shale in north-central Texas covers roughly 5000 square miles stretching north, south and west of Fort Worth. Shale is a dense geologic formation with pores filled, in the case of the Barnett, with natural gas. The Barnett Shale ranges in depth from the surface from 5000 feet to 8000 feet, and the attributes which are important for producing natural gas are the thickness of the shale, which ranges from tens to hundreds of feet, the porosity and the thermal maturity. These attributes vary slowly over space, and so, all else equal, proximate wells are likely to have similar production. Drilling Info, a commercial data provider to the oil and gas industry, has developed a grading of the Barnett Shale, that represents the quality of the formation. This can be seen in Figure 2.2, where green areas are the most favorable for producing natural gas and the red are the least favorable.

Geo-coded well production data and lease data also come from DrillingInfo. The lease

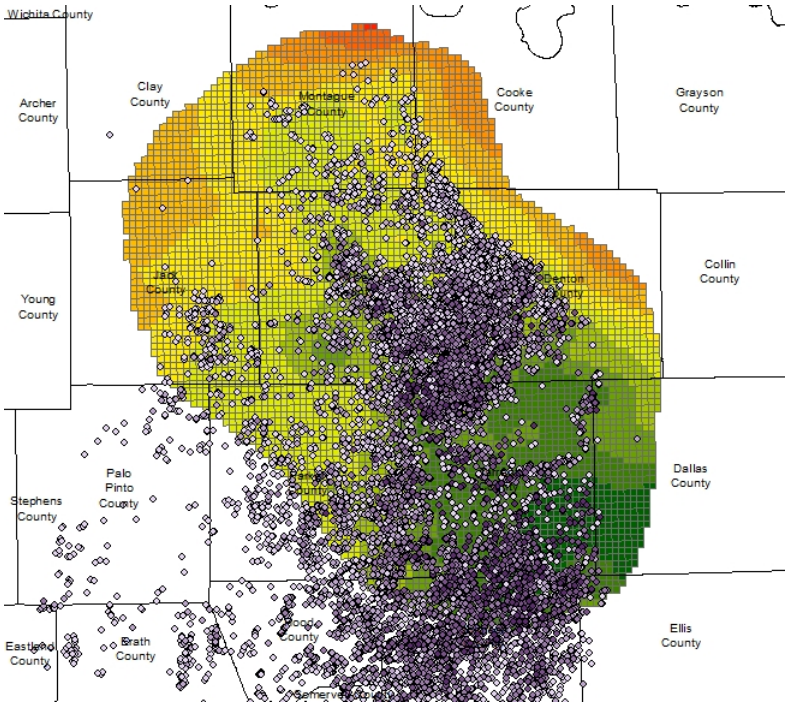

Figure 2.2: Barnett Shale showing Drilling Info acreage grades. Green areas are the most favorable locations and red are the least favorable.

data covers all leases in 8 counties of north, central Texas, which are the core counties of the Barnett Shale, from 1997 to 2014. These cover Fort Worth, Texas and the areas to the north, west and south of Fort Worth. The area covered by each lease vary greatly by county, as Tarrant county is urban and suburban, whereas Hood and Parker counties are largely rural. Figure 2.3 shows the number of leases signed in each county in each year, while Figure 2.4 shows the total acreage leased. Figure 2.5 shows how the median acreage of a lease changes over time in the Barnett Shale, demonstrating that the number of leases required to assemble 20 acres increases considerably as firms lease in increasingly densely populated locations.

Leases are comprised of a royalty, an upfront bonus and a term length. The royalties are a percentage of the gross natural gas produced by the well, prorated by the acreage under lease as the proportion of the entire unit. For example, if a mineral right lease is for an area of one acre of land out of the 20 acres assembled for the permit and the royalty rate is 15%, the mineral rights owner is paid the following:

$$Payment = RoyaltyRate * ProportionOfPermittedLand * WellProduction$$

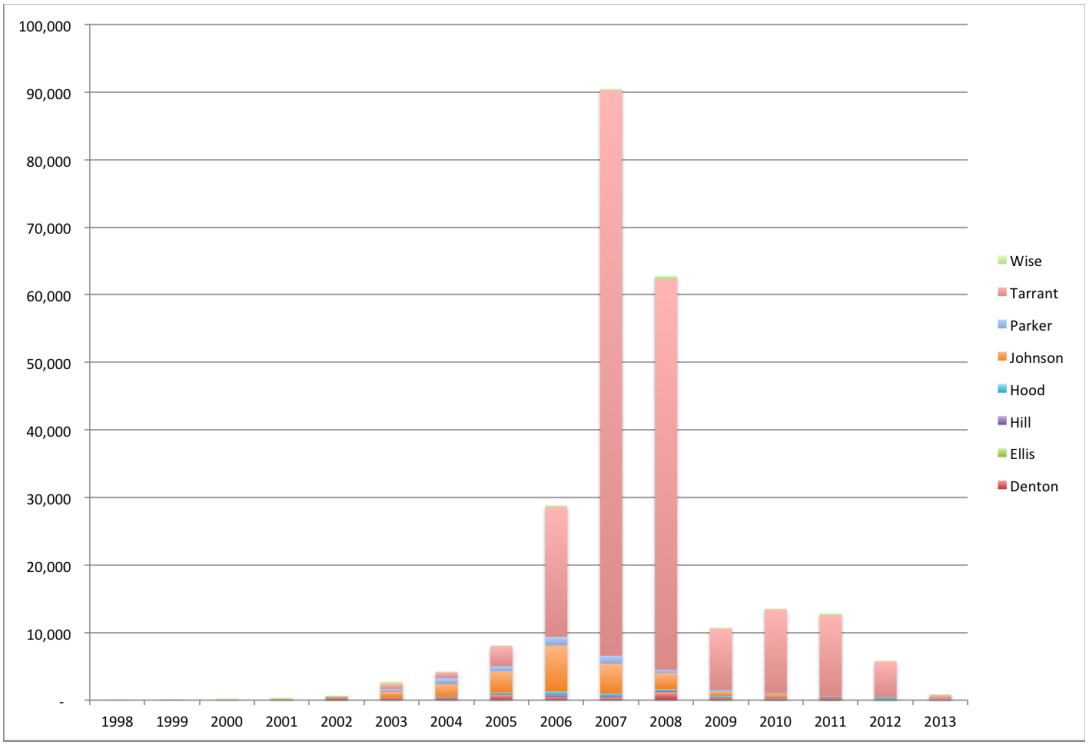

Figure 2.3: Number of leases signed in each year in each county

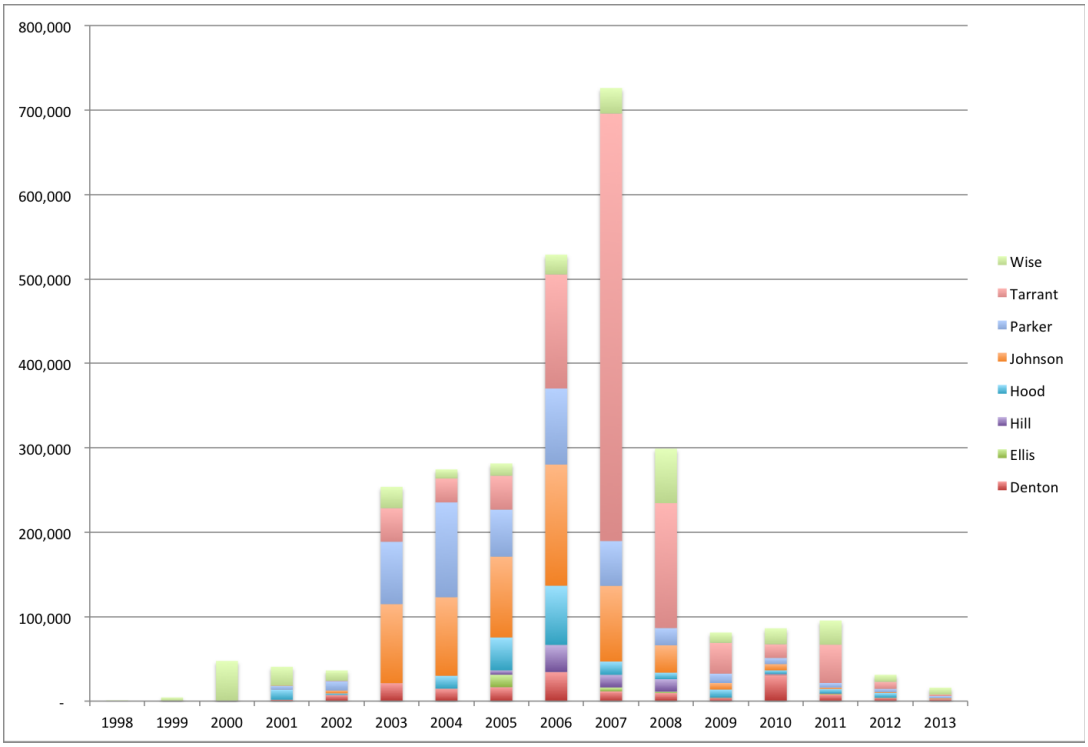

Figure 2.4: Total acreage lease in each year in each county

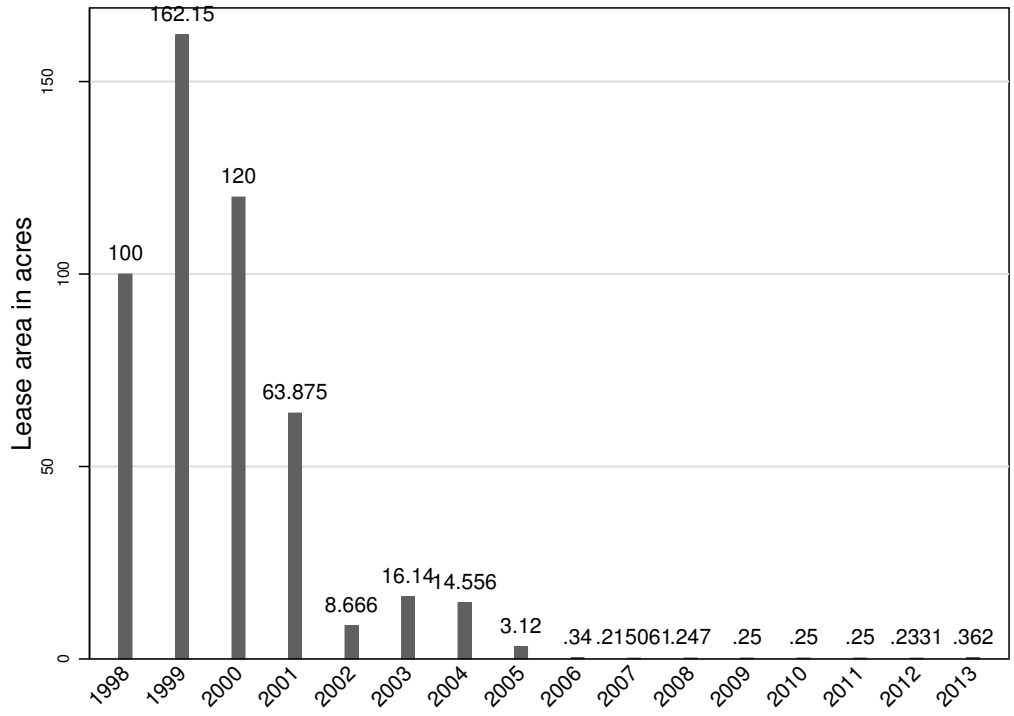

Figure 2.5: Median lease area over time

$$Payment = 15\% * (1/20) * WellProduction$$

Leases may include an up-front signing bonus, as well, but there is no requirement to report these to the regulator, so they are reported for only about 2% of the leases. Finally, the length of the lease stipulates how many months the firm has to start producing from a well on the land. If production starts during that lease term then it will be held by production (HBP) for the lifetime of the well. Once the well has started producing natural gas the firm retains the lease for as long as the well is productive.

Using GIS, I match the leases to operating wells and identifiable pooled units. About 38% of leases can not be matched to operating wells, which means that either wells were never developed on that lease or that my algorithm was unable to match them. Table 2.2 shows the summary statistics of the negotiated terms of both matched and unmatched leases. They

Table 2.2: Summary statistics for leases

	Leases matched to a unit			Leases not matched to a unit		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Royalty, %	145,673	22.83	2.65	92,085	22.94	2.59
Bonus, $	436	4,338.62	8,820.28	272	4,125.11	7,068.99
Lease length, months	145,673	44.32	13.10	92,085	44.70	12.62
Area, acres	145,673	11.28	107.16	92,085	10.71	99.69
Acreage grade	136,821	7.04	0.79	87,676	6.99	0.78

are statistically equivalent on all the dimensions presented in the table. Because wells which are not matched and did not become part of units with a producing well, I can not use those for understanding how they might contribute to holdout. My analysis, then, rests on the wells that are matched and do become part of a unit with a producing well.

2.3.1 Natural gas prices

The natural gas spot prices are from the US Department of Energy's Energy Information Agency. Natural gas prices are quoted at Henry Hub in southern Louisiana, which is the most commonly used natural gas benchmark price in the United States. Figure 2.6 shows Henry Hub natural gas spot prices in $ per MCF (thousand cubic feet) for the period of interest.

Natural gas production from shale follows a very regular pattern of decline. Geophysical models, corroborated by production data, show that the production rate declines as the inverse of the square root of time early in the lifetime of the well, and then shifts to a regime of exponential decline(Patzek et al. (2013)). My own estimation from monthly production data from over 10,000 natural gas wells in the Barnett Shale also shows a clear and highly significant pattern in Figure 2.7.

2.4 Model

I present a model similar in approach to Brooks and Lutz (2016) that presents some testable predictions. Assume that there is some initial endowment of parcels, some sized at $P^* = NP$, a size large enough to be permitted, and some sized at P, which are smaller than the required

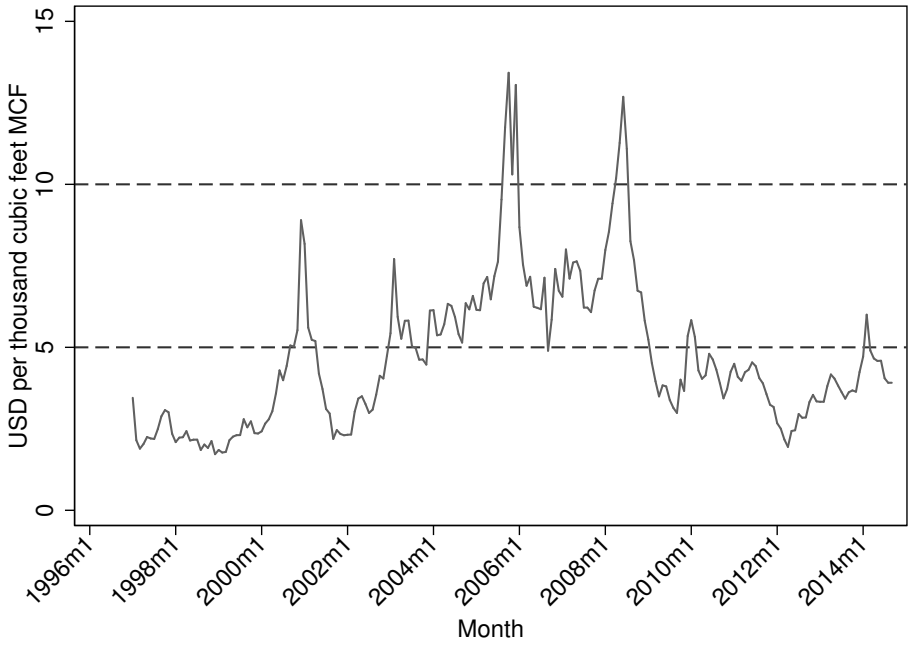

Figure 2.6: Henry Hub natural gas spot prices in $/MCF (thousand cubic feet)

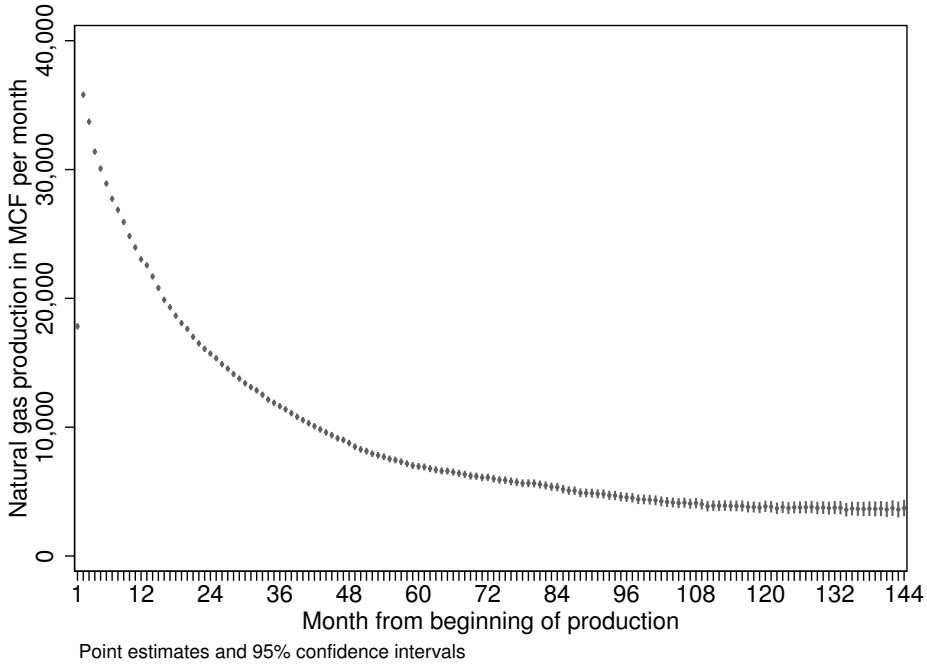

Figure 2.7: Lifetime production curve showing average natural gas produced in each month

size. Equation 2.1, shows that the relationship between the value of the parcels and their size is convex, where $V(x)$ is the value of parcel size x in terms of being able to produce natural gas from it. A firm will assemble N parcels together to get to a parcel of size P^* if the surplus from doing so, is greater than or equal to the surplus from not assembling the parcels. The surplus is shown in equation 2.2, where δ is the ordinary transaction cost for each parcel and r, the royalty paid to the mineral rights owner for a plot of size P. ϕ is some friction that could arise through the tragedy of the anticommons or additional negotiating power that the mineral rights owner gains through holdout, and only occurs when parcels are being assembled.

$$V(P^*) > NV(P) \tag{2.1}$$

$$V(P^*) - N\delta - N\phi - Nr(P) > NV(P) - N\delta - Nr(P) \tag{2.2}$$

The mineral rights owner, on the other hand, is maximizing $r(P)$, the royalty for their parcel, shown in equation 2.3. This is some function of the firm's value of the assembled parcels minus the ordinary transaction cost plus ϕ, which is some function of both the number of parcels needed to make up the critical size, N, and the order in which the parcel enters the pool, n.

$$r(P) = f(\frac{V(P^*)}{N}) - \delta + \phi(N, n) \tag{2.3}$$

$$\frac{V(P^*)}{N} = f(Revenue) = f(AcreageGrade, NaturalGasPrice) \tag{2.4}$$

I assume that δ, the ordinary transaction cost, is constant across leases and I normalize it to zero and that the value to the firm of the assembled parcels is a function of revenue from the well, as seen in equation 2.4. Therefore, the model that I take to the data is seen in Equation 2.5, where u is the pooled unit and n is the nth lease signed that makes up that unit. β_3 is, then, the coefficient of greatest interest, measuring the market friction. I

use as a spot price the Henry Hub price for natural gas in the month that the well starts production. The firms would also know the futures prices of natural gas at the time they start production. These futures contracts are thickly traded up to six months out. I have run the model incorporating futures prices and the results do not change appreciably.

$$R_{un} = \alpha + \beta_1 AcreageGrade_u + \beta_2 NGPrice_u + \beta_3 f(n, N) + \varepsilon \tag{2.5}$$

2.5 Results

2.5.1 The tragedy of the anticommons

In theory, the last land owner to sign a lease should be able to extract all of the rents from the project. This is forestalled in Texas by the possibility of applying for spacing exceptions from the Texas Railroad Commission, as illustrated in 2.1. This section shows results that demonstrate the existence of some frictions associated with a large number of leases making up a pooled unit, despite the mitigating effect of spacing exceptions.

Table 2.3 shows the average royalty of a pooled unit as a function of the number of leases that comprise that unit, controlling for revenue proxies of acreage grade and spot price. Even with year fixed effects, there is a strong positive relationship between the number of leases that make up a unit and the royalty rate paid for the those leases. This suggests that mineral rights owners are able to negotiate higher royalties when there are more leases and more transaction costs for firms.

2.5.2 Holdout

In order to understand, not only if the average royalty is higher for pooled units with more leases, but to see whether there is evidence of holdout in later leases, I order the leases within a unit by the sequence in which the leases were signed. So, the lease with the earliest date is number 1 in the sequence. In Table 2.4, we see that ordering of the leases is positive and significant in some specifications in explaining the negotiated royalty rate. In order to make sure that these results are not driven by the secular trend of higher royalties over time, I

Table 2.3: Average royalty in a pooled unit

VARIABLES	(1) Average royalty %	(2) Average royalty %	(3) Average royalty %	(4) Average royalty %	(5) Average royalty %
# Leases in Unit	0.00463***	0.00194***	0.00203***	0.00117**	0.00116**
	(0.000716)	(0.000607)	(0.000641)	(0.000487)	(0.000487)
Acreage grade				0.616***	0.532***
				(0.0847)	(0.130)
InterX acreage grade, spot price					0.0145
					(0.0197)
Mean Monthly NG Spot Price			-0.0146	-0.00797	-0.0982
			(0.0178)	(0.0156)	(0.127)
Constant	20.88***	16.85***	16.82***	13.90***	14.40***
	(0.197)	(0.605)	(0.677)	(0.680)	(1.055)
Observations	2,199	2,199	2,076	1,953	1,953
R-squared	0.062	0.312	0.313	0.377	0.378
Year FE	N	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the firm level.

introduce year fixed effects in columns (2) - (5). The revenue proxy controls of acreage grade and spot price are added in columns (3) - (5). In 2008-2011, when the median lease size was 0.25 acres, if all of the plots fulfilling the 20 acre regulatory requirement were of that size, it would require 80 leases to make up the unit. This means that the last lease in a pooled unit would have a royalty rate that is 0.018 percentage points higher than the first signatory, following the specification in column (3). This suggests that the spacing exceptions allowed by the Texas Railroad Commission may provide relief from holdout.

2.5.3 Changes in Field Rules

Table 2.1 shows the changes in field rules for the Barnett Shale. We should expect that decreasing the required acreage for optional spacing should reduce the bargaining power of the mineral rights owners because the firms now do not have to assemble as much land. I create two dummy variables, one for all pools that start production after the 1998 change and the 2005 change in field rules where the optional spacing decreases from 80 acres to 40, and 40 acres to 20, respectively. In Table 2.5, we see that the 2005 rule change has significant and negative effect on the average royalty paid in pooled units, which is what we would expect. The effect of the 1998 rule is also negative, but not statistically significant.

Table 2.4: Royalty rising with sequence in unit

VARIABLES	(1) Royalty %	(2) Royalty %	(3) Royalty %	(4) Royalty %	(5) Royalty %
Order in Pooled Unit	0.000990**	0.000234*	0.000231*	6.41e-05	7.37e-05
	(0.000414)	(0.000141)	(0.000139)	(6.14e-05)	(6.39e-05)
Acreage grade				0.569***	0.912***
				(0.0658)	(0.0841)
InterX acreage grade, spot price					-0.0497***
					(0.0123)
Mean Monthly NG Spot Price			0.0263**	-0.00248	0.345***
			(0.0128)	(0.0116)	(0.0827)
Constant	22.53***	15.62***	15.57***	13.92***	12.49***
	(0.267)	(0.000708)	(0.0265)	(0.214)	(0.331)
Observations	129,575	129,575	129,575	121,554	121,554
R-squared	0.020	0.331	0.331	0.356	0.357
Year FE	N	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the firm level.

This probably due to the paucity of data around that time period.

When looking at individual leases, rather than pool averages in Table 2.6, we see that the 1998 field rule has the expected effect, reducing royalty rates. The 2005 field rule's effect changes with specification, but in columns (4) & (5), which include the most complete set of controls the change from requiring 40 acres to 20 acres reduces royalty rates by more than 0.5 percentage points, which is economically meaningful.

2.6 Conclusion

This paper shows that, like Brooks and Lutz (2016), parcels for assembly garner a higher price. We see that average royalty for mineral leases are strongly associated with the number of parcels which make up a unit. This is consistent with the tragedy of the anti-commons in which each mineral rights owner realizes that the firms assembling land have no surplus until they pass a certain threshold of land. In other words, the mineral rights owners understand that the whole is worth more than the sum of its parts and so they all obtain some of that

Table 2.5: Impact of reducing required spacing on royalties

VARIABLES	(1) Average royalty %	(2) Average royalty %	(3) Average royalty %	(4) Average royalty %
# Leases in Unit	0.00470***	0.00470***	0.00242***	0.00242***
	(0.000748)	(0.000737)	(0.000484)	(0.000482)
New field rule 2005		-0.384**	-0.374**	-0.374**
		(0.191)	(0.162)	(0.161)
New field rule 1998		-0.338	-1.242	-1.242
		(0.372)	(0.803)	(0.804)
InterX acreage grade, spot price				-0.000491
				(0.0185)
Mean Monthly NG Spot Price			0.0267*	0.0298
			(0.0143)	(0.120)
Constant	21.10***	19.74***	14.29***	14.27***
	(0.637)	(1.118)	(1.305)	(1.483)
Observations	2,118	2,118	1,953	1,953
R-squared	0.062	0.064	0.250	0.250
Time Trend	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the firm level.

Table 2.6: Effect of reduced optional spacing on individual leases

VARIABLES	(1) Royalty %	(2) Royalty %	(3) Royalty %	(4) Royalty %	(5) Royalty %
Order in Pooled Unit	0.000559**	0.000468**	0.000335*	0.000136	0.000139
	(0.000228)	(0.000198)	(0.000179)	(8.41e-05)	(8.63e-05)
Acreage grade				0.603***	0.677***
				(0.0797)	(0.0922)
New field rule 2005		1.258***	0.0949	-0.521***	-0.543***
		(0.129)	(0.111)	(0.0565)	(0.0584)
New field rule 1998		-0.121	-2.435***	-3.956***	-4.153***
		(0.215)	(0.364)	(0.446)	(0.335)
InterX acreage grade, spot price					-0.0108
					(0.0125)
Mean Monthly NG Spot Price			0.314***	0.295***	0.372***
			(0.0144)	(0.0105)	(0.0855)
Constant	-13.05***	-9.534***	-17.85***	-19.19***	-19.53***
	(0.843)	(1.596)	(1.942)	(1.719)	(1.548)
Observations	129,569	129,569	129,569	121,529	121,529
R-squared	0.237	0.246	0.306	0.332	0.332
Time trend	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Standard errors are clustered at the firm level.

surplus in the form of higher negotiated royalties.

This does not seem to be driven by holdout, however, if we think about that as being the ability of the last participant to extract rents. While we see slightly higher royalty rates for the last leases joining a pool compared with the first, the magnitudes are small and in some specifications statistically insignificant. The spacing exceptions allowed under Texas regulation may help to prevent holdout through the following mechanism. Some mineral rights owners may be reluctant to sign a lease because of the disamenities, such as traffic and industrial activity, associated with a well. However, if their neighbors have all signed leases and the well will be drilled regardless of their participation, they will suffer disamenities whether they sign a lease or not. When the reluctant owners see that they are not preventing the development of a well, they may sign a lease in order to be compensated for the disamenities they would face regardless.

The introduction of new field rules twice during the period of study provide a natural test for both the anticommons and hold-up hypotheses. We see that the reduction in required unit size reduced both the average royalty for a pooled unit and the royalties on individual leases. This is consistent with the reductions in bargaining power by mineral rights owners.

Chapter 3

Innovation, Learning and Spill-Overs in the Shale Revolution

3.1 Introduction & Literature Review

In the 1980's and 1990's, the big US and international oil companies had decided that the continental US was tapped out and that the best new opportunities were overseas. At the same time, a medium-sized, independent natural gas producer in north, central Texas named Mitchell Energy was working to fulfill a contract to supply natural gas to Chicago. Mitchell also saw that their conventional sources of gas were tapering off and so they took notice when in 1982 one of their wells showed the presence of natural gas in an unexpected layer, the Barnett Shale. Shale formations have long been understood by geologists to contain oil and gas, but shale is so dense and with such small pockets of hydrocarbons that is was not given much consideration. Mitchell Energy persevered, however, and eventually developed the techniques that revolutionized the oil and gas industry. This paper tracks the development of an industry which in 2015 represented roughly half of both US natural gas and oil production and added 32% to global reserves of natural gas and 10% to reserves of oil (US Energy Information Administration, 2014). My main thesis is that this was the result of learning by doing. Mitchell Energy and then other firms got better at producing gas from

shale through experience, and I am able to observe this experience from its very beginnings in this relatively quiet corner of Texas oil and gas production. I believe this is the first paper to explore learning by doing in the research and development phase of a new technology and to be able to compare this to learning by doing as the technology matures. It is also one of the first learning by doing papers to give suggestive evidence of how learning is embodied in a firm and the mechanism by which spill-overs occur.

This paper is fundamentally about growth and how it occurs. Two strands of the literature inform this work, one focused on growth and the other dealing with productivity. As North and Thomas (1973) says, "The factors we have listed (innovation, ... education...) are not causes of growth; they *are* growth." (Emphasis in the original.) Innovation is at the heart of the Romer's theory of growth. Romer (1990) defines technological change as the "improvement in the instructions for mixing together raw materials" and posits that it is a result of intentional action taken by market-driven actors. With regards to the process of innovating, Bloom, Schankerman, and Van Reenen (2013) shows that there are significant spill-overs from research and development activity across a wide range of US firms. They estimate that the benefits from spill-overs to other firms are at least twice as large as the benefits that accrue to the innovating firm and that this benefit outweighs the negative effects of increased market rivalry.

Learning by doing is at the center of Arrow's endogenous growth theory (Arrow, 1962) and the literature has examined a number of industries to understand how firms learn from experience Thornton and Thompson (2001) and Levitt, List, and Syverson (2013) look at the manufacture of ships and automobiles, respectively, finding that with experience firms can become more efficient and less error-prone in their own production. Social learning, how a firm benefits from the cumulative experience of others, has been studied most thoroughly in the field of agriculture where Ryan and Gross (1943), Griliches (1957) and Conley and Udry (2010) document the adoption of new technologies or the improved usage of technologies by farmers, when they observe their neighbors doing so. Smieliauskas (2011) also shows a similar phenomenon of healthcare technology adoption amongst the most highly networked

physicians. In the oil and gas industry, Kellogg (2011) shows evidence of joint learning by doing for oil producing firms and drilling rigs working together. Covert (2015) documents the extent to which firms producing oil from shale formations in the US learn from their own experience, but do not incorporate the publicly available information about the experience of other firms as fully as they might to maximize their profits.

To deepen our understanding of the processes of innovation and learning by doing, as well as the how heterogeneity in learning leads to major market share differences, I conduct a detailed examination of the natural gas wells drilled in the Barnett Shale of Texas from 1983 to 2013. Here we can observe the birth of the shale gas industry from the earliest days of research and development of a new technology, through early commercial adoption and into maturity. In the learning-by-doing literature we have observed learning in mature industries, but research and development is usually proprietary. Because all oil and gas wells, whether they are profitable or not, must be reported to the state regulator, the earliest experimental wells are observable. Using geographical information systems (GIS) I am able to match wells with underlying geological characteristics of the shale formation. I then model a flexible production function where a firm's own experience and the experience of the rest of the industry are the primary inputs. I model two outputs using this function - well production conditional on location, and the choice of location itself which is important because of the underlying geological characteristics. Using k-fold validation, I verify that this model fits the data well. I then apply the model to the pre-commercial phase of development of the Barnett Shale (1983-1998), a period described as one of trial-and-error by industry participants[1]. Here, I find that learning in terms of maximizing well production is almost exclusively taking place at Mitchell Energy, the innovative first mover. I do not observe learning, however, in the decision of where to drill.

In the commercial phase, which I define as 1998-2013, I find a different pattern of learning by doing. In maximizing production, I see that social learning is of similar magnitude to learning from the firm's own experience. This differs considerably from the literature on

[1]Personal communication with Herb Magley, geoscience manager at Mitchell Energy during this period.

mature industries (Levitt et al., 2013; Kellogg, 2011) where spillovers are not a significant part of the learning story. In learning about geology quality and location, the commercial phase sees learning from a firm's own experience, especially by firms other than Mitchell. In this case Mitchell does not exhibit learning, but it has higher geology quality grades to begin with. In other words, Mitchell starts off this period with high quality locations and does not improve with experience.

The difference in patterns between learning how to maximize production and in learning to choose favorable locations, especially with regards to spillovers, suggests that different mechanisms are at work. I use data from the business-oriented social networking site LinkedIn were users can post their work histories in order to understand if the movement of workers from Mitchell, the innovative firm, to other firms is connected to higher outcomes at the receiving firms. I find that the move of an employee from Mitchell to another company is strongly linked to higher producing new wells for 12-18 months from the date of hire. This suggests that the learning and knowledge needed for designing a natural gas well is embodied in individuals who can move from one firm to another taking their knowledge with them. On the other hand, there does not seem to be any impact on the geology quality. The learning that takes place in order to assess geology quality and to choose well locations is not as easily transferred with the movement of an individual.

After exploring the effect of learning by doing on natural gas well production and suggesting a mechanism by which learning and spillovers occur, I turn to the effect of learning on the outcomes of firms. Of the 205 firms participating in the Barnett Shale over this period, 63 drilled only one well, most drilled fewer than 200 and 5 firms drilled more than 1000 wells. Can learning by doing explain this heterogeneity? I find that firms that are eventually prolific demonstrate different learning patterns than their nonprolific counterparts. The prolific firms exhibit much higher rates of learning from spillovers in their early wells even though their average production is actually lower than the nonprolific firms. In other words, the ability to learn in the early stages of a firm's participation in the Barnett is associated with whether the firm exits or not. The firms that are able to learn more end up with a

far greater market share. Foster and Rosenzweig (1995) shows that economic activity shifts towards firms with higher productivity. While I can not measure productivity directly, my results seem to suggest that activity shifts towards firms that exhibit higher learning even if their per-well production is initially lower.

This article asks four questions. First, does learning behavior change over different stages of technology development and adoption? Second, are different parts of the natural gas production process subject to different types of learning by doing? Third, in order to begin to understand the mechanisms of learning by doing, where is the learning embodied within the firm? Finally, why are some firms so much more prolific than others and can it be explained by heterogeneity in learning? In order to answer these questions, the paper is organized as follows. In the next section I discuss the institutional context of shale gas extraction, including the exploration and production process. In Section 3, I discuss the data used in my analysis. Section 4 presents the model of learning-by-doing; section 5 discusses results in the pre-commercial and commercial periods. Section 6 discusses a potential mechanism for the important spillovers seen in this context. Section 7 discusses the heterogeneity among firms in terms of performance and learning and section 8 concludes.

3.2 Institutional Context

Shale formations begin to be commercially exploited in the late 1990's with hydraulic fracturing, a technology that had been applied in other contexts since the 1950's. Mitchell Energy, which was later bought by Devon Energy, is credited with developing the techniques which made hydrocarbon production from shale formations feasible (Railroad Commission of Texas, 2016 & Gold, 2015). Mitchell Energy drilled its first well in 1983, and then drilled what is seen by the industry as the first commercially viable shale gas well in the Barnett Shale in 1998. This well used hydraulic slickwater fracturing in a vertical well. Around this time, firms also started to combine hydraulic fracturing with horizontal drilling, which had been developed in the 1990's. As it became clear that natural gas could viably be produced from shale, more firms entered the Barnett Shale in the early 2000's.

3.2.1 Shale gas extraction

Shale is a dense rock, which has long been known to contain hydrocarbons in its pore structure, but tapping these resources was considered economically unfeasible until the late 1990's and early 2000's when slick-water hydraulic fracturing was combined with horizontal drilling. This combination of technologies applied to shale formations led to a 24% increase in natural gas production in the US from 1998 to 2013 and and accounts for one third of oil production in the US.

Hydraulic fracturing or fracking is a technique that has long been used for enhanced recovery of oil and gas in conventional wells and as such is not a new technology, having been used since the 1950s.[2] Hydraulic fracturing involves forcing a mixture of water, sand and chemicals into a well at very high pressures in order to break the rock to release hydrocarbons. Water makes up about 92% of the mixture, sand about 5% and the balance is a mix of chemicals. The sand is forced into the fractures keeping them open so the gas can flow out. The chemical mix includes lubricants and bacteria-inhibiting agents that also facilitate the flow. The application of fracking to shale formations was novel and began in the Barnett Shale in the mid-1990s.

Directional drilling, the precursor to horizontal drilling has long been used to reach hydrocarbons which were not easily accessible by drilling vertically. It was used for oil production in the Austin Chalk and the Bakken Shale by the late 1980's. (King and Morehouse, 1993) Horizontal drilling entails drilling vertically to the desired depth and then turning the drill bit by 90 degrees so that the well continues horizontally for some distance, up to several thousand feet during the mid-2000s. The oil and gas industry gives credit to George Mitchell and his firm Mitchell Energy for developing the techniques in the Barnett Shale that allowed for the exploitation of shale resources, combining hydraulic fracturing and horizontal drilling (Gold, 2015).

The Barnett Shale is part of the Fort Worth Basin of northern Texas, underlying the

[2]Hydraulic fracturing is often referred to as fracking or fracing. Fracking is the preferred spelling in the popular press, whereas fracing is the spelling preferred by the industry. I will use fracking as it is the more common usage.

city of Fort Worth and areas to the north, west and south of the city, covering 18 or more counties, six of which are considered core counties. It was described by geologists by the early 1900s and lies at depths of 7000 to 8000 feet. The narrow Upper Barnett Shale is bounded by the Upper Barnett Limestone and the Barnett Limestone. Below this is the wider Lower Barnett Shale, which is bounded below by the Viola Limestone. The northeastern extent of the Barnett Shale is more than 1000 feet thick and thins to only tens of feet thick in the west (Pollastro et al., 2003).

3.2.2 Exploration and production process

Operating firms, such as oil majors like BP or Chevron or natural gas producers such as Chesapeake, start exploration and production by identifying promising locations. This can be done through a combination of performing geological studies and analyzing well performance of their own previous wells or publicly available well information of other firms. Then the firm must lease the mineral rights of the targeted areas. Leasing may be performed by the operating firm itself, but often firms contract third parties to do the leasing in order to mask their interest in the area. Once the firm has amassed at least 20 contiguous acres, the area required by the Texas Railroad Commission, the oil and gas regulator in Texas despite its name, for gas wells in the Barnett Shale, the firm applies for a permit to drill a well. This permit includes a plat showing the location of the proposed well, detailing the depth and length of the well and naming the target formation. Once approved these permits become public information through the Texas Railroad Commission website.

The operating firm's engineers and geologists design the well, but drilling is typically performed by firms that specialize in owning and operating drilling rigs. There is very low market concentration amongst drilling firms and these firms are typically contracted on a well-by-well basis (Kellogg, 2011). Hydraulic fracturing is often performed by yet another party. Service firms, which supply the equipment and perform the fracking, are firms like Halliburton, Baker Hughes or Schlumberger. This industry segment is quite concentrated with a few large service firms working with many different operators.

3.3 Data

3.3.1 Well data & firm characteristics

Once a well begins production the state of Texas requires monthly disclosure of production levels in order to determine the severance tax due. This information is publicly available through the Texas Railroad Commission website within about two weeks. The primary data used in this paper are well and permit data from the Texas Railroad Commission gathered by DrillingInfo, a commercial data supplier to the oil and gas industry. The data set includes all natural gas wells drilled in the Barnett Shale spanning the period from 1983 when the first well was drilled to the end of 2013. Figure 3.1 shows the number of new wells starting production in each month during this period. It shows fewer than 10 wells per month up to 2001 and then up to almost 300 wells per month in 2007 and 2008.

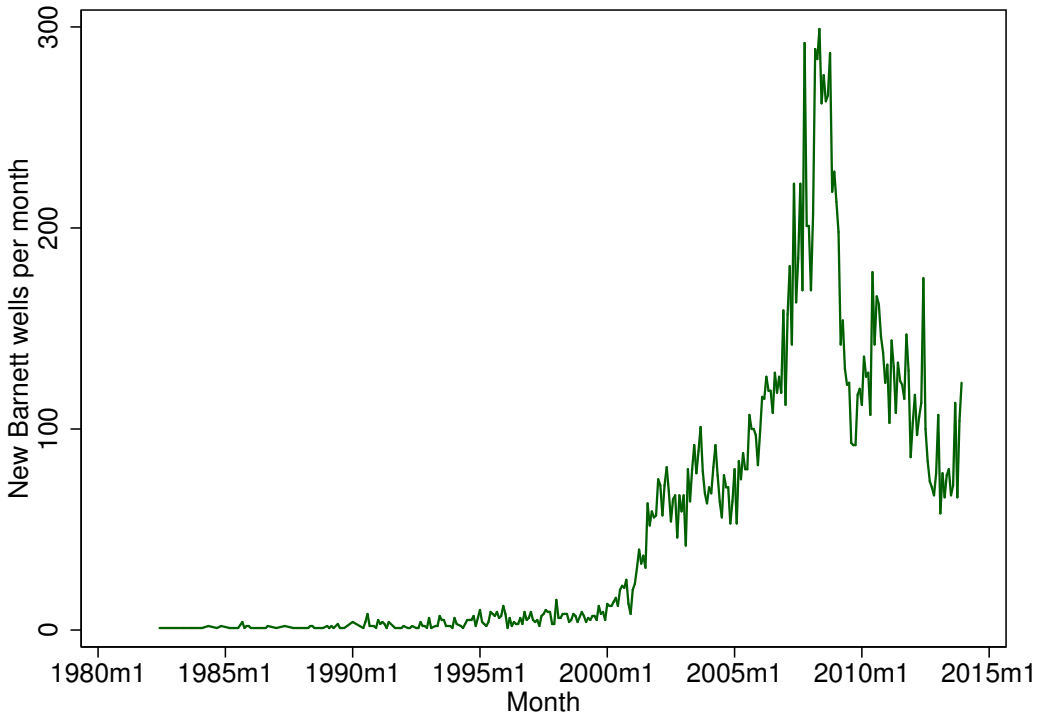

Figure 3.1: New Barnett wells starting production in each month

Most wells in the Barnett Shale are still actively producing at the time of writing, so measures of full cumulative production are not available. I use 24-month cumulative natural

gas production as my primary outcome variable. Production for wells in the Barnett Shale decline at a very regular and predictable rate (Pollastro et al. (2003); Pollastro (2007)), and roughly 80% of a well's life-time production occurs in the first 24 months. Using this measure allows me to include wells drilled to the end of 2013 and it captures the most economically meaningful period of the well's productive life. Figure 3.2 shows the 24-month production for each vintage of wells from 1983 to 2013.

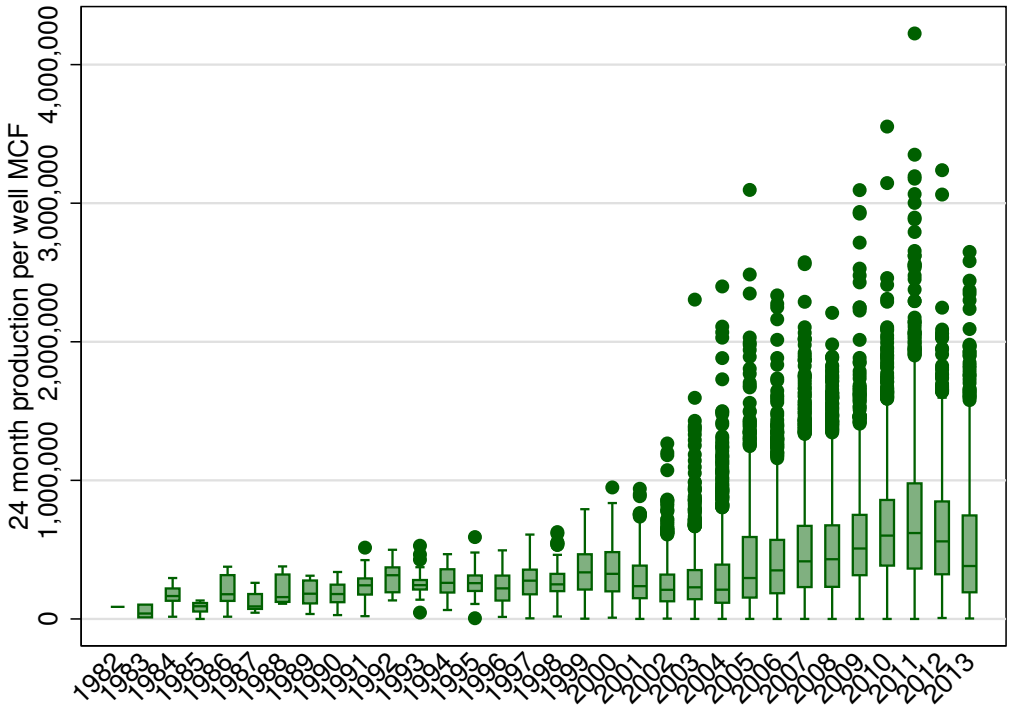

Figure 3.2: Production for natural gas wells starting in each year. The box represents the 75th percentile values, median and 25th percentile. The points show outliers.

I match production data with permit data in order to include the acreage allocated to each producing well. I drop the 3 unmatched wells, leaves a data set of 18,494 wells. In order to account for the acquisitions and mergers of firms, I attribute the wells and experience of an acquired firm to the acquiring firm. The wells in the Barnett are drilled by 205 operating firms. Sixty-three of those firms drilled only one well and eleven firms drilled more than 200 wells. The firms with the most wells are Devon Energy, which acquired Mitchell Energy, with 4,657 wells and Chesapeake Energy with 2,418 wells. Of the 205 firms active in the

Barnett Shale 10 of them are publicly traded. Many of the rest are small, independent oil and gas producers. The large, integrated oil and gas companies, like BP or Shell, are only present when they acquire smaller firms already active in the Barnett, such as the acquisition of XTO Energy by ExxonMobile in 2010.

The structure of the industry with so many different agents interacting at different points during the well drilling and fracking process fosters dispersion of know-how (Covert, 2015). Disclosure of neither the drilling firms nor the service firms are required by state regulation and so neither are available in the data.

3.3.2 Geological data

A number of factors determine the underlying productivity and accessibility of the shale formation. Thickness and thermal maturity determine the productivity, while depth is a major factor in accessibility. Drilling Info has created an index for grading shale acreage for oil and gas production potential.[3] The index is computed using a combination of geological characteristics and production data to eliminate the influence of how the well was drilled and completed to uncover the shale quality (Drilling Info Analytics, 2015). The index was made available to Drilling Info customers in early 2014, which is after my study period, and so could not have directly factored into the decision making process I am examining.

In this area the most productive acreage has a grade of 7 and the lowest least productive, lowest quality acreage has a grade of 2. I use the geocoded location of each well to match it to the acreage grade of the area as shown in figure 3.3 using GIS. Using only the wells with an acreage grade reduces my data set to 14,380 wells. Figure 3.4 shows the acreage grade for wells drilled over the years of the data set. In the early years of activity in the Barnett Shale, most wells were drilled in acreage graded 5, which is in the middle of the acreage grade range. Later years show greater activity in both higher and lower grade acreage.

[3]I initially used an acreage grade index from Rystad Energy, another data service provider in the oil and gas space. While the Rystad index has the advantage of being more simply calculated, I found that the correlation between the index and well production was -0.1 and therefor not a useful control in this setting. The Drilling Info index, on the other hand, has a correlation with production of 0.48.

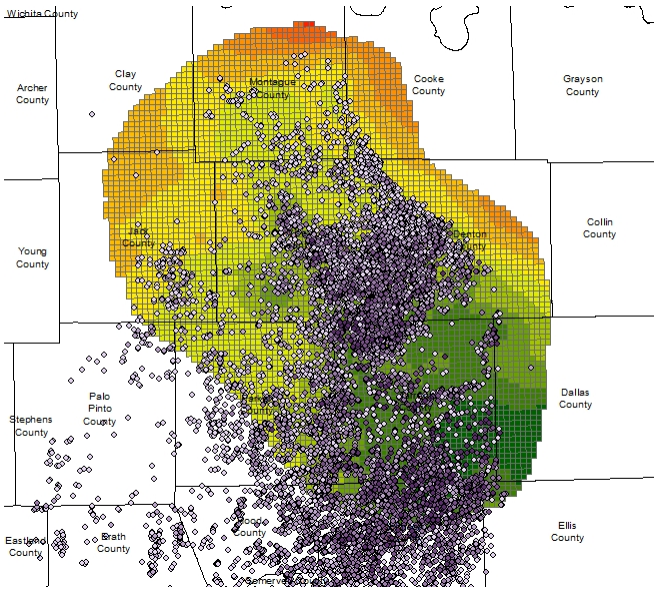

Figure 3.3: Barnett Shale showing natural gas wells and acreage grade scores. Green shows most the favorable locations and red the least favorable. Wells are shown in purple with darker wells having higher production.

3.3.3 Natural gas prices

Natural gas spot and futures prices are from the Energy Information Agency of the US Department of Energy, measured at Henry Hub, as seen in figure 3.5. Natural gas spot prices begin in 1989 and experience highs of over $10 per million Btu in December 2000 and over $13 in October and December 2005 during the supply disruptions following Hurricane Katrina. The price falls to just below $2 in April 2012 as shale gas production expands from the Barnett Shale to the Marcellus in Pennsylvania and the Haysnesville in Texas and Louisiana. The four-month futures prices start in February 1994 and show a similar pattern.

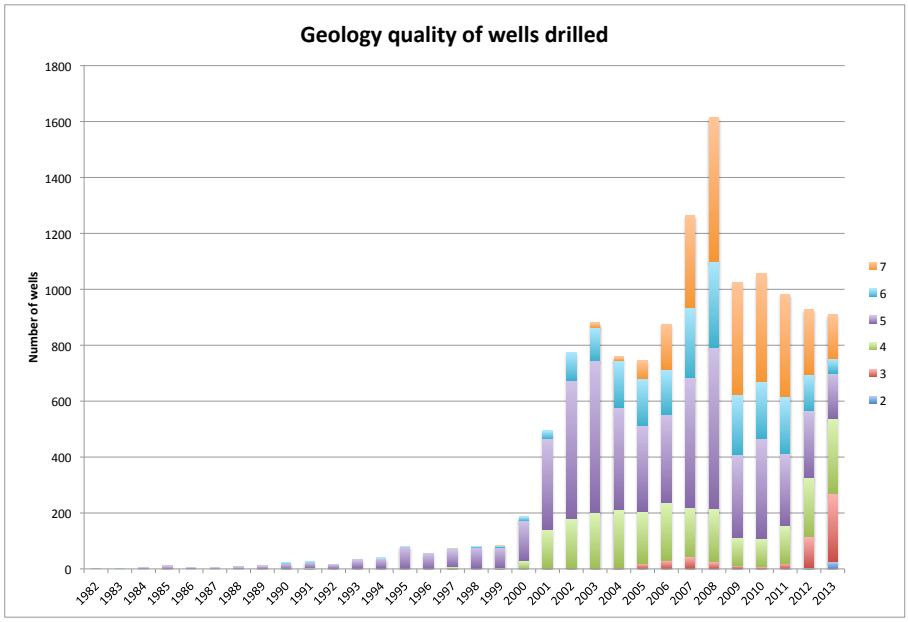

Figure 3.4: Distribution of geology grades over time, with 7 being the most favorable grade in the Barnett Shale and 2 the least.

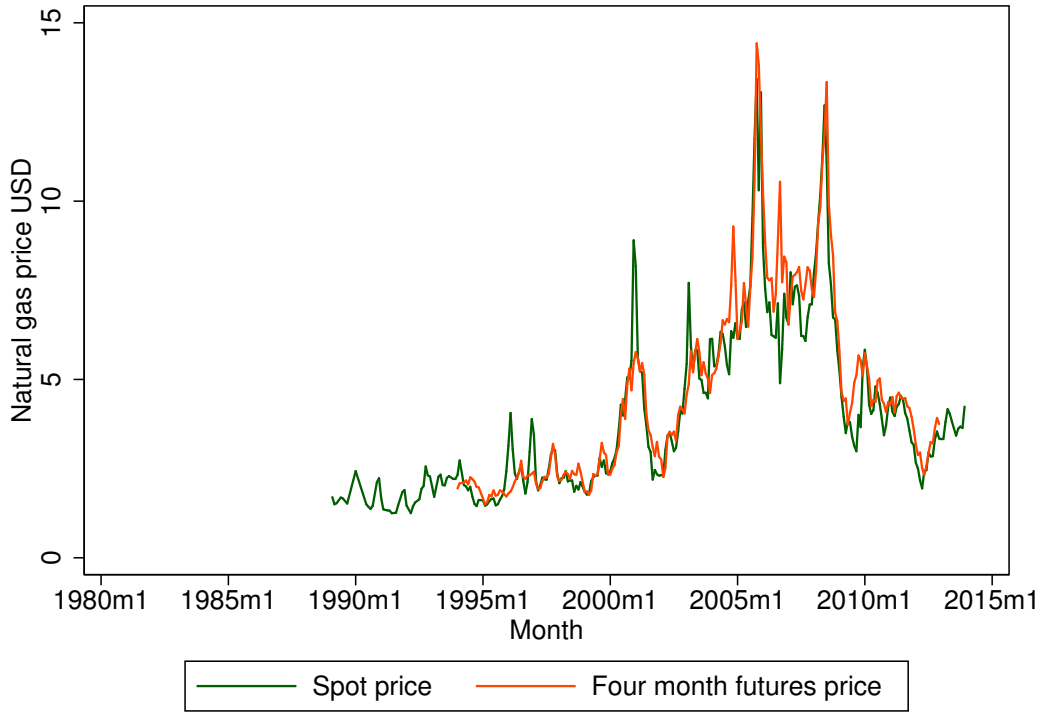

Figure 3.5: Natural gas spot and futures prices per MMBtu

3.4 Model of Learning by Doing

This paper assumes that the operating firm's objective is to maximize natural gas production. The model closely follows Kellogg (2011). Well production, y, is modeled by equation (3.1):

$$y = f(\Phi) \cdot \varpi \cdot \upsilon. \tag{3.1}$$

where Φ represents factors such as the skill and effectiveness of the operating firm, including decisions on how and where to drill the well. The function $f(\Phi)$ transforms these unobservable factors into their effect on the production of the well. ϖ represents the observable geological characteristics of the well which have an impact on production and υ is the disturbance term. Because the firm maximizes production it aims to optimize the skill and decisions which comprise Φ, seeking the $\Phi*$ which maximizes$f(\Phi)$. I model learning-by-doing as some function of experience, both the experience of the firm drilling the well and

the experience of the industry overall in the Barnett Shale, as shown in equation (3.2). E_{-f} denotes industry experience, as the number of wells drilled by all firms other than a firm f's own wells. E_f denotes the past experience of firm f, as the number of wells drilled by that firm and ρ represents all other factors of skill that are unobservable and orthogonal to experience.

$$f(\Phi) = g(E_f, E_{-f}, \rho). \tag{3.2}$$

The literature generally represents learning by doing with a log-log functional form. In Appendix A, I explore whether this fits the data by performing a k-fold validation exercise. I find that a log-log specification is the second best fit by a small margin and I use it here for parsimony and to ease comparison with the literature. Therefore, 3.1 and 3.2 can be combined into the following:

$$y = g(E_f, E_{-f}, \rho) \cdot \varpi \cdot \upsilon. \tag{3.3}$$

Taking logs leads to Equation 3.4, which I take to the data.

$$\log(y_{fti}) = \alpha + \beta_1 \log(E_{fti}) + \beta_2 \log(E_{-ft}) + \delta_f + \gamma X_{fti} + \varepsilon_{fti}. \tag{3.4}$$

y_{fti} represents the production from well i of firm f in month t. Similarly, E_{fti} is the experience of firm f during month t when well i starts producing. In other words, it is a count of all past wells drilled by that firm. Because the data shows only the month of first production, if a firm starts production from more than one well in a calendar month, I randomly assign an order to these wells. The model includes firm fixed effects, δ_f, and a vector of covariates, X_{fti}. The covariates include controls for geology quality, the acreage, and the natural gas futures price. The error term, ϵ_{fti}, is the sum of disturbance terms $\log(\nu)$ and $\log(\rho)$. As in Kellogg, in order to account for forgetting, I only include experience in the two years prior to the start of any well. This model does not include year fixed effects to account for technology change, because what what we are observing in this period is endogenous

technology development and change. For the same reason, although some wells are drilled vertically and some are drilled horizontally, I do not control for the direction of drilling or the length of the lateral extent. These are part of the result of the learning by doing process, rather than inputs to the model.

3.5 Learning Through the Stages of Technology Development

3.5.1 Research and development observed

This data provides us with a unique opportunity to observe the pre-commercial phase of technology development. Mitchell Energy, later bought by Devon Energy, began drilling gas wells in the Barnett Shale in 1983 using a variety of hydraulic fracturing techniques. They originally focused on "gel fracs", hydraulic fracturing using a viscous gel they believed would hold open the pores in the shale. Later they shifted to slickwater fracturing using sand to prop open the fractures, which proved to be much more successful. In addition, they began to attempt directional and horizontal drilling, which were techniques being developed in other contexts starting in the 1990s. This combination of slickwater fracturing, using sand as proppant, with horizontal drilling was to become the technique of choice.

Mitchell Energy and a small handful of other firms experimented with producing gas from the shale formation for many years until 1998, when Mitchell drilled what is considered by the industry to be the first commercially successful well, S.H. Griffin #4 . Because each well drilled in Texas must be registered with the state regulator, regardless of profitability, we have the opportunity to observe the research and development phase of technology development, which in most other industries is confidential and proprietary. Figure 3.6 shows the production of these early wells.

A total of 449 wells were drilled in the Barnett Shale in the pre-commercial period which I define as ending in July 1998 with the start of Mitchell's Griffin #4 well. Mitchell Energy accounts for 86%, or 387, of the wells drilled in the Barnett in this period, but 15 other firms also had some activity. Table 3.1 shows the active firms with the number of wells drilled in

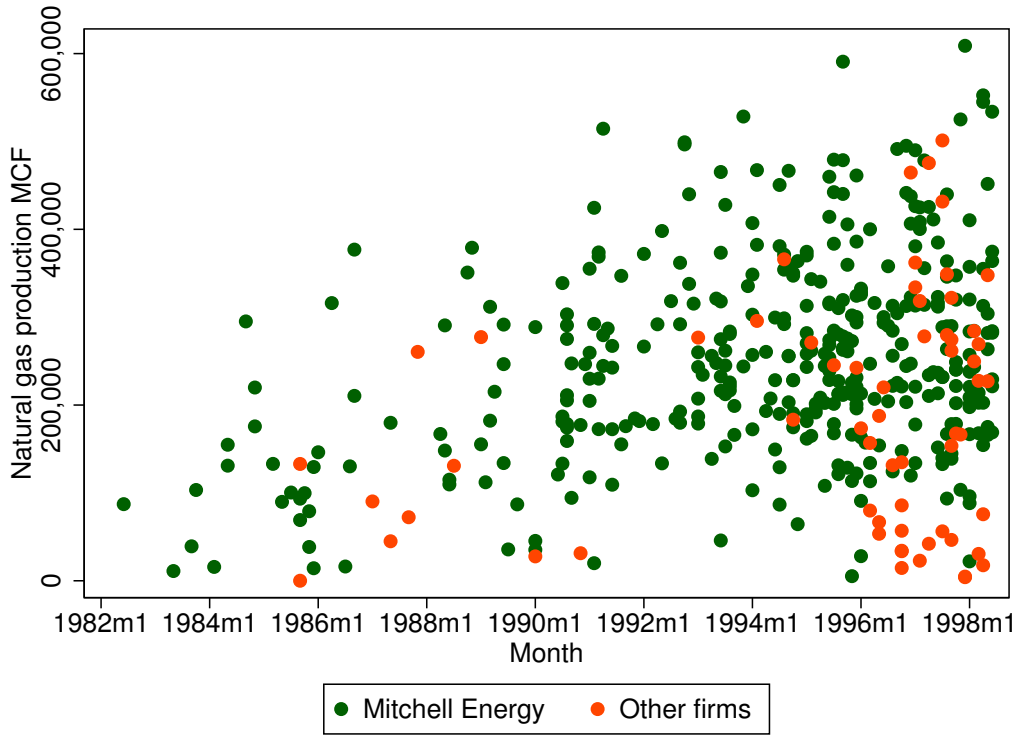

Figure 3.6: Production per well in the pre-commercial period. This figure shows 24-month natural gas production from wells drilled in the Barnett Shale from 1982 to 1998 by Mitchell Energy, the innovator, and other firms.

this period, their mean 24 month cumulative production and the average acreage grade of those wells. The higher acreage grade numbers indicated higher quality acreage.

Table 3.2 shows learning by doing in the precommercial phase of the Barnett Shale's development. The coefficient on firm experience corresponds to β_1 in Equation 3.4 and can be thought of as the learning due to a firm's own experience. The coefficient on industry experience corresponds to β_2 in Equation 3.4 and can be thought of as spill-overs in learning from other firms. Column (1) of Table 3.2, which includes all 16 firms, shows positive coefficients on both the firm's learning and social learning, but they are not significant. By contrast, in column (2), which includes only Mitchell Energy, we see positive and significant learning taking place at the firm level, without spillovers. This means that Mitchell is improving the production of their own wells with experience, but they are not learning from the experience of others in the industry. Column (3) looks at all of the firms except Mitchell. Here, surprisingly, we see a negative (albeit insignificant) coefficient on firm learning, which

Table 3.1: Firm characteristics, pre-commercial period including the number of wells drilled by each firm in the period 1982 to 1998, the mean 24-month production of those wells and the mean acreage grade for those wells.

Firm	N Number of wells	Mean Production (MCF)	Mean Acreage grade
BENGAL GAS TRANSMISSION CO.	1	14,555	5
BURLINGTON RESOURCES	2	36,162	5
DEVON / MITCHELL	387	252,571	5
ENCANA OIL GAS	1	131,527	5
ENERVEST OPERATING, L.L.C.	11	220,386	5
HESS, JERRY OPERATING CO.	2	51,380	7
J-W OPERATING CO.	21	156,586	5
L A PRODUCTIONS	1	27,748	5
LAKOTA ENERGY, LTD.	12	268,693	5
QUICKSILVER RESOURCES INC.	1	42,141	5
RYDER SCOTT MANAGEMENT, LLC	2	67,630	5
SAUDER MANAGEMENT COMPANY	1	130,876	5
SCOUT ENERGY MANAGEMENT LLC	1	132,952	5
TEXXOL OPERATING COMPANY, INC.	3	269,629	5
VICTORY EAGLE UTILITY SV Y, INC.	2	296,390	6
WESTERN PRODUCTION COMPANY	1	84,744	6
Total	449	242,669	5

suggests that one's own experience is detrimental to well production. There is a positive coefficient on spill-overs however suggesting that firms are learning from the experience of others. Combined, these suggest that learning in the pre-commercial period is happening primarily by Mitchell Energy. Other firms may be benefiting somewhat from Mitchell's experience.

Table 3.2: Pre-commercial learning by doing

VARIABLES	(1) All firms recent	(2) Mitchell/Devon only	(3) Excl. Mitchell/Devon
Log firm experience	0.181***	0.238***	-0.257
	(0.0560)	(0.0521)	(0.309)
Log industry experience	-0.0234	-0.0440	0.564
	(0.0350)	(0.0312)	(1.278)
Constant	11.78***	11.58***	9.520
	(0.224)	(0.218)	(5.820)
Observations	410	366	44
R-squared	0.039	0.073	0.021
Number of Opid	9	1	8
Firm FE	Y	Y	Y
Geology controls	Y	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Outcome is the log of 24 month production MCF.

Table 3.3: Learning by doing in the commercial phase

VARIABLES	(1) All firms recent	(2) Mitchell/Devon only	(3) Excl. Mitchell/Devon
Log firm experience	0.0198	-0.446***	0.0141
	(0.0130)	(0.0387)	(0.0150)
Log industry experience	0.178***	0.305***	0.307***
	(0.0107)	(0.0161)	(0.0213)
Constant	9.674***	13.67***	8.631***
	(0.186)	(0.406)	(0.227)
Observations	13,519	3,978	9,541
R-squared	0.176	0.153	0.204
Number of Opid	96	1	95
Firm FE	Y	Y	Y
Geology controls	Y	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Outcome is the log of 24 month production MCF.

3.5.2 Commercial production phase

In the period after Mitchell's Griffin #4 well, which I call the commercial phase, the number of wells and number of participating firms increase considerably. Table 3.3 shows learning by doing for the commercial period of 1998 to 2013. Column (1) shows all firms in the Barnett Shale during the commercial period with more than 1 well, a total of 139 firms. During this time frame firms exhibit learning by doing both through their own experience and through spill-overs. This differs from previous literature where spillovers are small or negligable. Here, we see spillovers on the same order of magnitude as the firm's own learning. Column (2) shows the results if only Mitchell Energy (later Devon), the industry leader, is considered. Unlike in the precommercial period, Mitchell is now the beneficiary of spill-overs from other firms and, surprisingly, its own experience has a negative effect on well production. Column (3) looks at the rest of the industry, where again we see significant spillovers. The coefficients suggest that the bulk of the learning in the Barnett Shale is occurring through spillovers and that a firm's own experience does not contribute to production gains.

3.5.3 Learning about geology quality

The analysis of well production has taken geology quality as given, but choosing the right

Table 3.4: Learning about geology quality in the precommercial and commercial periods

VARIABLES	(1) Precommercial Acreage grade	(2) Precom Mitchell Acreage grade	(3) Precom excl. Mitchell Acreage grade	(4) Commercial Acreage grade	(5) Com Mitchell/Devon Acreage grade	(6) Com excl. Mitchell/Devon Acreage grade
Log firm experience	0.0229	-0.103	0	0.121***	-0.145	0.113***
	(0.0664)	(0.195)	(0)	(0.0179)	(0.215)	(0.0201)
Log industry experience	0.0364	0.0772	0	-0.00599	0.0572	0.440***
	(0.0516)	(0.0722)	(0)	(0.0514)	(0.167)	(0.107)
Log permitted acres	0.0463**	0.0488**	0	-0.0546***	-0.0233*	-0.0749***
	(0.0219)	(0.0240)	(0)	(0.00882)	(0.0138)	(0.0111)
Constant	4.660***	5.086***	5.026	5.286***	5.592***	3.456***
	(0.357)	(0.763)	(0)	(0.305)	(0.626)	(0.682)
Observations	306	268	38	13,059	3,862	9,197
R-squared	0.056	0.061		0.104	0.073	0.124
Number of Opid	5	1	4	95	1	94
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Outcome variable is acreage grade which is a measure of geology quality.

location to drill a well is also a process subject to learning by doing. The model I use is very similar to Equation 3.4, except that the outcome variable is acreage grade and I include year fixed effects to take into account that the choice set changes with time, given that some land has already been leased. Here A_{fti} is the acreage grade for well i drilled by firm f in month t. The year fixed effects are denoted as ϕ_T. Z_{fti} represents the log number of acres on the permit for the well to take into account the possibility that firms may trade want to trade off acreage grade for a larger, more workable area.

$$A_{fti} = \alpha + \beta_1 \log(E_{fti}) + \beta_2 \log(E_{-ft}) + \delta_f + \phi_T + \gamma Z_{fti} + \varepsilon_{fti}. \tag{3.5}$$

In Table 3.4 the independent variable in the acreage grade, reflecting the quality of the shale. Columns (1) - (3) show learning by doing in pre-commercial phase for all firms, for Mitchell and for all others. While both the firm experience and industry experience have positive coefficients neither are significant. Columns (4) - (6) reflect the commercial period of development in the Barnett. Here, we see learning at the learning at the level of the firm, but unlike like learning in well production, there is no evidence of spillovers. We also see that the learning in geology quality is not driven by Mitchell/Devon, but by the rest of the industry. Here, the market leader, does show evidence of learning, although we can see by the value of the constant that their acreage grade was higher on average, but did not improve with experience.

3.6 Potential Mechanisms for Learning by Doing

The learning by doing literature has largely been silent on the mechanisms by which learning occurs. Because spillovers are important in this setting, I examine the potential mechanism of labor movement. I use data gleaned from LinkedIn, a professionally-oriented social networking website where people can post their work histories. LinkedIn does not represent the full universe of labor. In 2016 it has 433 million users, 128 million of which are in the US (Linkedin, 2016). A quarter of internet users are also users of LinkedIn, with demographics higher than average income and education (Pew Research Center, 2015). My dataset is made up of the 163 LinkedIn users who self-report work experience at Mitchell Energy between the years of 1982, when the first Barnett well was drilled, and 2002, when Mitchell was purchased by Devon Energy.

I examine the Mitchell employees who switch companies to join another firm active in the Barnett Shale. Excluding the switches from Mitchell to Devon during the acquisition, but including people who later leave Devon, I test whether the addition of an employee from Mitchell during it's innovative phase shows any correlation with increased well production at the receiving firm. Table 3.5 suggests that the move of a Mitchell employee to another company increases production from that firm's new wells even after taking into account the firm's own experience and industry experience. Column (1) shows that wells drilled by a firm in the first year after receiving a Mitchell employee are more productive. The effect does not persist as strongly after 12 months. This suggests that the learning about how to drill a productive well is embodied in the individuals, that in this case the firm's capital is human capital. This contrasts the results shown in column (2) showing that the acreage grade for wells drilled after the receiving a Mitchell employee does not change. This, and the fact that we don't observe spillovers for geology quality learning, likely stems from the fact that firms develop firm-specific algorithms to determine desirable locations. Especially as firms get bigger, no one single individual is likely to be able to replicate that software when they move to another firm. So, in the case of identifying well locations, the firms learning by doing is embodied in the software developed over time by a number of individuals and is

Table 3.5: Effect of Mitchell employee move on receiving firm

VARIABLES	(1) Log production	(2) Acreage grade
Log firm experience	0.0356***	0.147***
	(0.0126)	(0.0157)
Log industry experience	0.120***	0.0543***
	(0.00815)	(0.0101)
First 12 months	0.220**	-0.145
	(0.101)	(0.127)
12 to 18 months	0.0827	0.115
	(0.123)	(0.154)
Constant	10.03***	4.366***
	(0.179)	(0.0634)
Observations	13,929	14,105
R-squared	0.178	0.025
Number of Opid	97	102
Geology controls	Y	
Firm FE	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

not easily transferrable.

3.7 Firm Heterogeneity

There is considerable firm heterogeneity in the sample. Of the 205 firms that have been active in the Barnett Shale at any time, 63 drilled only one well, 11 have more than 200 wells, 5 of which have more than 1000. What makes some firms so much more prolific than others? Because only 10 of the firms are publicly traded, the capitalization and access to debt of most of these firms are not public information. However, we can observe their initial well performance and their ability to learn. We can see from Table 3.6 that the wells drilled by firms who eventually have a large number are more productive, even if they are located in places with less desirable or equivalent acreage grades. In Table 3.7 we can see that even the first wells of prolific firms look different than those of firms that drill only one well. While first wells have similar acreage grades across firms, figure 3.7 shows the prolific firms are much earlier to begin drilling in the Barnett with a median first well in 2001 compared with 2004 for firms with 2 to 200 wells and 2006 for firms with one well. The firms with only

Table 3.6: Well characteristics by firm size

Number of wells per operator	Number of firms	Cumulative production (MCF)		Acreage grade	
		Mean	Median	Mean	Median
Only 1 well	63	88,375	53,039	5.1	5.0
2 to 200 wells	131	244,502	163,233	5.2	5.0
More than 200 wells	11	505,166	416,607	4.4	5.0
Total	205	473,655	381,302	4.4	5.0

Table 3.7: First well characteristics by firm size

Number of wells per operator	Cumulative production (MCF)		Acreage grade		Year
	Mean	Median	Mean	Median	Median
Only 1 well	88,375	53,039	5.2	5.0	2006
2 to 200 wells	139,394	87,059	5.2	5.0	2004
More than 200 wells	114,466	79,885	5.2	5.0	2001
Total	124,009	79,295	5.2	5.0	2005

one well are less productive in that well than the firms that go on to drill more wells, with median production at 60%-67% of other firms.

Firms that eventually have 200 or more wells in the Barnett Shale, which I'll call prolific firms, exhibit different learning-by-doing behavior than nonprolific firms with fewer than 200 wells. Column (1) of Table 3.8 shows that firms with fewer than 200 wells total in the sample do not exhibit any statistically significant learning from their own experience. They benefit from spill-overs to a greater degree, however, than firms with more than 200 wells and as we see in Table 3.7 their first wells are initially more productive, even when controlling for geology quality. Column (2) shows results for just the first 160 wells of the prolific firms, which is roughly the average total number of wells drilled by the nonprolific group, so that they can be compared more directly with the nonprolific group. When comparing column (2) to column (1) we see that the prolific firms learn from their own experience, whereas the nonprolific firms learn from industry experience.

Prolific firms also show different learning behavior in terms of choosing where to drill. Table 3.9 shows in column (2) that nonprolific firms have a lower average acreage grade and, even with year fixed effects to account for the fact the choice set for land is changing over time, that they seem to be choosing lower quality acreage with greater firm experience. Prolific firms, shown in column (2), have both a higher average and benefit from spillovers in their early wells.

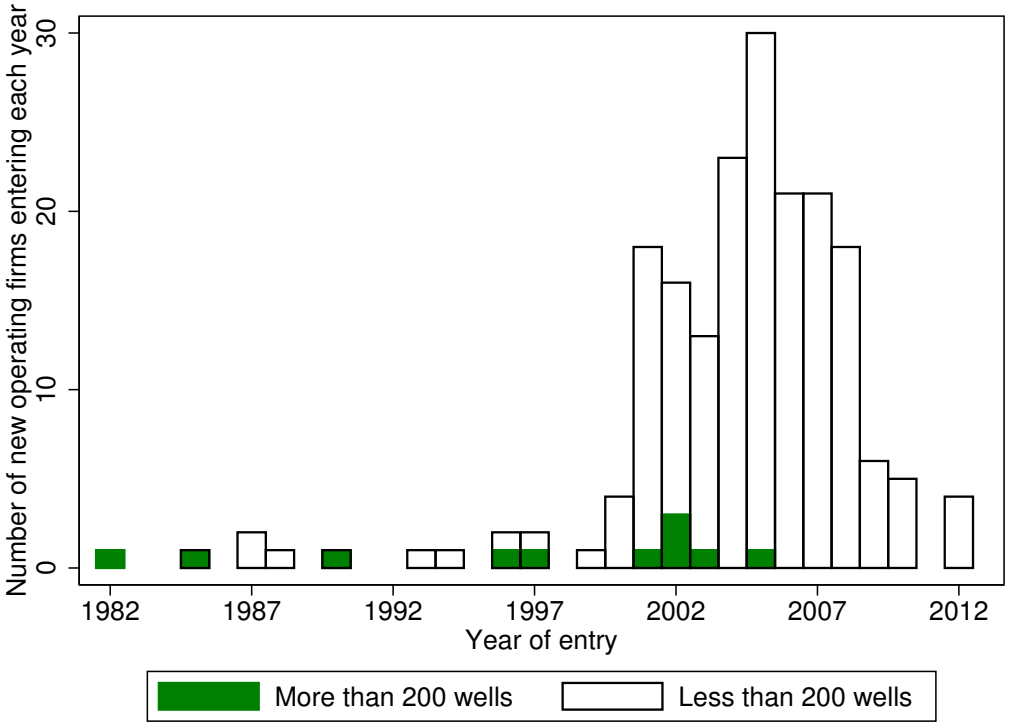

Figure 3.7: Year of firm entry to Barnett Shale, by final number of wells

Table 3.8: Learning by eventual number of wells

	(1)	(2)
		Prolific firms
VARIABLES	Nonprolific firms	Early wells
Log firm experience	0.0486	0.113***
	(0.0322)	(0.0357)
Log industry experience	0.124***	-0.0161
	(0.0436)	(0.0595)
Constant	10.49***	11.30***
	(0.344)	(0.360)
Observations	1,470	1,230
R-squared	0.204	0.145
Number of Opid	86	11
Firm FE	Y	Y
Acreage grade dummies	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Outcome is the log of 24 month production MCF.

3.8 Conclusion

This paper examines learning-by-doing at the birth of an industry with novel technology. I find that in the research and development phase learning takes place within one innovative firm. Later, during the commercial phase of the technology, learning how to make productive wells is largely a function of spill-overs, whereas learning which acreage is most fruitful comes from a firm's own experience and spill-overs are not important. The spill-overs that I find in well production are of much larger magnitude than those found in the rest of the literature. The firms that eventually become large players in the Barnett Shale show different patterns of learning even in their early wells than the firms which drop out. The non-prolific firms do not show learning from their own experience and do not benefit from spill-overs, whereas the prolific firms exhibit learning both from their own experience and the experience of other firms. I explore the mechanism of labor movement and find suggestive evidence that the movement of individuals from the innovative firm to other firms in the industry is correlated with higher well production at the receiving firms for up to 18 months after the move.

Table 3.9: Learning about geology by eventual number of wells

	(1)	(2)
		Prolific firms
VARIABLES	Nonprolific firms	Early wells
Log firm experience	-0.114***	0.0637
	(0.0314)	(0.0444)
Log industry experience	0.348	0.350**
	(0.243)	(0.170)
Constant	3.953***	4.847***
	(1.077)	(1.206)
Observations	1,540	1,231
R-squared	0.128	0.164
Number of Opid	91	11
Firm FE	Y	Y
Year FE	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Outcome is the acreage grade - a measure of geology quality.

Bibliography

Arrow, Kenneth. "The Economic Implications of Learning by Doing." *The Review of Economic Studies* 29.3 (1962): 155–173.

Ayres, Ian and Siegelman, Peter. "Race and Gender Discrimination in Bargaining for a New Car." *American Economic Review* 85.3 (1995): 304–321.

Becker, Gary S. *The Economics of Discrimination*. Chicago: University of Chicago Press, 1957.

Black, Sandra E and Brainerd, Elizabeth. "Importing Equality - The Impact of Globalization on Gender Discrimination." *Industrial & Labor Relations Review* 57.4 (2004).

Bloom, Nicholas, Schankerman, Mark, and Van Reenen, John. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica* 81.4 (2013): 1347–1393.
URL `http://doi.wiley.com/10.3982/ECTA9466`

Brooks, Leah and Lutz, Byron. "From Today's City to Tomorrow's City: An Empirical Investigation of Urban Land Assembly." *American Economic Journal: Economic Policy* 8.3 (2016): 69–105.
URL `http://pubs.aeaweb.org/doi/10.1257/pol.20130399`

Conley, Timothy G. and Udry, Christopher R. "Learning about a new technology: Pineapple in Ghana." *American Economic Review* 100.1 (2010): 35–69.

Covert, Thomas R. "Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale.", 2015.
URL `http://ssrn.com/abstract=2481321`

Drilling Info Analytics. "Graded Acreage Methodology." Tech. rep., Drilling Info, 2015.

Flabbi, Luca. "Gender discrimination estimation in a search model with matching and bargaining." *International Economic Review* 51.3 (2010): 745–783.

Foster, Andrew D. and Rosenzweig, Mark R. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103.6 (1995): 1176–1209.

Gold, Russell. *The Boom: How Fracking Ignited the American Energy Revolution and Changed the World*. New York: Simon & Schuster, 2015.

Griliches, Zvi. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25.4 (1957): 501–522.

Heller, M A and Eisenberg, R S. "Can patents deter innovation? The anticommons in biomedical research." *Science* 280.5364 (1998): 698–701.

Heller, Michael A. "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets." *Harvard Law Review* 112.3 (1998): 622.

Hornbeck, Richard and Keniston, Daniel. "Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872." *NBER Working Paper* .

Kellogg, Ryan. "Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch." *The Quarterly Journal of Economics* 126.4 (2011): 1961–2004.
URL http://qje.oxfordjournals.org/cgi/doi/10.1093/qje/qjr039

King, Robert F. and Morehouse, David. "Drilling sideways: a review of horizontal well technology and its domestic application." *Gas Energy Review* .April (1993).

Levitt, Steven D., List, John A., and Syverson, Chad. "Toward an Understanding of Learning by Doing Evidence from an Automobile Assembly Plant." *Journal of Political Economy* 121.4 (2013): 643–681.
URL http://www.jstor.org/stable/info/10.1086/671137

Linkedin. "About Us | LinkedIn Newsroom." 2016.
URL https://press.linkedin.com/about-linkedin

Murphy, Kevin M and Welch, Finis. "Empirical age-earnings profiles." *Journal of Labor Economics* 8.2 (1990): 202–229.
URL http://www.jstor.org/stable/2535096

North, Douglass C and Thomas, Robert Paul. *The Rise of the Western World: A New Economic History*. Cambridge: Cambridge University Press, 1973.

Patzek, Tad W, Male, Frank, and Marder, Michael. "Gas production in the Barnett Shale obeys a simple scaling theory." *Proceedings of the National Academy of Sciences of the United States of America* 110.49 (2013): 19731–6.

Pew Research Center. "LinkedIn Demographics." 2015.
URL http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19{_}social-media-update{_}10/

Pollastro, Richard M. "Total petroleum system assessment of undiscovered resources in the giant Barnett Shale continuous (unconventional) gas accumulation, Fort Worth Basin, Texas." *AAPG Bulletin* 91.4 (2007): 551–578.
URL http://pubs.er.usgs.gov/publication/70029918

Pollastro, Richard M., Hill, Ronald J., Jarvie, Daniel M., and Henry, Mitchell E. "Assessing Undiscovered Resources of the Barnett-Paleozoic Total Petroleum System, Bend Arch Fort Worth Basin Province, Texas." 2003.
URL http://www.searchanddiscovery.com/documents/pollastro/

Railroad Commission of Texas. "Barnett Shale Information." 2016.
URL http://www.rrc.state.tx.us/oil-gas/major-oil-gas-formations/barnett-shale-information/

———. "Oil & Gas Division." 2017.
URL http://www.rrc.texas.gov/about-us/organization-activities/divisions-of-the-rrc/oil-gas-division/

Romer, Paul M. "Endogenous Technological Change." *Journal of Political Economy* 98.5 (1990): S71–S102.

Rosenthal, Jean-laurent. "The Development of Irrigation in Provence , 1700-1860 : The French Revolution and Economic Growth." *The Journal of Economic History* 50.3 (1990): 615–638.
URL http://www.jstor.org/stable/2122820

Ryan, Bryce and Gross, Neal C. "The Diffusion of Hybrid Seed Corn In Two Iowa Communities." 1943.

Smieliauskas, Fabrice. "Peer effects and service differentiation in technology adoption : Evidence from physician markets.", 2011.

Thornton, Rebecca Achee and Thompson, Peter. "Learning from experience and learning from others : An exploration of learnin ..." *American Economic Review* 91.5 (2001): 1350–1368.

US Energy Information Administration. "Shale oil and shale gas resources are globally abundant." 2014.
URL http://www.eia.gov/todayinenergy/detail.php?id=14431

Vissing, Ashley. "Private contracts as regulation: A study of private lease negotiations using the Texas natural gas industry." *Agricultural and Resource Economics Review* 44.2 (2015): 120–137.

Woodward, Susan E. "A Study of Closing Costs for FHA Mortgages A Study of Closing Costs for FHA Mortgages." : 270.
URL http://papers.ssrn.com/sol3/papers.cfm?abstract{_}id=1341045

Appendix A

K-fold Validation

The literature has commonly modeled learning-by-doing with a log-log model in experience. To verify if this functional form fits the data, I undertake a k-fold validation exercise in the spirit of Murphy and Welch (1990), which measures out-of-sample fit for each proposed model. The two primary explanatory variables are firm experience, the cumulative count of all wells drilled by a firm up to the well in question, and industry experience, all wells drilled by all other firms in the Barnett Shale up to that point. In order to perform the k-fold validation, I drop all firms with only one well in the Barnett Shale, as they cannot exhibit learning. However, I keep those wells in the count of industry experience. Next, I randomly assign the remaining 142 firms into 28 groups of five firms and one group of two firms. For each proposed specification, I fit the model to all but one group of firms and then estimate the root mean squared error when the model is applied to the excluded, or out-of-sample, group of firms. I execute this process 29 times for each model so that each group of firms is excluded once. The average root mean squared error for each model is then used to determine which model fits the excluded data best. This method of excluding one group of firms at a time avoids the pitfall of over-fitting the data by adding an increasing number of higher order terms.

The dependent variable in all the models is the 24-month production of each well, or its log. The right-hand side variables includes functions of firm experience and functions

Table A.1: Best learning-by-doing models for the pre-commercial period

VARIABLES	(1) Log production	(2) Log production	(3) Log production
Log of industry experience	0.169***		-0.00830
	(0.0399)		(0.0834)
Acreage grade	-0.201*	-0.222*	-0.374**
	(0.115)	(0.115)	(0.150)
Log of firm experience		0.226***	
		(0.0325)	
NG 4 month futures price			0.0730
			(0.139)
Constant	12.71***	12.31***	14.03***
	(0.588)	(0.599)	(0.804)
Observations	436	448	300
R-squared	0.048	0.111	0.023
Number of Opid	16	16	11
Firm FE	Y	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Outcome is the log of 24 month production MCF.

of industry experience. Additional controls tested are firm fixed effects, geology controls, natural gas spot prices, sp_t, and four-month futures prices, fp_t. The error term, ϵ_{fti}, is the sum of disturbance terms $\log(\nu)$ and $\log(\rho)$. As above, let f denote the operating firm, and -f denote the rest of the industry. Let t denote the month the well began production and i denote the number of the well in the sequence of operating firm's experience. In months where an operating firm has two or more wells beginning production in the same month I randomly assign their order, since I only observe the month. The model includes firm fixed effects, δ_f, and geology quality dummies, D_{fti}. The error term, ϵ_{fti}, is the sum of disturbance terms $\log(\nu)$ and $\log(\rho)$.

Table A.1 shows the models found through k-fold validation on the pre-commercial wells that best fit the data with specification (1) exhibiting lowest average root mean squared error and specification (2) with the second lowest. The results suggest that the data in the precommercial period shows collinearity, given that 387 of the 449 wells are drilled by the same firm, Mitchell. This suggests that learning is only occurring within Mitchell and not for the other firms.

Table A.2: Best fit model

Model	Average RMSE
$\log(y_{fti}) = \alpha + \beta_1 \log(E_{ft}) + \beta_2 (\log(E_{ft}))^2 + \beta_3 \log(E_{-ft}) + \delta_f + D_{fti} + fp_t + \varepsilon_{fti}$	0.1301
$\log(y_{fti}) = \alpha + \beta_1 \log(E_{ft}) + \beta_3 \log(E_{-ft}) + \delta_f + D_{fti} + fp_t + \varepsilon_{fti}$	0.1306
$\log(y_{fti}) = \alpha + \beta_1 \log(E_{ft}) + \beta_2 (\log(E_{ft}))^2 + \beta_3 \log(E_{-ft}) + \delta_f + \gamma D_{fti} + fp_t + \varepsilon_{fti}$	0.1312

Table A.2 shows that the model with the smallest average root mean squared error is Equation A.1 below, which is a log-log model including a second order term in log firm experience and a first order term in industry experience.

$$\log(y_{fti}) = \alpha + \beta_1 \log(E_{ft}) + \beta_2 (\log(E_{ft}))^2 + \beta_3 \log(E_{-ft}) + \delta_f + D_{fti} + fp_t + \varepsilon_{fti} \quad \text{(A.1)}$$

I perform the k-fold validation exercise again for the commercial phase of shale production in the Barnett. This reflects all wells drilled after July 1998. Table A.2 shows the three highest performing models that were tested and their corresponding average root mean squared error. Table A.3 shows the same models with specification (1) being the lowest, followed by (2) and (3). While the typical log-log specification generally used in the literature can be seen as specification (2), the model with the best fit includes a second-order term in log-firm experience. The negative coefficient on that term, despite not being significant suggests greater concavity than the standard model.

Table A.3: Best learning-by-doing models

VARIABLES	(1) Log production	(2) Log production	(3) Log production
Log of firm experience	0.177***	0.155***	0.180***
	(0.0294)	(0.0135)	(0.0294)
Sq of log firm exp	-0.00174		-0.00265
	(0.00336)		(0.00335)
Log of industry experience	0.0710***	0.0704***	0.0722***
	(0.0149)	(0.0110)	(0.0148)
Acreage grade			0.251***
			(0.00816)
NG 4 month futures price	-0.00599*	-0.00510*	-0.00597*
	(0.00326)	(0.00308)	(0.00326)
Constant	8.369***	8.849***	9.369***
	(0.798)	(0.783)	(0.165)
Observations	12,161	12,161	12,161
R-squared		0.162	
Number of Opid	136	136	136
Firm FE	Y	Y	Y
Acreage grade dummies	Y	Y	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Outcome is the log of 24 month production MCF.