Spring 5-5-2012

Oil and Gas Production: An Empirical Investigation of the Common Pool

Andrew T. Balthrop

PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this dissertation may be granted by the author or, in his or her absence, the professor under whose direction it was written or, in his or her absence, by the Dean of the Andrew Young School of Policy Studies. Such quoting, copying, or publishing must be solely for scholarly purposes and must not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential gain will not be allowed without written permission of the author.

Signature of Author

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:
Andrew Travis Balthrop
1258 Woodfield Drive
Jackson, MS 39211

The director of the dissertation is:
Dr. Kurt E. Schnier
Andrew Young School of Policy Studies
Georgia State University
P. O. Box 3992
Atlanta, GA 30302-3992

Users of this dissertation not regularly enrolled as students at Georgia State University are required to attest acceptance of the preceding stipulations by signing below. Libraries borrowing this dissertation for the use of their patrons are required to see that each user records here the information requested.

Name of User	Address	Date	Type of use (Examination only or copying)

OIL AND GAS PRODUCTION: AN EMPIRICAL INVESTIGATION OF THE COMMON POOL

BY
ANDREW TRAVIS BALTHROP

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree
of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2012

ACCEPTANCE

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair:	Dr. Kurt E. Schnier
Committee:	Dr. H. Spencer Banzhaf
	Dr. James C. Cox
	Dr. Timothy W. Fitzgerald

Electronic Version Approved:
Mary Beth Walker, Dean
Andrew Young School of Policy Studies
Georgia State University
April 2012

ACKNOWLEDGMENTS

I could not have completed this dissertation without the help, support, and friendship of many people.

The data for this dissertation were provided HPDI Corporation, along with plenty of willing counsel. Without HPDI support, this research would not have been possible. I will long owe a debt of gratitude to Jordan Cole and March Nippes for their friendship, their knowledge of the industry, and their patience.

Shared experience forges strong friendships. I was lucky to be classmates with some great people. I must thank Merlin Hanauer for the many great conversations about empirical methodologies over cheap beer at Manuel's, Zack Hawley for ever being contrarian, and Kelly Wilkin for discussions about "*all this*" while sipping afternoon coffees at Samir's.

For much of the time I was researching, I was interning at the Federal Reserve Bank of Atlanta. I am thankful to all my friends and colleagues there for so much of my intellectual development, among them Chris Cunningham, Anastasios Karantounias, Federico Mandelman, Pedro Silos, and Eric Smith. I was especially lucky to spend a great deal of time working with Julie Hotchkiss and Melinda Pitts–had they not taken men in, I might still be lost.

My second essay benefited from the support of the Property and Environmental Research Center. The summer I spent in Bozeman was one of the most intellectually stimulating summers of my life. PERC is a special place, and I hope to be able to contribute more in the future.

I am deeply thankful for the help of my committee. While in Bozeman, I spent some time discussing economics in Tim Fitzgerald's office; it was a lot of fun, and I am jealous that the students at Montana State can do that year round and I cannot. I owe Jim Cox for his time and for his expansive knowledge. I am also thankful for the mentorship of Spencer Banzhaf; one day I hope to display a portion of the kindness, enthusiasm and curiosity he does so daily.

By far my largest debt of gratitude is to Kurt Schnier. He taught me much of the trade of economics–I hope this dissertation will demonstrate that I managed to learn some of it.

Any faults that remain in this dissertation are, of course, my own.

This Dissertation is Dedicated to my Parents

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

ABSTRACT

OIL AND GAS PRODUCTION: AN EMPIRICAL INVESTIGATION OF THE COMMON POOL

BY

ANDREW TRAVIS BALTHROP

March 2012

Committee Chair: Dr. Kurt E. Schnier

Major Department: Economics

This dissertation focuses on the spatial aspects of oil and natural gas production to investigate the extent and effects of inefficient and unnecessary spatial competition. Because oil and natural gas are migratory, operators can cause hydrocarbon resources to flow from a neighboring property onto his or her own through rapid extraction. This problem is compounded when productive leases are comparatively small, as is the case in Texas.

Following an introduction and literature review, the third chapter takes advantage of a natural experiment to demonstrate how spillovers in production limit total cumulative recovery, and how the assignment of secure property rights can enhance economic outcomes. The chapter examines production from wells in Oklahoma and Texas near the panhandle border. While wells on either side of this line have similar geologies and so should be similarly productive they are exposed to different treatments: Oklahoma has a much higher rate of unitization (a contractual scheme where competing owners hire a common operator and share profits), whereas the unitization rate in Texas is lower. Using regression discontinuity design, I find that Oklahoma wells are produced more slowly early on, and that this results in greater cumulative recovery over the course of a well's life (150% more relative to Texas). These results are robust after controlling for reservoir specific effects, and across parametric, semi-parametric and nonparametric specifications.

The fourth chapter quantifies the degree to which competing owners interfere with each other's production through spatial spillovers. I use a spatial econometric model that controls for spatial autocorrelation and spatial dependence and can therefore identify the spillovers in production. Additionally, by comparing leases owned by competing producers to leases owned by a common producer, I show empirically how securing property rights through common ownership can alleviate the externality in production. A priori, one would expect that when a common producer owns adjacent leases, the producer has the incentive to fully account for how spillovers in production affect neighboring wells. Conversely, when adjacent landowners are in competition to extract the resource, they will not account for the damage rapid production causes at neighboring wells. After controlling for secondary injection I find that this is indeed the case for Slaughter field of West Texas.

The fifth chapter investigates the statistical properties of oil and natural gas production. I find striking evidence that both oil and natural gas production are power-law distributed with the exponent approximately equal to one. This distribution might arise from disequilibrium in production and exploration. Highlighting this distribution is important because it has potential consequences for the political economy of regulation as well as for resource management. For example, because the most productive wells lie in the far-right tail of the distribution, regulation geared to prevent a Deepwater Horizon scale spill need fall on a vanishingly small percent of wells. The distribution also has consequences for management because a company profitability depends disproportionately on how it manages its most productive wells.

The sixth chapter provides a short conclusion.

Chapter I

INTRODUCTION

Oil and natural gas are economically vital commodities accounting for over half of US primary energy consumption according to the Energy Information Administration. This will remain the case for the foreseeable future. While the fact that so much of US energy consumption is based on non-renewable resources is a potential cause for concern, that these non-renewable resources are currently being managed less than optimally most definitely is. This dissertation demonstrates that common pool externalities remain an issue in domestic production, quantifies the effectiveness of policies designed to abate the common pool problem, and characterizes the statistical properties of the distribution of oil and gas lease productivities.

Why the common pool externality in production is important to address is that it prevents optimal resource exploitation along four pathways:

1. The resource is not exploited efficiently in a spatial sense because correlative rights are infringed upon.

2. The resource is not exploited efficiently dynamically because extraction occurs too early relative to the price rule.

3. There is physical inefficiency because rapid extraction damages the reservoir.

4. There is economic inefficiency because too much of the rents from the resource are dissipated in the variable factors of recovery.

When property rights to the resource are secure, it is in the producer's best interest to fully account for the exhaustibility, and manage the resource in a way that is in the best interest of society. The trouble arises when producers do not have exclusive rights to the resource. Oil and gas reservoirs are large and the resources flow freely within them, but leases in the study area are small. It is possible, then, for competing producers to poach each others' hydrocarbons (1).This incentivizes rapid extraction (2), which can actually damage the reservoir (3), and wastes resource value in excess extraction and storage capacity (4).

A straightforward solution to this common pool problem is to unify reservoir management under a single operator, unitization. Unitization secures property rights, maximizes the resource rents, and enhances cumulative recovery. Yet this simple solution has not been universally adopted. In my third chapter, I exploit a natural experiment to quantify the difference in cumulative recovery that results from differential rates of unitization. The difference in recovery is stark: unitization enhances cumulative recovery.

The fourth chapter uses a spatial econometric model to explicitly characterize how the production of nearby but competing producers influences own production. The result is consistent with the predictions of theory. When nearby leases are managed by competitors, operators produce at a higher rate than when they themselves own those nearby leases. Fractionalization of reservoir management thus results in dynamically inefficient production–a race to extract.

The fifth chapter characterizes the statistical properties of the distribution of cumulative oil and gas production for the state of Texas. The preceding chapters are based on averages, but it is not clear beforehand that it makes sense to focus on the average lease. An astonishing amount production is carried out by only a few wells and leases. Because of this managers might not care about what is happening on the average lease. The US Army Corps of Engineers does not decide how high to build their levees based on the average flood. It would be folly to do so because the distribution of floods is so wildly varying. The same may be true for oil and gas managers and the government regulators–there is no point in worrying about the average well because that is not where the profits come from. I find that the average lease is descriptive of the distribution. The numbers produced in the proceeding chapters are therefore economically relevant. Securing rights to the resource *in situ* can result in economically meaningful gains.

Chapter II

LITERATURE REVIEW: THE ECONOMICS OF OIL AND GAS PRODUCTION

Introduction

The literature on oil and natural gas production, to which this dissertation seeks to add, is extensive and covers a great variety of topics. Maximizing the economic value of the resource involves addressing both spatial and dynamic considerations: one must decide where to harvest the migratory resource, and which periods of time to do the harvesting. Throw in other economic agents who may be competing for the resource, and the problem grows in complexity, becoming worthy of the research effort it has received.

In this chapter, I review the state of the literature on petroleum and natural gas recovery as it exists at present. My focus is on the managerial side, the micro-level decisions made by the agents involved directly in production, and the relevant constraints these agents face. As such, I will not touch on the broader macroeconomic literature which looks at how the oil and gas industry affects the broader economy, and is affected by it. However, as will become clear, especially with Hotelling-type models, this distinction can get hazy.

I begin my analysis with the basic theory of non-renewable resource recovery. My initial focus is on the dynamic issues of recovery, then the spatial ones, then a combination of the two approaches. I then review the important empirical findings.

Given the complexity of the theory, even with so many simplifying abstractions, it is perhaps not surprising that the results of empirical tests have been mixed. The final section focuses on a separate branch of the literature concerning the managerial side of oil and natural gas production, and the constraints and incentives the agents face. This section has less to do with a generalizable theory of nonrenewable recovery, instead the focus is on applying core economic principles to oil and gas recovery. I conclude by discussing what this dissertation seeks to contribute.

Theory

Oil and natural gas are finite, exhaustible resources, and are nonrenewable on relevant economic timescales. The resources form in places on the seafloor where dead algae and zooplankton accumulate at a rate faster than they decay. On geologic time scales, sediment is deposited overtop this organic material, later forming rock. As more and more rock accumulates, the building heat and pressure transforms the organic material first into kerogen, then into the more familiar hydrocarbons: oil and natural gas. This process is slow. The petroleum deposits used for spatial analysis in the third chapter were formed during the Permian period, over 250 million years ago.

Given the resources are non-renewable, the first aspect of the theory I examine deals mainly with the opportunity cost of extraction. Resource extracted at present is unavailable for future extraction. What then is the optimal rate of extraction? Another property of oil and natural gas is that they are migratory resources that are subject to spatial competition, which is also examined in this section. I conclude the section by reviewing articles that simultaneously examine both the spatial and dynamic properties of extraction.

Theory of Dynamic Nonrenewable Recovery

The article that is commonly credited with beginning the literature on nonrenewable extraction is Hotelling (1931). Hotelling consider the basic problem of the producer: how to maximize the present value of the future stream of profits given a fixed quantity of exhaustible resource. The solution to this problem is known as Hotelling's rule, and is fundamental in the theory of nonrenewable resource economics:

$$\frac{\dot{p}}{p} = (1 + r). \tag{1}$$

The percentage rate of change of the price of the resource is on the left, the rate of return of the next best asset is on the right. It is assumed the only benefit from holding the nonrenewable resource is that it will appreciate in value–it can be sold tomorrow for more than it was bought for today. If the resource gains value at greater than the rate of return on the next best asset, there is no reason to extract it at present. If on the other hand, the resource is appreciating at less than the rate of the next best asset, then one should extract the whole stock of the resource at present, sell it, and invest in the better asset. In equilibrium, what must be the case is the RHS and LHS of equation 1 are equal.

Hotelling's model is useful in that it demonstrates clearly a central feature of non-renewable resource recovery: producers recognize the opportunity cost of extraction as less extraction tomorrow. It can be shown under the assumptions of the model, that producers will engage in the socially optimal amount of conservation. Hotelling (1931) also considers the case of the monopolist, finding the monopolist will be excessively conservative, although Stiglitz (1976) and Khalatbari (1977) show that this isn't inherently the case.[1] The welfare implications of the

[1] It is the demand elasticity that determines whether a producer will delay production in an effort to raise prices.

discount rate on conservation and future consumption are considered more in depth by Vousden (1973) and Dasgupta and Heal (1979).

Serious criticism on the Hotelling model can be made along two fronts. First, equation 1 means that the price of the resource must be rising exponentially, which has been difficult to find evidence for in the data. Part of the reason that the theoretical predictions have not been validated is that the theory is based on assumptions that do no accurately characterize the oil and natural gas industry. The Hotelling model is based on the assumptions that the stock of the resource is independent of the rate of production, as well as the production of other resources; it is assumed that the stock is known, fixed, and homogeneous, and that production is costless, unconstrained, and that technology is unchanging; it is also assumed that demand for the resource is fixed and known. Well-known papers have extended the Hotelling model in various ways, and demonstrated that the trajectory of price can be shaped much differently with more realistic assumptions about pumping cost, uncertainty (Pindyck 1980), exploration (Pindyck 1978), joint production with natural gas (Pindyck 1982), among other things.

The more trenchant criticism is that equation 1 does not contain a term for production in it. Something must be coordinating individual production so that aggregate production behaves correctly and price rises at the rate of interest, and yet in equilibrium each individual producer is indifferent as to his or her level of production–how is it that in aggregate the individuals produce the just the right amount?[2]

Well-known papers have demonstrated that the trajectory of price can be influenced by pumping cost, uncertainty (Pindyck 1980), exploration (Pindyck 1978), joint production with natural gas (Pindyck 1982), among other things.

[2]Of course, maybe Hotelling is telling us that nothing is coordinating production, and overshooting is exactly what to worry about.

Spatial Aspects of Recovery

The reservoirs in which oil and natural gas are found can be expansive, covering many thousands of acres. On the other hand, the leases from which the hydrocarbons are withdrawn are frequently much smaller, so that no single owner has property rights to the resource *in situ* (while it is in the ground). Why this is a problem is that the oil and gas are not fixed in space, they can be moved. In fact, rapid extraction at neighboring pump sites can actually cause hydrocarbons to flow off the lease. The spatial competition means the oil in the ground is a fugacious resource, more similar to fish and wild game than to other non-renewables such as minerals and old growth forest. The result is that oil and gas deposits must be frequently modeled not as private goods, as with the basic Hotelling models, but as a common pool resource.

The theory of common pool resources was first developed for explaining problems with overfishing. Common pool resources are defined as goods which are *rival*, in that one agent's consumption precludes the consumption of another agent, but *nonexclusive*–other's consumption can't be prevented, except by gobbling up all the good oneself. Gordon (1954) demonstrates this issue with a productive fishing ground. A single fisher would harvest until the marginal product of the last unit of effort equals the marginal cost of expending that effort; however, this can only happen when the fishing ground is exclusive. When the fishing ground is open, the fishers instead harvest until the average product equals marginal cost, driving the profit stemming from the productive ground to zero. Scott (1955) further develops the work of Gordon. The same argument holds for oil, except the productive grounds become reservoirs and the fishing boats become wells.

Brown (1974) analyzes a renewable common pool in a dynamic setting, providing a nice segue to the final theory subsection. Before getting there, however, there is another issue with the common pool that is important to highlight: the lost profit, and why the common pool remains a common pool. Why, given the economic damages, is the common pool not privatized? Demsetz (1967) argues that the reason the commons are not privatized is that (1) the benefits from privatization are low, or (2) the cost of privatization must be great. This critique must be taken seriously. On the other hand, in explaining why common pool externalities persist, one must consider who wins in a common pool. Weitzman (1974) makes this argument elegantly, showing that the economic rents are dissipated in excess capital (or, more generally, in the non-fixed inputs to production). Thus, while the lease owner may lose out from failing to unify the common pool, and the profitability of land is dissipated in the drilling of too many wells on the property, the roustabouts, roughnecks, and drill-rig operators are making out just fine.

Spatial-Dynamic Models of Recovery

In light of Weitzman's findings, it is useful to turn back to Brown (1974). The first issue is that given common pool nature, the producers leave too little resource stock for the future. A manager can alleviate this problem by charging the appropriate fee for a unit of stock. The second issue is that the producers use too much of the variable input to produce too little. A tax on the variable instrument could be used to deal with this congestion externality. The stock, still, for all intents, has no spatial dimension. Most of the non-renewable literature has continued to model it as such. Long (1974) deals with substantively the same problem as Brown, but from the perspective of non-renewable resources. Long (1975) equates the common pool externality with the threat of future resource expropriation, finding that this has the same affect on recovery as a higher rate of discount: extraction proceeds at a higher rate, and the resource is exhausted more quickly.

Levhari and Mirman (1980) retain the assumption that resource stocks are independent of space, but do model the strategic behavior between two agents given a common renewable stock in the presence of dynamic feedback. Lewis and Schmalensee (1980) consider similar strategic behavior on the part of multiple producers from a non-renewable perspective, but without common pool or dynamic feedback. Instead, firms are assumed to be able to commit to initial production plans, which are in Nash equilibrium (Levhari and Mirman (1980) compute dynamic reaction functions so that there is a Nash Equilibrium at each point in time). Reinganum and Stokey (1985) model a common pool with both plan commitment and dynamic feedback. The authors find that the length of contract commitment can be an important avenue for addressing the common pool. As the the period of contractual commitment approaches zero, however, the agents extract all of the resource instantaneously.

The spatial dynamic aspects of extraction have been more realistically dealt with in the groundwater literature, albeit with less emphasis on the strategic interactions of producers. Neumann (1972) develops a single-celled aquifer where the aquifer's response to pumping is not instantaneous. Alley and Schefter (1987), fearing that a single-cell aquifer exaggerates the common pool problem, build a multiple cell model that effectively slows the lateral movement of water. The qualitative results of the multi-cell model are identical to the single cell, since any water savings alters the gradient in favor of neighbors. Brozovic et al. (2006) builds a model that explicitly accounts for the spatial distribution of water users, as well as their history of extraction, finding such assumptions make the resource less public. The authors then compare the model to traditional single-cell aquifers and show that predictions can diverge.

Empirical Findings

This section review empirical findings related to the Hotelling model. The Hotelling framework lends itself to a variety of frameworks for empirical tests, including via structural estimation from restricted cost-functions, asset arbitrage models, time series analyses of market prices, and exploitation of what is known as the Hotelling valuation principle. I examine each of these methods in turn. The results are decidedly mixed, and are based on a variety of nonrenewable resource data.

Halvorsen and Smith (1991) and Chermak and Patrick (2001) employ the restricted cost function approach. One implication of Hotelling's theory is that the value of the resource *in situ* should rise at the rate of interest. In the restricted cost function approach this hypothesis is implemented as a parametric restriction on an estimated cost function. An unrestricted cost function is also estimated, and a Hausman test is used to determine whether the restriction is valid.

In particular, these two papers take advantage of the vertical integration of many resource companies. The identification strategy is based on companies that produce both intermediate and final goods. The intermediate good is extracted from the ground and then some processing takes place at the wellhead before the final good is put in the pipeline and sold. The price of the intermediate transaction is not directly observed, but the final cost is diminishing in intermediate production. That is, for any given level of final production, costs are lower when more of the intermediate input is used up. Both studies observe final and intermediate output. They can then estimate the price of the intermediate good that results in the producer choosing to consume the observed value of the intermediate good. The restricted cost function restricts this intermediate price to rise at the rate of interest.

The weakness of this type of estimation is that errors in estimation of the virtual cost parameter results in potentially erroneous conclusions about Hotelling's rule. Halvoresen and Smith (1991) uses macro data on Canadian metals industry from 1954- 1974 , and reject Hotelling's rule. They admit that macro data may not be appropriate, especially in the presence of heterogeneous deposits. Moreover, the authors recognize that their restriction is based on *ex post* interest rates, when the valid rates are actually *ex ante*. The authors conclude that Hotelling's model is not strictly valid, and the assumptions of complete certainty and perfect arbitrage need to be relaxed. They also note that substitutes for the resource or discovery of new stock (Pindyck, 1978, 1980) would shift the price path, and lead to a rejection of Hotelling's theory.

Chermak and Patrick (2001) employs micro data and finds no evidence against Hotelling's rule. The authors examine 29 natural gas wells from 1987- 1991 to make only 449 observations, which does not give them much power. The authors do take engineering considerations seriously, accounting for pressure constraints and reserves that are dependent on the path of extraction.

Heal and Barrow (1980) uses a model of asset arbitrage in which resources and capital assets are explicitly linked. The authors estimate a reduced form equation of supply and demand for the resource flow which is a function of resource price, income, and return on the resource relative to return on other assets. The authors examine the return on copper, lead, tin and zinc from 1965-1977. Krautkraemer (1998) comments that this sort of model explains mineral price behavior relatively well. Heal and Barrow find that movement in resource prices are related to a change in the rate of interest rather than being a function of interest levels as Hotelling's theory predicts.

Hotelling's theory makes predictions about the value of the resource *in situ*, and not about the resource price itself. So one must tread carefully in interpreting price movement in the context of Hotelling's framework. What cannot be is for price to have a unit root. A unit root occurs when price is regressed on lagged price and the coefficient of lagged price is equal to one. This implies the effect of the past shock never diminishes and the price path never reverts to a stationary trend–it is totally stochastic. For Hotelling's theory to hold there must be some deterministic aspect of price. Lee et al. (2006) runs time series regressions on 11 nonrenewable commodity prices from 1870- 1990. Using a new technique, they allow for two endogenously determined structural breaks. The authors are then able to reject unit roots for all 11 commodities (using the same data that others had used and found unit roots). The authors find that prices are trend stationary until a large shock comes along. Encouragingly, the endogenously determined structural breaks occur at significant points in history (WWI,II and Energy Crisis). Livernois (2008) criticizes the use of long time series since they must be deflated by some price index as follows, presenting new difficulties.

Miller and Upton (1985) demonstrates that, as a consequence of the Hotelling principle, the value of a unit of reserves today depends on current prices and extraction costs regardless of when the reserves are extracted. This is because, following the Hotelling principle, the producer must be indifferent between extracting the resource today and extracting it tomorrow–the present value of extraction must be the same in all periods. Miller and Upton then regress the value of a unit of reserves on the net price of extraction. The coefficient should be equal to one. Costs, which are cumulative in production, decreasing returns to scale, and non-competitive prices are all allowable and would be incorporated into the intercept. Taxes are not explicitly included and so the coefficient may be somewhat less. A weakness is in how Miller and Upton (1985) determines the value of reserves, which is taken as stock value divided by reserves. Ideally, what is needed is the value of the lease. Stocks are a problem because many companies are diversified into other assets besides reserves, and so a correction must be made to the data. After this Miller and Upton (1985) finds that the Hotelling Valuation Principle (HVP) does hold for a sample of natural gas and oil producing companies over the period 1979-1981.

In a theoretical exercise, Davis and Cairns (1999) shows that the HVP coefficient equaling one is an upperbound. This is counterintuitive, since discounted cash flows of reserves in general provide a lower bound. The authors criticize Miller and Upton (1985) for lacking sufficient realism, particularly in regard to engineering and regulatory constraints and the option value of future production. These considerations explain the lower coefficients other studies found on Hotelling Valuation Principle (in the range of 0.5). Berck (1995) criticizes Miller and Upton (1985) for not making allowances for size of reserves. He argues that small scale extraction and large scale extraction are not the same, nor should they be treated as such.

Given the failure of Hotelling models to deal with spatial considerations (which are important in the case of oil and natural gas recovery) it is perhaps not surprising to find mixed results. Nevertheless, in a working paper, Pfeiffer and Lin (2009) are able to empirically identify spillovers in water extraction from the High-Plains Aquifer. Instrumenting for neighbor's production with permitted maximum production and weighting according to Darcy's Law, they find externalities to comprise 2% of extraction. Despite the spatial externalities farmers are found to extract water in a dynamically optimal manner consistent with Hotelling (1931).

Applied Oil and Gas Research

While the last section has focused on nonrenewable extraction for oil and gas recovery, extraction is not the only relevant variable. In addition to that, economic agents and public policy makers have a variety of tools at their disposal to alleviate common-pool problems, and influence extraction profile of gas and oil. This section focuses on research particular to the oil and gas industry.

One possible private response to such a common pool problem in the oil and gas industry is the unitization agreement, whereby leaseholders over a common pool effectively merge their leases and hire a single operator. Production is then allocated among leaseholders according to a pre-arranged allocation mechanism;

Libecap and Smith (1993) and Libecap and Smith (1999) list the characteristics of these successful unitization agreements. Alternatively, producers may buy out their competitors or negotiate pro-rationing agreements that limit production. Libecap and Wiggins (1984) considers the determinants of unitization in five oil fields in Texas and Oklahoma, finding that, barring military occupation, fields with large numbers of producers face significant barriers. Such contractual failure may be partly explained by diverging views on lease productivities between producers (Wiggins and Libecap 1985), or by a heterogeneous resource consisting of two substances, (Libecap and Smith 1999). Regulations in Texas systematically favor small landholders (Libecap and Wiggins 1985), (Libecap 1989b) so that government intervention has been unable to alleviate the problem. When contractual response fails, lease owners will even split individual leases among competing operators to increase inflow onto the property (Yuan 2002)

Much of the failure for private response can be attributed to initially diffuse landholdings (Libecap 1989a). Some of the diffusion may be explained by exploration. Hendricks and Porter, among others, have undertaken intensive study of exploration, bidding, and initial land allocation in the Outer Continental Shelf Auctions of the US Department of the Interior. Of particular interest are Hendricks and Porter (1993), Hendricks and Porter (1996) and Lin (2009). These studies examine the hazard rates for drilling on OCS wildcat tracts to see if firms are more likely to drill when drilling occurs on neighboring tracts, thus explaining initial congestion.

Since oil deposits are spatially correlated, one would expect an informational externality. Hendricks and Porter (1993) describe a war of attrition resulting from this externality, where firms have the incentive to wait for neighbors to undertake the cost of drilling in order to see if the property is worthwhile. The evidence for an informational externality is mixed in Hendricks and Porter (1993), while Lin (2009), having specified what it is to be a neighbor in a different manner, finds against any such externality. One possible explanation for the failure to find results is the importance of future production externalities. Waiting to drill is costly: with each passing day production is lost irrevocably to neighbors.

Conclusion

Such is the present state of research on the oil and natural gas industry. This dissertation contributes in three ways. First, it explicitly quantifies the physical wastage of hydrocarbons as a result of common pool externalities. The possibility of such wastage is discussed extensively by Libecap and Wiggins, but not quantified. Second, the dissertation demonstrates that partial unitization is effective in abating spillovers, whereas there is evidence of a race to extract in non-unitized areas. Finally, the dissertation observes that the distribution of well-productivities is particularly skewed towards the right-tail. This is important to understand because the heavy tails have consequences for management, regulation and statistical analysis.

Chapter III

A REGRESSION DISCONTINUITY APPROACH TO OIL AND NATURAL GAS REGULATION

Introduction

From their first discovery, oil and natural gas have suffered from externalities in production.This stems from the nature of the resource. Oil and natural gas exist jointly in subterranean geological strata, and these strata often span large areas so that they underlie multiple productive leases. The upshot is that no single producer has rights to the entire reservoir. Property rights to the resource are assigned according to the "rule of capture" so legal ownership begins only when the resource is extracted. Thus we have a classic common-pool resource: rival, but non-exclusive. Previous literature has illustrated that secure property rights can be used to minimize production externalities (Libecap and Wiggins 1985) (Wiggins and Libecap 1985). The present research uses a quasi-experimental approach to empirically test these findings. Our results are striking: wells in Oklahoma, where policies encourage unitization, making property rights more secure, produce between 3,360 and 4,217 more barrels of oil than comparable wells in Texas.

Common-pool externalities typically lead to too much extraction and the dissipation of rents with excess capitalization (Gordon 1954) (Scott 1955) (Hardin 1968) (Weitzman 1974) (Brown 1974). Unitization, where competing producers contract a single common operator and share profits, is a potential remedy. Kaffine and Costello (Kaffine and Costello 2011) demonstrate that in the presence of common pool externalities, unitization can lead to the first-best outcome. And, although unitization has been a tool favored mainly by oil and natural gas producers, it can be used to secure a variety of migratory natural resources, including wild game and fish. The Chignik Salmon Cooperative of Alaska, the *shiroebi* shrimp fishery of Toyama Bay, Japan provide salient examples of unitization concepts applied in fisheries (Costello and Deacon 2007). In oil and natural gas production, where cumulative recovery depends on the rate of extraction, a race to extract can be particularly damaging. The common-pool externality results not only in economic inefficiency, but also limits physical recovery.

Texas is widely known to suffer from common-pool externalities due to regulations that are biased in favor of small landowners. Oklahoma, in contrast, has been more effective in securing the property rights of producers (Libecap and Wiggins 1985). By comparing wells in Oklahoma to similar wells in Texas, we find Oklahoma producers are more successful in terms of cumulative oil recovery than their Texas counterparts. This is because Oklahoma producers extract at a slower, more sustainable rate. These results are driven by the difference in regulatory policy.

The paper is organized as follows. In section 2 we provide background on the nature of oil and natural gas production and summarize previous literature. Section 3 describes the relevant policy differences between Oklahoma and Texas. Section 4 describes our regression discontinuity design. In section 5 we detail our unique data set made available by the HPDI Corporation. Section 6 presents results; a summary and conclusion is found in the final section, section 7.

Background

Oil and natural gas *in situ* exist jointly under phenomenal pressures thousands of feet below the surface. The tremendous weight of the overlying sediment forces the oil and natural gas into solution within the pore space of the rock. In "primary recovery," a well is drilled, creating an area of low pressure in the resource bearing strata. It is the expansion of the compressed gas that forces the mixture through the pore space of the rock, toward the volume of low pressure at the well face, then drives the resource up the well to the surface. The compressed gas provides the energy for recovery.

Pressure drawdown and production are the same thing (Nind 1981). Increasing the rate of production means lowering the pressure at the well face. As the reservoir ages the pressure across the reservoir tends toward equilibrium and thus falls toward the pressure of the well face. The pressure at which natural gas leaves solution is called the "bubble point." When the pressure at the well face is below the bubble point, the reservoir pressure may also fall below the bubble point, which is particularly damaging to the reservoir in terms of productivity. The reason is that at the bubble point natural gas leaves solution– it literally bubbles out. "Then, because gas is lighter and travels more quickly than oil, it is expelled first, leading to a too-rapid decline in subsurface pressure per barrel of oil produced. As gas is drawn off in condensate fields, it clogs pore space in the reservoir, permanently trapping large quantities of oil." (Libecap and Wiggins 1985).

In "secondary recovery," producers inject water, carbon dioxide, and other substances, into the resource bearing strata to maintain the field pressure, or to drive the resource to the producing wells. In this stage of recovery, it is often the case that wells formerly used for production are switched over to injection.

In fields with many competing producers, recovery is limited along two pathways. First, extraction proceeds too rapidly, leaving large quantities of oil in the ground. Extraction problems with endogenous stocks have been modeled by (Chermak and Patrick 1995) (Chermak and Patrick 2001). Second, because of the common-pool externality, we can expect too little effort in secondary recovery because the benefits of injection may accrue to neighboring producers (Wiggins and Libecap 1985). Thus failure to address the common-pool externality may result in large losses in potential recovery over the life of the well.

Leaseholders can privately deal with the externality through outright purchase, unitization, or pro-rationing agreements in production (Libecap and Wiggins 1984). Perhaps the most promising recourse for producers is the unitization agreement, whereupon the different leaseholders come together to contract a single operator to produce the field. The profits from production are then shared by the leaseholders according to the terms of the unitization agreement. Not surprisingly, the chances of settlement depend on the concentration of ownership (Libecap and Wiggins 1984), with more concentrated fields quicker to reach an agreement.

Libecap and Wiggins (Libecap and Wiggins 1985) consider unitization under the different regulatory environments of Wyoming, Oklahoma, and Texas. In Wyoming, where drilling is often on federal leases, unitization is encouraged prior to production. Oklahoma has compulsory unitization whereby, when 63% of leaseholders (weighted by acreage on the field) agree to unitize, the field must then be produced as a unit. In Texas, unitization agreements must be unanimous. Libecap and Wiggins find that the institutional arrangements do matter, with a much higher percentage of Wyoming wells unitized than Oklahoma, and, in turn, a higher percentage of Oklahoma wells unitized than in Texas. Unitization rates in Table 1 are consistent with the findings of Libecap and Wiggins. We see from the second column that for a sample of wells within five miles of the Texas-Oklahoma border at 36.5° latitude, the unitization rate of Oklahoma is at 15 %, while Texas has no units.

The difference in unitization across Oklahoma and Texas brings two further questions: (1) why is it that Oklahoma was able to pass legislation making it more easy to unitize than Texas, and (2), is this regulation really welfare improving.

Libecap (Libecap 1989b) discusses the different political histories of Texas and Oklahoma. Relative to federal lands, which had large acreages, Texas and Oklahoma had a large number of small landholders. Oklahoma was, however, able to pass well spacing regulations in 1935 and compulsory pooling regulations in 1941, paving the way for Oklahoma's first unitization law in 1945.[3] The average Texas firm in 1930 was only 63 % the size of the average Oklahoma firm (Libecap 1989a) and so was even more resistant to unitization. Compounding this was the discovery of the East Texas Field resulting in a boom in small producers. Prorationing regulations were implemented on a per well basis, and so did little to rationalize drilling to make it easier for future unitization. Owing to the large number of small firms, prorationing regulations were favored, and no compulsory unitization regulations were ever passed (Libecap 1989a).

One can also dig deeper into the question of why, if unitization is welfare improving, the parties involved cannot contract to solve the problem. Employing Demsetzian logic, property rights are only established when the value of doing so exceeds the cost (Demsetz 1967). Economist have taken it for granted that the costs for establishing rights for the resource *in situ* is large, which is why rights to the resource were given according to the rule of capture in the first place (Lueck 1995). Furthermore, the nature of oil and natural gas recovery, the fact that seismic technology is imperfect and that the extent and nature of the resource can only be know by drilling wells, etc., provides good reason to believe that contracting is costly. Our paper attempts to shed light on the other side–to quantify in physical terms the benefits of unitization. Regression discontinuity design allows us to identify the treatment effect of Oklahoma 's unitization policy, which results in between 3,360 and 4,217 bbls more cumulative oil recovery (relative to Texas), with no significant difference in natural gas recovery. This is relative to a Texas fitted

[3]The first unitization law was repealed as well as challenged in court, the 63 % threshold passed in 1963, see (Libecap 1989a), chapter 6.

production at the border of 1,026 bbls of oil. Wells along the border seem to be slightly less productive than in terms of oil than Texas wells, Kellogg finds that median lifetime production to 8,625 bbls (Kellogg 2010).

Policy

The failure to account for the externalities in production leaves the potential for government intervention. In both states, the production of hydrocarbons is highly regulated. Petroleum production in Texas is overseen by the Texas Railroad Commission; in Oklahoma, by the Oil and Gas Conservation Division. Both agencies use production quotas and royalties, as well as command and control policies. We examine each of these measures in turn.

Command and Control

Command and control type policies include a great variety of regulations.[4] Of these regulations, well spacing regulations and regulations on the inclination of drilling (slant and horizontal drilling) are the most relevant in addressing issues of common pool production.

To prevent outright theft, neither Texas nor Oklahoma allows for slant drilling without special permission. Texas has a statewide spacing rule, which disallows the drilling of wells within 467' of a property line or within 1200' of an existing well. While Oklahoma has no explicit spacing requirement, wells must be located at the center of standard leases, and according to the size of the lease, must be a certain minimum distance from the nearest border. For a standard 40-acre lease,[5] Oklahoma wells must lie at least 330' from the border. Thus, it is not immediately clear which spacing regulations are more onerous. Since we have restricted our data

[4] Texas Administrative Code, Title XVI, Part 1, Chapter 3; Oklahoma Administrative Code, Title 165: Chapter 10

[5] Our definition of standard is as defined in the Texas regulatory code: Texas Administrative Code, Title XVI, Part 1, Chapter 3

sample to be of equal areas on either side of the border, we can sum the number of wells within a given latitude, to get an idea of well density. Oklahoma has a much higher well density. Within two miles of the border, Oklahoma drilled 1104 wells to Texas's 658. Within half a mile of the border, Oklahoma drilled 217 wells to 166 in Texas. Combining the figures translates into a density difference of 0.75 to 1.64 wells per square mile between Oklahoma and Texas. Given the higher density of wells in Oklahoma along the border, we would expect this to have a downward bias on our empirical estimates of the treatment effect of unitization because each well in Oklahoma sweeps out less area. This lends further credence to the positive impact we observe in our analysis.

Well spacing exceptions may be granted to protect correlative rights or to prevent resource waste. In the former case, a producer would be allowed to drill closer to a property line if drilling according to regulation would result in substantial portion of the resource underlying the lease to be captured by neighboring producers. In the latter case, exception may be granted if the oil could not otherwise be recovered. Yet these two goals frequently conflict when production tracts are small, as in the case of Texas. Until the decision *Halbouty vs. Texas Railroad Commission* (1962), small leaseholders were given a greater production allowable, to cover the costs of drilling plus a reasonable profit, even at the expense of neighbors' production (Lowe 2003). The alternative to well spacing exceptions, preferred by most states, is forced pooling.

Taxes and Quotas

Monthly quotas on production are assigned in both states as a percentage of a maximum allowable production for the well. Maximum allowable production is based on the depth of the well, and the lease size, with these allocation guidelines being similar in both Oklahoma and Texas. Texas's maximum allowables are more restrictive for substandard leases, less so for standard leases and larger.[6] Because of the mature stage of development of the fields in the analysis, quotas are not likely to bind.

An important difference between the two states are the royalties on natural gas and oil production. In Oklahoma, these royalties are 7 % of the value of production, for both natural gas and oil. In Texas, natural gas is subject to a royalty of 7.5% the value of production, while oil is subject to a lower royalty of 4.6% the value. While these taxes are not consistent with Pigouvian taxation to address the common pool externalities (Dasgupta and Heal 1979), economic theory predicts royalties slow the rate of extraction (Gamponia and Mendelsohn 1985). Nevertheless, given the endogeneity of the petroleum reserves, the joint nature of production, the presence of common-pool externalities and the relative difference of the royalties, it is not entirely clear the effect that royalties will have in our context.

Policy Analysis via Regression Discontinuity

Oklahoma and Texas both suffer from common-pool production externalities, but have implemented different policy measures to mitigate these damages. Oklahoma, through its emphasis on securing property rights through unitization should have slower rates of extraction than Texas (Long 1975), and because of the physical dynamics of production discussed earlier, greater cumulative recovery. Lueck and Schenewerk (Lueck and Schenewerk 1996) model and simulate how unitization affects the trajectories of extraction. Their model leads to the following hypothesis: (1) production in unitized reservoirs will be tilted toward the future; (2) for a single reservoir the production rate should decline after a unit is formed; (3) recovery should be greater in unitized reservoirs; (4) the productive life of unitized reservoirs should be longer; (5) for nonunitized reservoirs, cumulative recovery should decline

[6] Specifics are available upon request.

in the number of firms. In this paper, we are able to test hypotheses (1), (3) and (4) using production data from Texas

In order to test these hypotheses, it would be ideal to set up an experiment with exposure to the different policies (the experimental treatment) being random across the different wells. The randomization gives a valid counterfactual so we may know what would have occurred in the absence of the policy. Wells that are otherwise identical could be randomly exposed to different property rights schemes in an effort to scientifically identify the impact property rights have on production profiles and cumulative extraction. By comparing wells across Oklahoma and Texas we can identify the impact of unitization. We would not, however, want to compare all wells in Oklahoma to all wells in Texas, as these populations are likely to have different unobserved geological characteristics. Identification of the treatment effect of policy requires wells be comparable along both observed and unobserved variables.

Regression discontinuity design provides the framework to address comparability concerns. Since oil and natural gas production are geologically driven, and geology is spatial, wells that are near one another in space are comparable. By looking at the border region of two states, we may find that wells close together, and therefore similar, are exposed to different policy treatments. For this study, we examine wells in Texas and Oklahoma along 36.5° latitude.

Define the treatment as exposure to Oklahoma's policies. That is, $OK = 1$ for wells lying in Oklahoma, and $OK = 0$ for wells lying in Texas. We can then think of the potential productivity for a particular well, i, whether it is exposed to the Oklahoma policies or to the Texas policies. Letting Y_i being some measure of production, we have,

$$Y_i = \begin{cases} Y_i(0) & \text{if } OK_i = 0 \\ Y_i(1) & \text{if } OK_i = 1. \end{cases}$$

$Y_i(1) - Y_i(0)$ gives the treatment effect of Oklahoma's policy relative to that of Texas for well i. While we can think of the potential outcomes for well i under each treatment, in the real world we cannot simultaneously expose well i to both the Texas and Oklahoma treatments; therefore, we can never actually observe the difference. Instead, we must find the average treatment effect $Y(1) - Y(0)$ over some population of wells. For identification of the average treatment effect (ATE) to be valid, those wells receiving the treatment must be comparable in all respects to those wells that do not receive the treatment, excepting for the dimensions of treatment and outcome.

Regression discontinuity design provides the framework for choosing the population over which to calculate the average treatment effect. Wells that are close together are geologically similar, and should therefore be similarly productive. Let X_i be the coordinates of well i in space, in particular, its latitude. Near the border, where $X = c = 36.5°$ the only difference in the wells is the treatment to which they are exposed: $OK = 1$ when $x \geq c$ and $OK = 0$ when $x < c$. The average treatment effect is then (Lee and Lemieux 2010) (Imbens and Lemieux 2008),

$$ATE = \lim_{\epsilon \downarrow 0} E[Y_i(1)|c \leq X_i < c+\epsilon] - \lim_{\epsilon \uparrow 0} E[Y_i(0)|c \geq X_i > c+\epsilon] = E[Y_i(1) - Y_i(0)|X_i = c].$$

The equation states that by examining wells close enough to the border (within an epsilon) we have identified the average treatment effect. Identification of the ATE is based on the assumption of continuity of the conditional expectations at the border, $E[Y(1)|X = c]$ and $E[Y(0)|X = c]$ (Imbens and Lemieux 2008). Intuitively, this means that moving an Oklahoma well slightly south into Texas while maintaining $OK = 1$ cannot cause any jump in outcomes. The same must hold true for Texas wells exposed to the Texas treatment but moved slightly north into Oklahoma. An alternative way of stating this assumption is that the border at $36.5°$ must be exogenous to the production of hydrocarbons.[7]

Data

The data for our analysis is provided by HPDI Corporation, which collects, compiles, and publishes oil and natural gas production data for 31 US states, 4 Canadian provinces and the Federal offshore areas in the Gulf of Mexico and the Pacific. We limit the analysis to production in the Anadarko basin of Texas and Oklahoma to wells within five miles of the Oklahoma-Texas border and within a longitude of $[-100°, -101.9°]$.[8] Figure 1 displays the well locations within our sample. The sample includes monthly observations of production from 1980-2009. The Anadarko Basin is in a mature state of exploration and development for what the USGS labels "conventional resources", where the natural gas and oil have accumulated in discrete traps (USGS 2011).

Oklahoma production observations are entirely at the well level. Texas production, however, is reported at the lease level for wells classified as producing oil, and at the well level for wells classified as producing natural gas. In order to make the data sets comparable, we transform Texas lease production data into well level data by assigning to the lease the average production per active well, appropriately weighted by the number of wells on the lease. Gas production predominates in the Anadarko basin; over 80% of the wells in both states are classified as natural gas.

Price data are obtained from Haver Analytics. For oil, we use the Cushing, OK spot price in dollars/barrel; for natural gas, we use the EIA computed US Natural

[7] This border at $36.5°$ or $36°30'$ is a relic of the Missouri Compromise of 1820. As such, it was set long before oil was discovered in Texas or Oklahoma.

[8] West of -101.9 longitude wells get sparse, and lie mostly in Oklahoma.

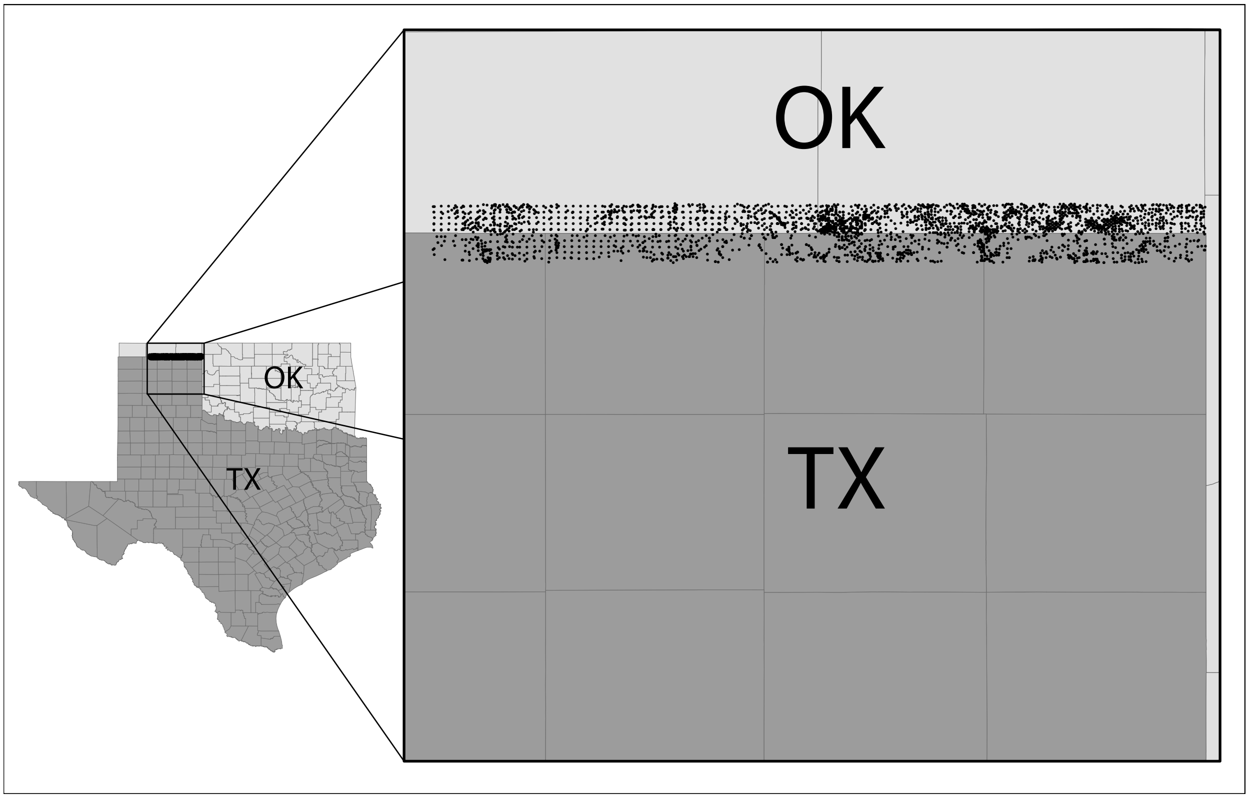


Figure 1: Well locations.

Gas Price in nominal dollars per thousand cubic feet. Both prices are then deflated by the Consumer Price Index.

Table 1 provides means for variables of interest for a variety of samples; standard deviations for each variable are reported below and highlighted in gray. Columns represent four subsamples: the first two columns are wells located within a half-mile of the border for Texas and Oklahoma, respectively; the next two columns are for wells lying within five miles of the border; the fifth and sixth columns are a sample of wells within the first two years of production *and* within five miles of the border; the final two columns are observations of a well's last reported production for wells within five miles of the border.

Table 1: Summary statistics: Texas-Oklahoma border

	Half-mile		5-mile		First production		Last production	
	TX	OK	TX	OK	TX	OK	TX	OK
log gas	6.62	5.15	6.51	6.39	7.36	6.94	3.95	2.63
	2.07	3.15	2.11	2.53	2.61	3.08	2.63	2.76
log oil	2.41	2.41	2.55	1.12	3.54	2.24	1.39	1.85
	2.36	2.75	2.36	2.18	2.92	3.01	1.82	2.31
log monthly revenue	8.09	7.55	8.02	7.81	9.17	8.77	5.53	5.86
	1.31	2.23	1.45	1.71	1.77	2.04	2.18	2.43
well depth (ft)	7315.8	7635.8	7534.0	6295.1	7825.9	7525.3	7846.4	7492.1
	1643.4	1208.3	1678.4	2419.5	1852.5	1646.7	1614.4	1347.0
months produced	122.0	120.0	114.5	138.4	11.2	11.3	106.9	160.3
	89.3	84.9	88.1	96.7	6.9	6.9	87.5	106.0
longitude	-100.77	-100.72	-100.72	-100.91	-100.75	-100.78	-100.71	-100.69
	0.53	0.41	0.49	0.55	0.52	0.57	0.47	0.44
cumulative revenue	13.22	13.26	13.17	13.15	11.78	11.33	12.51	12.55
	1.34	1.82	1.45	1.56	1.89	2.12	1.90	1.99
unitization rate	0	0.36	0	0.15	0	0.10	0	0.18
	0	0.48	0	0.36	0	0.30	0	0.38
log cumulative oil	6.43	8.02	6.99	4.96	9.74	9.35	10.56	9.16
	3.97	3.75	3.68	4.30	2.97	3.59	2.96	4.43
log cumulative gas	11.59	10.61	11.50	11.48	5.90	4.38	6.80	7.13
	2.38	3.36	2.30	3.10	3.63	3.94	3.62	3.61
log gas-oil ratio (gor)	4.21	2.74	3.96	5.28	3.82	4.70	2.56	0.77
	3.49	4.56	3.38	3.93	3.88	4.66	3.09	4.55
observations	12560	12427	132129	189038	13050	10162	493	585

Notes: Sample means for the variables are above and standard deviations are below, highlighted in gray. Four samples are presented: the half-mile sample represents wells within a half-mile of the border; the five-mile sample, wells within five miles of the border. The table contains a sample of wells during their first 2 years of production, and a sample of wells on the last month of production. Log gas it the log of natural gas production per well (units are log(thousand cubic feet)) , log oil is oil production in log(barrels). Log monthly revenues is real revenue from monthly oil and gas production. First seen is the day of first production, units are days withs Jan. 1, 1960=0. Months produced is the number of months a well has been in active production. Longitude is the longitude of the well, in decimal degrees. Cumulatives represents undiscounted sums over the entire production history, and are then logged. Unitization rate is the percent of wells that are unitized. Log gas to oil ratio is monthly gas production divided by monthly oil production (thousands of cubic feet/ barrels) which is then logged. Texas observations represent lease average per well, with averages weighted by wells/lease.

Looking first at the half-mile sample, it is apparent from the means of log gas and log oil production, as well as gas to oil ratio (GOR),[9] that Oklahoma produces more oil per unit of gas than does Texas. Also note that the average age of a well (months produced) is fairly close across the states–it differs by only two months. These statistics are consistent with the idea that Oklahoma wells are produced more sustainably: gas is kept in the reservoir to maintain pressure. Oklahoma also leads Texas in cumulative oil recovery, and trails in cumulative natural gas recovery. Well depth is the total depth of the well (the deepest well in the case of lease-level data). Longitude is the longitudinal coordinate of the well in decimal degrees.

[9]This is gas-oil ratio measured at the well-head: log gas production minus log oil production

Oklahoma's regulations lead us to expect Oklahoma wells to have longer life spans, to produce at a lower rate of production and GOR early on, but to have greater cumulative recovery by the time the well is finally exhausted. To examine these hypotheses, samples of wells both at the beginning and end of life are required. These samples are provided in the last two columns of Table 1. Particularly striking in these cases is the discontinuity in unitization.

For the sample of wells within the first two years of production, we see Oklahoma wells are produced at a lower rate in terms of monthly gas and oil production, although the ratio of these two is slightly higher in Oklahoma. As with the full five-mile sample, the sample of young Oklahoma wells is shallower and more westerly than the young Texas sample.

For the sample of wells in the final month of production, notice that the Oklahoma wells are more than thirty months older than the sample of Texas wells. The sample of exhausted Texas wells are younger on average than the wells in either the half-mile or the five-mile sample. For both states production rates are lower than for the sample of young wells, which is expected. In terms of cumulative recovery, Oklahoma leads in gas, Texas in oil, although the difference in oil recovery is not significant.

Results

Parametric Specifications

Summary statistics in the previous section presage the main findings, but do not tell a causal story in themselves. Graphical analysis indicates that wells become shallower, less productive and more gas-producing as we move to more northerly latitudes. To control for the spatial heterogeneity and to implement the regression discontinuity design we estimate

$$Y_{it} = \alpha + \beta * X_i + \tau * OK_i + \gamma_t + \epsilon_{it} \tag{2}$$

where X_i is a vector of polynomials of distance to the border interacted with the dummy variable for treatment, OK, allowing estimated slope coefficients to differ on either side of the border. The parameter τ gives the average treatment effect of the Oklahoma's policy relative to those of Texas; γ are time-specific fixed effects. Additionally, to control for unobserved heterogeneity in reservoir geology, reservoir fixed effects can be added. To determine the highest order distance polynomial in the specification, we begin with the first-order polynomial, adding higher order polynomials so long as F-tests indicate their joint significance.

Table 2 presents parametric results for nine different dependent variables using the full sample of wells within five miles of the border[10]. Only the coefficients for average treatment effects are shown; their robust standard errors clustered at the lease level are reported in parentheses underneath. Without controlling for reservoir fixed effects, in all cases first order polynomials for distance are the preferred specification. Results indicate Oklahoma wells produce almost 60% less gas and an order of magnitude less oil. Not surprisingly, when output is weighted by their respective prices and summed to arrive at log monthly revenues, Oklahoma wells generate significantly less revenue. As wells age, production tends to decline, and this explains some of the difference in production: the coefficient on months produced indicate that at the border, Oklahoma wells have been producing for 619 more days than wells in Texas. With the wells in Oklahoma almost two years older, it is not surprising that they would be less productive. Yet, despite the Oklahoma

[10]The analysis of this chapter and the next acknowledges that the independent variables are not necessarily independent from each other, *i.e.,* that there is simultaneity in the production of oil and natural gas, which is not explicitly specified.

wells being older, they recover less in terms of cumulative oil (significant at the 10 % level).

Table 2: ATE of Oklahoma policy: full sample

Polynomial Order DEP. VAR.	(1)	(1)	(2)	(3)	(4)
log gas	-0.597**	-1.432***	-0.912	-1.882***	-2.404***
	(0.238)	(0.518)	(0.604)	(0.552)	(0.516)
log oil	-1.137***	-1.024***			
	(0.331)	(0.383)			
log rev	-0.549***	-0.932**			
	(0.162)	(0.369)			
log cum rev	-0.178	-0.183			
	(0.172)	(0.299)			
log cum gas	-0.35	-1.096**	-0.351		
	(0.24)	(0.486)	(0.559)		
log cum oil	-0.941*	-0.010			
	(0.559)	(0.546)			
well depth	-866.322***	-316.849*			
	(250.139)	(168.987)			
longitude	-0.154**	-0.103			
	(0.065)	(0.064)			
produced time	619.538***	1,175.506***			
	(177.667)	(202.353)			
reservoir FEs	No	Yes	Yes	Yes	Yes
observations	321167				

Notes: The columns represent average treatment effects estimates from OLS regression for different orders of polynomial distance interactions. Rows represent different independent variables. Robust standard errors clustered at the lease level are reported in parentheses. Only specifications where the polynomials are jointly significant are reported.

The parametric regressions make possible a specification test. For the regression discontinuity design to be valid samples should be comparable both along observables as well as unobservables. While it is impossible to test for the comparability of unobservables, spatial variables measuring the longitude and depth of the well are observed and, by the same arguments made for the exogeneity of latitude, should be exogenous to treatment. A simple falsification test is to take these other dimensions in space and plug them into the parametric specification to see if there is a discontinuity. The statistically significant estimates for the ATE for both longitude and well depth indicate that Oklahoma and Texas samples are observably different.

Reservoir specific fixed effects do not qualitatively change estimates in terms of monthly production. Controlling for reservoir differences does, however, seem to make the samples slightly more comparable in terms of longitude and well depth, although the treatment effect for well depth remains significant at the 10% level. Controlling for reservoir fixed effects, we no longer observe differences in either cumulative oil or natural gas recovery.

Reservoir dynamics and economic incentives predict that Oklahoma wells should be produced at a lower rate early in the life of the well, and that this should result in enhanced cumulative recovery and a longer well life relative to Texas (Lueck and Schenewerk 1996). By failing to control for well age, it is impossible to determine the impact of policy. Therefore we examine two smaller samples: wells in the first two years of production, and a sample of wells on the final day of production.

Table 3 presents parametric specifications for the sample of young wells. The results are qualitatively similar whether reservoir fixed effects are included. Falsification tests on longitude and well depth come up clean, in that the parameter for ATE is insignificant at conventional levels. Second order terms are needed for log gas and log cumulative gas production, and, when added, result in an insignificant estimate for the average treatment effect. First order distance polynomials are sufficient for log oil, log cumulative oil production, log revenue and log cumulative revenue. While all the coefficients are negative, only the coefficients on monthly and cumulative revenues are significant. This evidence indicates that early in the life of a well Oklahoma wells are produced more slowly in the joint terms of oil and natural gas and price. It may be the case that Texas wells are completed more pro-cyclically, so that they take advantage of high prices early in the life of the well, when it is the most productive.

Table 3: ATE of Oklahoma policy: young sample with reservoir fixed effects

Polynomial Order DEP. VAR.	(1)	(2)	(1)	(2)
log gas	-0.911***	-0.317	-1.938***	-0.912
	(0.347)	(0.545)	(0.420)	(0.613)
log oil	-0.515		-0.799	
	(0.433)		(0.545)	
log rev	-0.827***		-1.282***	
	(0.276)		(0.413)	
log cum rev	-0.803***		-1.259***	
	(0.278)		(0.412)	
log cum gas	-0.763*	-0.074	-1.922***	-0.755
	(0.402)	(0.627)	(0.488)	(0.702)
log cum oil	-0.141		-0.598	
	(0.531)		(0.675)	
well depth	232.160		-138.284	
	(209.045)		(163.322)	
completion	-4.810		-91.892	
	(54.610)		(63.979)	
longitude	0.040		-0.075	
	(0.065)		(0.073)	
reservoir FEs	No	No	Yes	Yes
observations	23212			

Notes: The columns represent average treatment effects estimates from OLS regression for different orders of polynomial distance interactions. Rows represent different independent variables. Robust standard errors clustered at the lease level are reported in parentheses. Only specifications where the polynomials are jointly significant are reported.

Table 4 contains findings of the parametric specification for the sample of wells on their final day of reported production. This sample is particularly useful for examining cumulative recovery, both in terms of revenue and quantity. Oklahoma recovers significantly more oil in cumulative terms, with no evident impact on gas recovery. What is perplexing is the difference in cumulative quantities does not show up significantly in terms of cumulative revenues. Again, if Texan operators are particularly successful at drilling and rapidly producing wells when prices are high, there may be no difference in cumulative revenue, while there is in terms of quantity recovered. The sample of old wells is consistent with this story: Oklahoma recovers much more oil, but there is no statistical difference in cumulative revenue. As in the sample of young wells, the sample of old wells passes falsification tests based on exogenous variables: there is no border discontinuity in the depths of wells or in their longitude. Finally, wells appear to be decommissioned at the same age, in contrast to what summary statistics had indicated.

Table 4: ATE of Oklahoma policy: old sample with reservoir fixed effects

Polynomial Order DEP. VAR.	(1)	(2)	(3)	(1)	(2)
log gas	-2.186***			-2.419***	
	(0.456)			(0.601)	
log oil	1.384***			1.089***	
	(0.321)			(0.399)	
log rev	0.734**			0.393	
	(0.319)			(0.426)	
log cum rev	0.261	0.170	0.640	0.401	
	(0.314)	(0.442)	(0.579)	(0.462)	
log cum gas	-0.873			-1.301	
	(0.618)			(0.863)	
log cum oil	2.721***			2.135***	
	(0.649)			(0.713)	
well depth	254.181			-156.054	
	(256.424)			(258.131)	
longitude	0.326***	0.117		0.040	
	(0.075)	(0.111)		(0.074)	
prod. time	795.856**	-354.397		1,252.190**	332.347
	(377.662)	(509.033)		(503.883)	(654.543)
reservoir FEs	No	No	No	Yes	Yes
observations	1078				

Notes: The columns represent average treatment effects estimates from OLS regression for different orders of polynomial distance interactions. Rows represent different independent variables. Robust standard errors clustered at the lease level are reported in parentheses. Only specifications where the polynomials are jointly significant are reported.

Nonparametric Specifications

Parametric regressions have the weakness that they use data far from the border in order to estimate the discontinuity, adding potential bias. Regression discontinuity design thus lends itself to nonparametric methods (Lee and Lemieux 2010). For the purposes of robustness we employ two nonparametric estimators: local linear, and rectangular kernel. The local linear estimator is given according to

$$min_{\alpha_{OK},\beta_{OK}} \sum_{i:c\leq X_i<c+h} (Y_i - \alpha_{OK} - \beta_{OK}(X_i - c))^2 \tag{3}$$

and

$$min_{\alpha_{TX},\beta_{TX}} \sum_{i:c-h<X_i\leq c} (Y_i - \alpha_{TX} - \beta_{TX}(X_i - c))^2. \tag{4}$$

The average treatment effect is then calculated as[11]

$$\tau = \alpha_{OK} - \alpha_{TX}. \tag{5}$$

The average treatment effect for the rectangular kernel estimator is given by

$$\tau = \frac{\sum_{i:x_i\geq c} Y_i * K(\frac{x_i-c}{h})}{\sum_{i:x_i\geq c} K(\frac{x_i-c}{h})} - \frac{\sum_{i:x_i<c} Y_i * K(\frac{x_i-c}{h})}{\sum_{i:x_i<c} K(\frac{x_i-c}{h})} \tag{6}$$

where $K()$ represents the rectangular kernel. Each of these estimators requires choosing a bandwidth, h. To show robustness, we use bandwidths of a half-mile and of two miles. Additionally, we perform estimates based on the optimal bandwidth that minimizes the cross validation criterion,

$$CV_y(h) = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_{-i}(X_i, h))^2. \tag{7}$$

[11] For local linear estimates the fits are calculated at the border, i.e., x=c, so that only the intercept terms are used. See (Imbens and Lemieux 2008).

This procedure is very close to minimizing the estimate of residual variance, except for two alterations. First, in estimating $\hat{Y}_{-i}(X_i, h)$, the ith observation must be left out, else the bandwidth would shrink until the predictions perfectly match the data (Yatchew 1998). Second, the goal is to choose h, so as to get the best possible estimates at the border. But to make predictions for $Y(1)$ and $Y(0)$ at the border, we can use observations from only one side. Only observations to the north are used for predicting $Y(1)$ (observations in Oklahoma); Only observations to the south in predicting $Y(0)$ (observations in Texas). Therefore, in fitting $Y\hat{(1)}_{-i}$ we use observations only to the north, in fitting $Y\hat{(0)}_{-i}$ we use observations only to the south. This is the procedure used by Ludwig and Miller (Ludwig and Miller 2007) and recommended by Imbens and Lemieux (Imbens and Lemieux 2008).

Also, adapted from Imbens and Lemieux (Imbens and Lemieux 2008), the asymptotic distribution of the average treatment effect is given by

$$\sqrt{Nh}(\hat{\tau} - \tau_0) \rightarrow \mathcal{N}(0, \frac{4}{f_X(c)}(\sigma^2_{Y_{OK}} + \sigma^2_{Y_{TX}})$$

where $f_x(c)$ is the an estimate of the density. Consistent estimates of the variances and the density are given by

$$\hat{\sigma^2_{Y_{TX}}} = \frac{1}{N_{h,TX}} \sum_{j:c-h \leq X_j \leq c} \hat{\epsilon^2_j}$$

$$\hat{\sigma^2_{Y_{OK}}} = \frac{1}{N_{h,OK}} \sum_{j:c<X_j \leq c+h} \hat{\epsilon^2_j}$$

$$f_X\hat{(}x) = \frac{N_{h,OK} + N_{h,TX}}{2Nh}$$

where, $\hat{\epsilon_j}$ are the estimated residuals.

Nonparametric estimates for the average treatment effect of Oklahoma policy for the sample of young wells are given in Table 5. Before estimation, the sample is averaged over the first two years of production for each lease, resulting in 1,079 observations. Each column represents a different bandwidth and/or specification. We find that at the border, and across all bandwidths and specifications, there is no significant difference in either longitude or depth of well–two variables that we expect to vary exogenously with policy. Also, by design, there is no significant difference in the time the well is produced (measured here in days). Treatment effects for measures of the dependent variables of interest, log oil and log gas production, log cumulative revenues and log cumulative production are nearly always negative. The variables that are particularly important are log gas and log oil. Treatment effects of Oklahoma policy for log gas range from an insignificant -20.3 % to a large and highly significant difference of -364 %, (1,966 MCF of natural gas, evaluated at the border). In circumstances where the local linear and Nadayara-Watson-style rectangular kernel give different estimates, it is the local linear estimates that are more reliable, the rectangular kernel estimates more likely to be biased (Fan 1992).

Table 5: Nonparametric estimates: sample of young wells

	Opt. Loc. Lin.	Opt. Rect. Kern.	Loc. Lin Bin=1/2 mi.	Loc. Lin. Bin=2 mi.	Rect. Kern. Bin=1/2 mi.	Rect. Kern. Bin=2 mi.
log gas	-1.210***	-0.907	-3.637***	-0.727	-0.203	-0.802
	(0.333)	(0.741)	(1.140)	(0.540)	(1.144)	(0.541)
log cumulative gas	-1.141***	-0.838	-3.360**	-0.506	0.151	-0.769
	(0.399)	(0.886)	(1.379)	(0.652)	(1.370)	(0.648)
log oil	0.034	-0.545	-0.111	-0.097	-0.173	-0.500
	(0.335)	(0.557)	(1.107)	(0.524)	(1.063)	(0.502)
log cumulative oil	0.176	-0.398	-1.126	-0.197	-0.152	-0.269
	(0.552)	(0.753)	(1.479)	(0.699)	(1.442)	(0.679)
depth	365.425	110.396	1106.31	48.078	18.572	43.335
	(238.048)	(434.497)	(718.773)	(338.996)	(712.940)	(335.788)
longitude	0.014	-0.022	0.080	-0.056	0.052	0.039
	(0.121)	(0.143)	(0.223)	(0.105)	(0.222)	(0.105)
produced time	2.955	6.800	26.705	14.440	42.147	6.696
	(11.933)	(13.479)	(36.172)	(17.073)	(36.043)	(17.010)
cumulative rev	-0.733**	-0.568	-2.534***	-0.735*	-0.51	-0.507
	(0.315)	(0.386)	(0.797)	(0.377)	(0.780)	(0.367)
observations	1079	1079	1079	1079	1079	1079

Notes:Columns report estimates of ATE from different nonparametric regression specifications. Bandwidth in the first two columns is chosen endogenously to minimizes the cross-validation criterion.

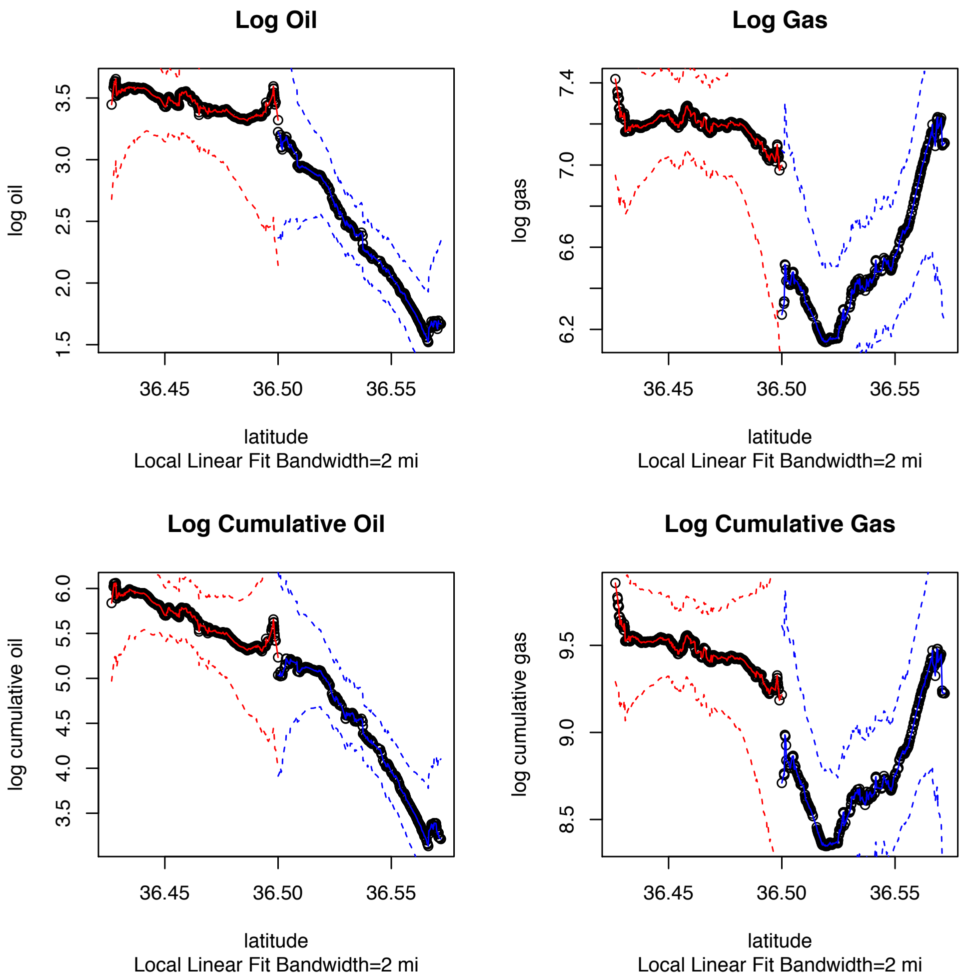
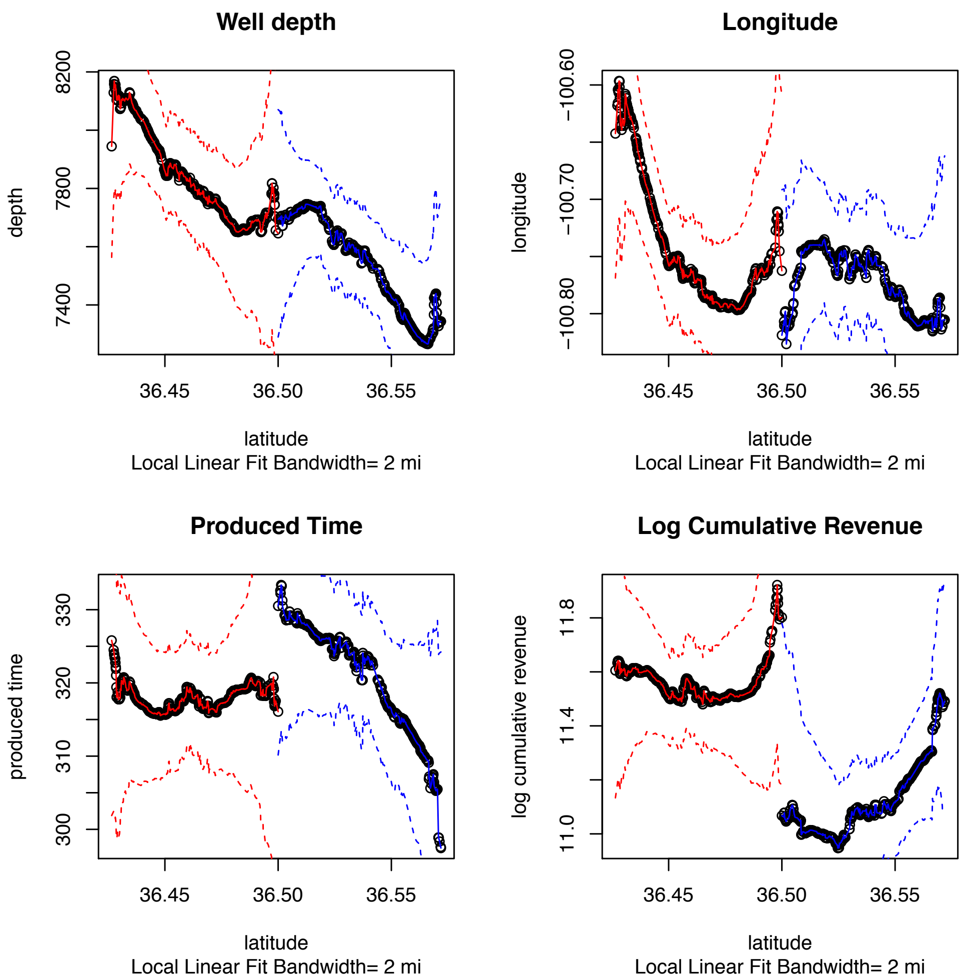


Figure 2: Local linear regression: sample of first 24 months of production, endogenous variables. Solid lines represent the fitted model based on local linear estimates with two-mile bandwidth. Confidence intervals based on bootstrapped standard errors from 500 draws are represented with dashed lines. Texas is to the left in red, Oklahoma to the right, in blue.

In terms of log oil the rectangular kernel and local linear estimates are closer together, and with one exception, negative in sign. No bandwidth or estimate gives significant difference between Texas and Oklahoma oil production. These results combined with those of log gas indicate that, while there is little difference in initial production of oil between Oklahoma and Texas, Texas owners do seem to bleed away more gas in this initial phase of recovery.



!htpb]

Figure 3: Local linear regression: sample of first 24 months of production, exogenous variables. Solid lines represent the fitted model based on local linear estimates with two-mile bandwidth. Confidence intervals based on bootstrapped standard errors from 500 draws are represented with dashed lines. Texas is to the left in red, Oklahoma to the right, in blue.

From Section 2, the initial degassing of wells should have consequences for cumulative recovery. Indeed, Table 6 shows the impact on cumulative oil recovery is large. Significant estimates (at conventional levels) indicate that Oklahoma recovers between 145.3% and 159.0 % more oil than Texas. Moreover, there is no statistical difference between Oklahoma and Texas in terms of cumulative gas recovery. Analysis of treatment effect estimates for longitude and well depth show no evidence that the Texas and Oklahoma samples differ in exogenous observable covariates. Looking at production on the last day, it also seems clear that the threshold for production is lower in Oklahoma, as coefficients are again almost uniformly negative (significantly so for gas). Oklahoma wells are produced significantly longer (around 5 years longer, coefficient estimates are in days).

Table 6: Nonparametric estimates: sample of old wells

	Opt. Loc. Lin.	Opt. Rect. Kern.	Loc. Lin Bin=1/2 mi.	Loc. Lin. Bin=2 mi.	Rect. Kern. Bin=1/2 mi.	Rect. Kern. Bin=2 mi.
log gas	-1.405***	-1.392*	-3.013***	-1.393**	-1.697	-1.457***
	(0.513)	(0.732)	(1.083)	(0.551)	(1.086)	(0.552)
log cumulative gas	-0.369	-0.663	-2.039	-0.953	-0.47	-0.768
	(0.986)	(1.317)	(1.512)	(0.772)	(1.509)	(0.774)
log oil	-0.575	-0.384	0.776	-0.546	-0.472	-0.052
	(0.371)	(0.534)	(0.836)	(0.424)	(0.821)	(0.417)
log cumulative oil	1.453***	1.095	1.590**	1.548	1.920	0.894
	(0.508)	(0.936)	(1.442)	(0.734)	(1.435)	(0.730)
depth	26.659	-274.81	584.368	391.009	136.55	-186.211
	(205.160)	(234.825)	(601.531)	(304.504)	(599.401)	(303.674)
longitude	-0.003	-0.028	-0.291	-0.045	0.017	0.029
	(0.074)	(0.127)	(0.185)	(0.093)	(0.185)	(0.093)
produced time	1389.959***	1565.842	896.196	1528.266**	2396.666**	1668.687***
	(493.764)	(961.956)	(1,192.887)	(603.627)	(1,188.109)	(602.149)
cumulative rev	-0.128	0.089	-0.134	-0.241	0.314	0.016
	(0.316)	(0.367)	(0.783)	(0.398)	(0.778)	(0.395)
observations	1078	1078	1078	1078	1078	1078

Notes: Columns report estimates of ATE from different nonparametric regression specifications. Bandwidth in the first two columns is chosen endogenously to minimizes the cross-validation criterion.

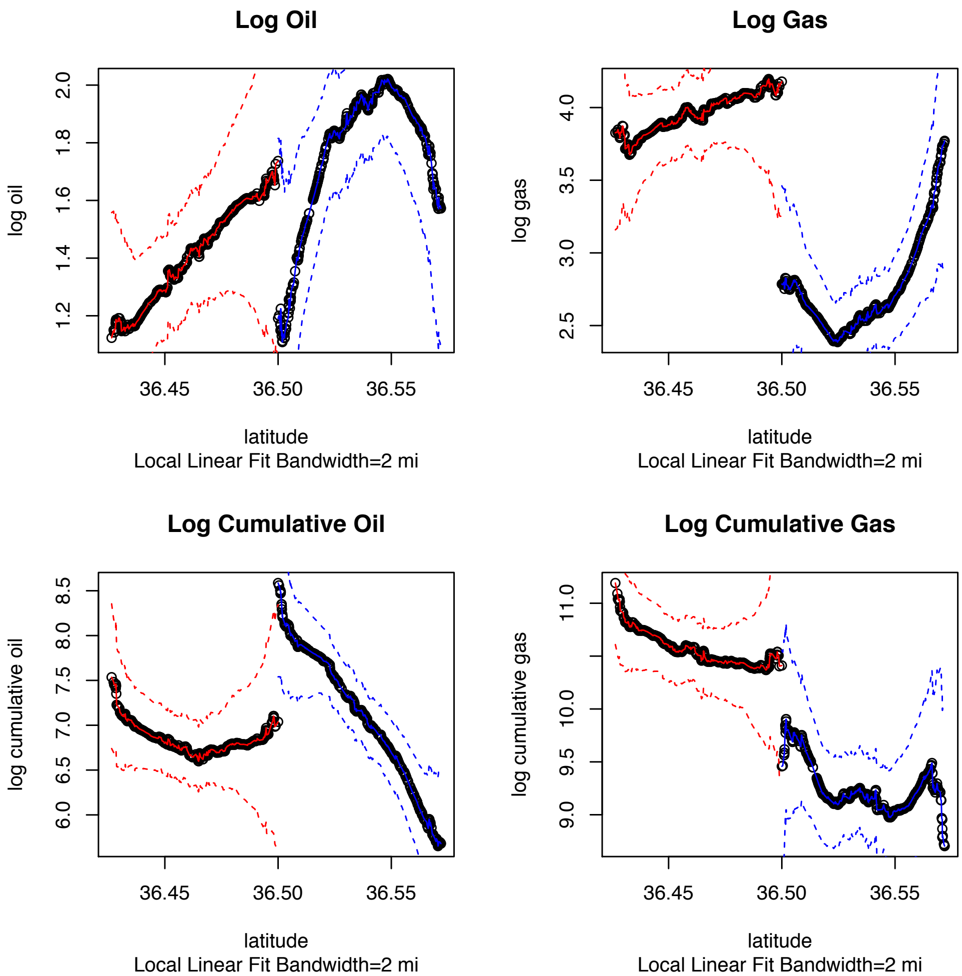


Figure 4: Local linear regression: sample of last reported production, endogenous variables. Solid lines represent the fitted model based on local linear estimates with two-mile bandwidth. Confidence intervals based on bootstrapped standard errors from 500 draws are represented with dashed lines. Texas is to the left in red, Oklahoma to the right, in blue.

Partial Linear Model

In an effort to control for the well age, while estimating the impact of well location nonparametrically, we estimate a differencing model similar to the Robinson (1988) partial linear model (Yatchew 1998). In particularly, we estimate the following equation separately for Oklahoma and Texas,

$$Y_{it} = \alpha + \beta * X_{it} + \gamma_t + f(latitude_i) + \epsilon_{it}. \tag{8}$$

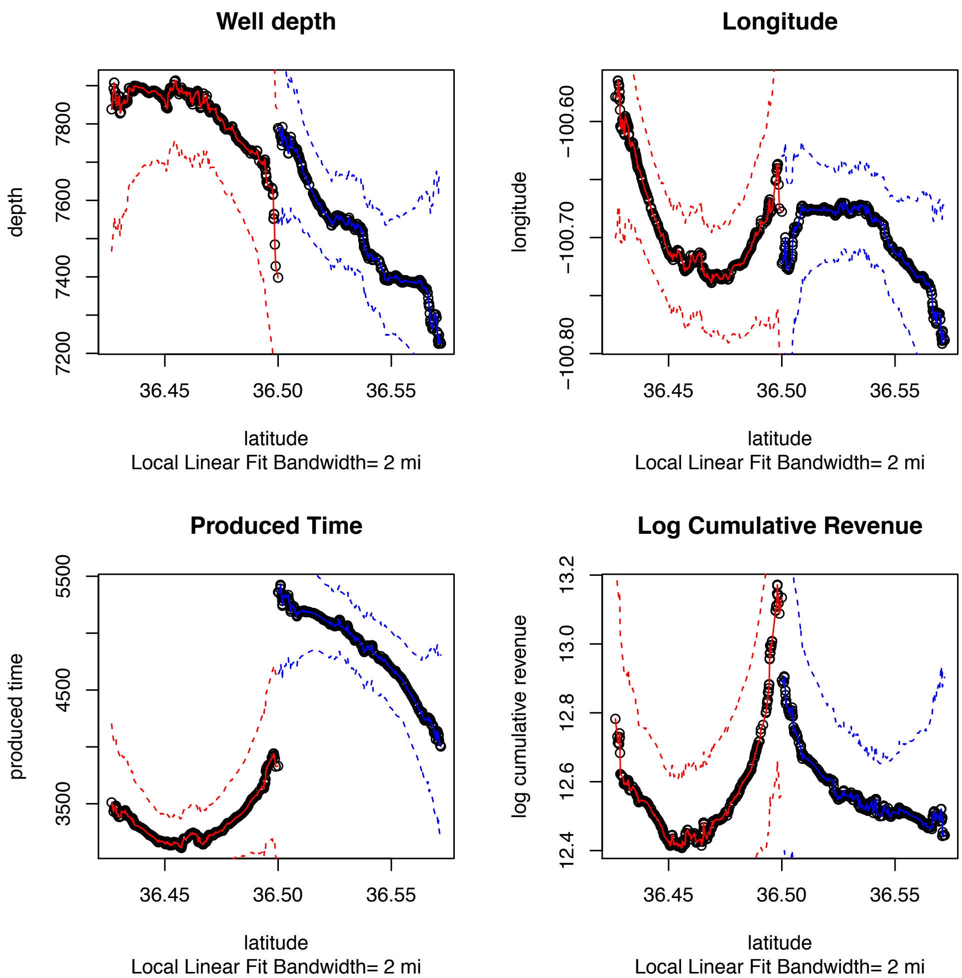


Figure 5: Local linear regression: sample of last reported production, exogenous variables. Solid lines represent the fitted model based on local linear estimates with two-mile bandwidth. Confidence intervals based on bootstrapped standard errors from 500 draws are represented with dashed lines. Texas is to the left in red, Oklahoma to the right, in blue.

The vector X_{it} includes well age, the square of well age, longitude, and well depth; γ_t is a vector of year-month dummy variables. It is the function $f()$ that is of particular interest, as it represents the component of production accounted for by latitude. The equation is estimated by ordering the data according to latitude and then estimating β and τ with the difference estimator (Yatchew 1998) to get the linear parameters.[12] These parameters in hand, we can then back out fitted values for $\hat{f}(latitude_i)$ which we smooth with the local linear smoother. Then by taking the difference $\hat{f}_{OK}(latitude = 36.5^{\circ}) - \hat{f}_{TX}(latitude = 36.5^{\circ})$ we isolate the average treatment effect, after controlling for differences in well age and depth, which may confound our results. These treatment effects are presented in Table 7.

Table 7: PLM estimates: full sample

	Loc. Lin Bin=1/2 mi.	Loc. Lin. Bin=2 mi.
log gas	-2.302***	-0.792
	(0.458)	(0.459)
log cumulative gas	-1.302***	-0.187
	(0.486)	(0.489)
log oil	0.475	-0.412
	(0.396)	(0.399)
log cumulative oil	-0.38	-0.206
	(0.592)	(0.594)
observations	300,946	300,946

Notes:Discontinuity of spatial component of partial linear model given in equation 7. Fits local-linearly smoothed with a bandwidth of a half-mile (1st column) and two miles (2nd column).

Under the identification assumption that the productivity of the wells does not change discontinuously at the border, there should be no difference between the contribution of latitude to production based on whether the well is in Oklahoma or Texas. Yet the first column of Table 7 indicates Texas wells are producing more gas.

[12]The data is ordered by latitude and then differenced by latitude to eliminate the impact of latitude, any difference in Y must be the result of difference in the other covariates.

This is consistent with the story of the rapid degassing of wells seen in the sample of young wells, and reported in Table 5.

False-border test

To illustrate that estimates for average treatment effects at the Oklahoma-Texas border are attributable to policy and not merely to our spatial identification strategy, we perform two false-border tests. That is, we test for average treatment effects across latitudes where there should be none. Significant estimates would cast doubt on our identification strategy. So as not to contaminate the analysis with real policy differences, we split the sample of old wells into Oklahoma and Texas subsamples. It is possible to imagine many false borders. We pick two lines of latitude at the midpoints of the Oklahoma and Texas data. The Oklahoma false border is given according to the formula, $[36.5^{\circ} + max(latitude_{OK})]/2 = 36.536^{\circ}$; the Texas false border is given by $[36.5^{\circ} + min(latitude_{TX})]/2 = 36.463^{\circ}$. Lines of latitude are chosen for the false border to maximize the number of relevant observations.[13] We use the midpoint formula in selecting this line so that we have roughly the same number of observations on either side of the line.

[13]Because of the distribution of our sample, very many points would lie "far" from any given line of longitude, or diagonal.

Table 8: Nonparametric false border test: old wells

	Rect. Kern. Bin=1/2 mi. TX	Loc. Lin. Bin=1/2 mi. TX	Rect. Kern. Bin=2 mi. TX	Loc. Lin. Bin=2 mi. TX	Rect. Kern. Bin=1/2 mi. OK	Loc. Lin. Bin=1/2 mi. OK	Rect. Kern. Bin=2 mi. OK	Loc. Lin. Bin=2 mi. OK
log gas	-0.413	-1.337	0.026	-0.105	0.463	1.199	0.214	0.072
	(1.070)	(1.069)	(0.535)	(0.533)	(0.945)	(0.940)	(0.504)	(0.500)
log cumulative gas	-0.281	2.355**	-0.082	0.084	-0.601	-1.052	-0.069	-0.819
	(1.207)	(1.201)	(0.603)	(0.605)	(1.520)	(1.521)	(0.812)	(0.802)
log oil	0.122	-0.09	0.302	0.094	-0.437	-0.175	0.206	-0.242
	(0.748)	(0.747)	(0.375)	(0.375)	(0.800)	(0.798)	(0.423)	(0.420)
log cumulative oil	-0.617	1.215	0.010	-0.115	-0.451	0.223	-0.595	-0.392
	(1.489)	(1.487)	(0.745)	(0.743)	(1.227)	(1.221)	(0.652)	(0.646)
depth	-344.112	870.918	-43.566	-131.366	-10.073	19.263	-83.73	61.663
	(654.663)	(652.372)	(326.544)	(326.129)	(466.442)	(464.743)	(245.830)	(244.409)
longitude	0.017	0.334*	-0.007	0.030	-0.112	-0.065	-0.056	-0.108
	(0.192)	(0.190)	(0.096)	(0.095)	(0.154)	(0.153)	(0.081)	(0.080)
produced time	48.013	1677.273*	211.08	-117.500	158.776	644.666	-133.955	-63.105
	(1088.963)	(1083.912)	(544.371)	(541.561)	(1113.380)	(1111.505)	(587.362)	(586.213)
cumulative rev	-0.356	1.544**	-0.067	-0.182	0.108	0.42	0.147	-0.078
	(0.769)	(0.766)	(0.387)	(0.386	(0.694)	(0.693)	(0.365)	(0.365)
observations	493	493	493	493	585	585	585	585

*Notes:*This table tests the sample of Texas and Oklahoma wells for evidence of treatment effects at a false border at 36.463° and 36.536° , respectively. Columns represent different samples, specifications and bandwidths.

Nonparametric estimates for average treatment effect at these false borders are presented in Table 8 for the sample of old wells from Texas and Oklahoma. Estimation is carried out using the same formulas as in the previous section. As there are 64 different estimates for an ATE in the table, we expect some to come up significant. Indeed, there are two significant results at the 95% level in the second column, yet these results do not hold across columns, nor do even the directions of the estimates. We conclude that there is not convincing evidence for discontinuities where there is no policy shift.

Analysis

Parametric regressions of Table 2 show Oklahoma produces its wells at a slower rate, both in terms of monthly oil and gas production and current revenue. We also see that Oklahoma wells are older. The first explanation is exponential decline of well productivity with well age–Oklahoma wells are older and so they should produce less. The puzzle is that Oklahoma wells also seem to have recovered less oil (though we say this with somewhat less confidence). An equally compelling story is that Oklahoma wells are, on average, produced at a lower rate of production, and that the wells are still young enough that the full benefits in terms of cumulative production are not yet apparent. Finally, it is possible that we have made a bad assumption in spatial identification, and for some reason Oklahoma wells are simply less productive.

As for the reasonableness of the assumption, the false-boundary tests indicate that intrinsic well productivity does not discontinuously change across space where policy shifts are not involved. If we are to differentiate between the other two hypotheses–of whether Oklahoma wells might reap benefits down the road from reduced production early on, or whether they are merely less productive–we must eliminate the confounding influence of the age of the well. To do this we focus on a sample of observations early in the life of a well (within the first 24 months) and a group of observations for wells on their final day of production. Parametric regressions of wells early in the life of the well indicate that Oklahoma wells are produced less intensively in terms of revenues, but that there is no statistical difference in terms of quantities of oil and natural gas produced. The revenues may be interpreted as a lowered joint rate of production. The sample of old wells indicates that there is no cumulative difference in natural gas recovery, but that Oklahoma produces 270 % more oil than Texas. The parametric results point toward Oklahoma wells being produced more sustainably early in the life, and that this leads to greater cumulative resource recovery.

Our identification of the impacts of policy is based on the spatial contiguity of wells, and so a worry with the parametric regressions is that estimates of average treatment effects for wells near the border might unduly influenced by wells far from the border. To guard against this, we repeat the analysis using nonparametric specifications. The general story is the same: Oklahoma wells seem to produce at a lower rate in the first years of life, and in particular, are subject to less degassing. The benefits of this policy are apparent late in the life of a well in terms of a significant, but more modest estimate of 145 % to 153 % greater recovery.

The difference in recovery is stark. At one level, the Anadarko Basin predominately produces gas. Yet the difference in cumulative oil production evaluated at the border over the course of the life of the well is 3,360 barrels (in comparison to the Texas lifetime well production of 1,026 barrels at the border). [14] Yet for effective secondary recovery operations to occur, fields must have concentrated ownership. This is even more important for the Anadarko basin, considering its mature state of development. And so it seems that for our sample of

[14] Without accounting for tax difference or discounting, and assuming a conservative price of $90/bbl (the WTI Cushing price as of April 15, 2011 exceeded $108/bbl) this amounts to a lump sum benefit of $302,400 for drilling in Oklahoma rather than Texas.

wells near the border, the common-pool damage may be two pronged–damage to the well as a result of wasteful competition, limiting full primary recovery, and through fractured ownership, which impedes secondary recovery. Especially in light of the latter, the large difference in cumulative oil recovery is not unreasonable.

We have focused on the difference in unitization rates as the policy difference that drives the difference in output, yet at the Texas-Oklahoma border many policies change discontinuously, including, well density, and tax regimes, and so it can be argued that the impact of unitization is not wholly identified. Yet the impact of well density should bias downwards our estimates for the treatment effect. Oklahoma has a higher well density, and is thus sweeping out less area per well, and should therefore recover less. Yet we found Oklahoma wells recover more. Texas has a higher royalty on natural gas, and a lower royalty on oil, this should result in the postponement of natural gas production into the future relative to Oklahoma, yet we see Texas produces more natural gas early on.[15] We therefore conclude our estimates for the average treatment effect of the difference in unitization at the border are a lower bounds of the true average treatment effects.

Conclusion

The physical dynamics of oil and natural gas production maintain that a slower rate of production leads to enhanced primary recovery by conserving the gas pressure of the field. The importance of unitization to secondary recovery also suggests that unitized fields should be more successful in terms of cumulative recovery during this phase of production. Oklahoma, with its greater emphasis on unitization, should be more successful than Texas in terms of cumulative physical recovery. To test this hypothesis, we employ regression discontinuity design. Because the production of hydrocarbons is geologically driven, wells close together in space should be similarly

[15] This analysis uses a simple model of taxation (Gamponia and Mendelsohn 1985). We do not account for the endogeneity of cumulative recovery, or the joint nature of production.

productive. A line at 36.5° latitude randomly exposes wells that are otherwise similar to different policies; those wells north of the line receive the Oklahoma treatment hypothesized to yield greater cumulative recovery, those below the line receive the Texas treatment. Discontinuous jumps in outcomes at the border identify treatment effects of the Oklahoma policy relative to Texas.

Looking at a sample of young wells and old wells to disentangle the impact of policy over the course of a well's life cycle, we find evidence that Oklahoma wells are produced more slowly early on (in joint terms of oil and natural gas), are more successfully unitized, and as a result are much more successful in terms of cumulative oil recovery.

Chapter IV

SPATIAL SPILLOVERS IN OIL AND NATURAL GAS PRODUCTION

Introduction

The goal of this paper is to quantify the spillovers in production and injection in oil and natural gas recovery in Slaughter field of West Texas. Oil and natural gas deposits are concentrated spatially in reservoirs, yet it is often the case in Texas that no single producer has rights to the entire reservoir (Libecap and Wiggins 1985). Because property rights to the oil and natural gas are administered according to the "rule of capture," ownership is not fully secured until the resource is extracted. While in the ground, it is an example of a common good: rival but nonexclusive. Under these circumstances the resources can be the subject of fierce competition as neighboring producers race to extract. Economic rents are dissipated in the drilling and operation of more wells than are needed to efficiently drain the reservoir (Weitzman 1974). Additionally, owing to the dynamics of recovery, overly rapid extraction can result in damage to the reservoir and lowered cumulative recovery (Dake 2001).

The present age of the majority of Texan fields only compounds the common pool inefficiencies. In order to maintain the productivity of a maturing field, operators shift production wells into injection. These injection wells pump substances (*e.g.*, water, carbon dioxide, natural gas) into the reservoir to drive the resource towards neighboring production wells where it is then extracted. Injection is costly, and it makes little sense to undertake injection when ownership of the reservoir is highly fractured and neighboring wells are likely to be owned by competing operators. The resulting miserly secondary injection can lead to substantial losses in recovery (Libecap and Wiggins 1985). Injection wells may also be used to offset the production at neighboring wells. A neighbor's production creates a cone of depression in the resource bearing strata; oil and natural gas tend to flow towards this depression. To prevent the resource from escaping the lease, injection wells may be drilled along the border to halt resource emigration. These offset injection wells are unnecessary for production and represent another economic cost of intra-field competition (Libecap 1989a).

The spatial interdependence of oil and natural gas production brings the potential for economic inefficiency because competing producers discount the value of leaving resource *in situ* for future periods resulting in a "race to extract." The goal of this paper is to demonstrate how a race to extract can be prevented through unitary ownership. I use a spatial econometric model to explicitly characterize spillovers in production while controlling for unobserved spatial autocorrelation. This type of spatial model has been used recently to estimate spillovers in production of fossil groundwater aquifers (Savage and Brozovic 2011) (Brozovic et al. 2006) (Pfeiffer and Lin 2009). The model is estimated using extensive and novel data provided by HPDI Corporation. The main result is as expected: when neighboring wells are under unitary ownership, extraction proceeds at a comparatively slower pace than when wells have competing operators.

Background

Spillovers in the production of oil and natural gas, where one producer's extraction interferes with that of another, have been pervasive since the initial discovery of the resources (Yergin 2008). Addressing these spillovers in production are economically important because resource rents can be dissipated in excess effort and capital (Gordon 1954) (Scott 1955) (Hardin 1968) (Weitzman 1974) (Brown 1974). Whereas it might take only a few wells to efficiently drain a field, competing producers may drill many more in order to extract and secure the resource so that it is not lost to neighboring producers. The capital tied up in excess wells could be more efficiently used elsewhere in the economy. Additionally, the common pool nature of the hydrocarbons undermines the producer's incentive to conserve and so is dynamically inefficient (Eswaran and Lewis 1984) (Khalatbari 1977) (Long 1974) (Long 1975) (Dasgupta and Heal 1979) (Reinganum and Stokey 1985).

What is peculiar to oil recovery, however, is that the race to extract can cause damage to the reservoir, limiting ultimate recovery (Libecap and Wiggins 1985) (Chermak and Patrick 2001). Thus, the consequences of the common pool are not limited to economic inefficiencies of too costly extraction, too soon, but to physical inefficiencies as well. Overly rapid production destroys the resource. Oil and natural gas exist in solution, and it is the expansion of natural gas that drives the oil to the well-face and then up to the surface. Rapid extraction can cause the natural gas to bubble out of the mixture. The natural gas is more mobile than oil, and is quickly drawn off. Meanwhile, the oil becomes increasingly viscous, and so difficult to move as to be permanently unrecoverable. It may well be that it is economically efficient to sacrifice cumulative recovery in favor of present extraction (Clark 1973), but this aspect of oil and natural gas exploitation has yet to be studied by economists.

The spillovers in production are essentially issues of property rights. When production spillovers are large and involve a small number of agents, it is reasonable to expect private contracting to solve the problem. In a series of papers, Libecap and Wiggins describe the contracting failure in Texas. Libecap and Wiggins (1984) consider three mechanisms through which leaseholder can address production spillovers: lease consolidation, unitization, where competing leaseholders hire a single operator to jointly develop the field, and prorationing agreements on output. The authors examine five oil fields in Texas, and find that firm concentration is an important determinant of private contracting. Bargaining costs increase with the number of firms, inhibit unitization and consolidation, and in some cases, the ownership of the field is so fractured as to even prevent prorationing agreements. Libecap and Wiggins (1985) study the impediments to unitization agreements. Comparing Wyoming, Oklahoma and Texas, the authors find Texas to be particularly poor at unitizing fields because the unananimity required for unitization creates a holdout problem. Wiggins and Libecap (1985) model unitization negotiations, and test the model empirically, finding that imperfect information about reserves when combined with diffuse landholding prevents unitization. When contractual response fails, lease owners will even split individual leases among competing operators to increase inflow onto the property (Yuan 2002). The work of Libecap and Wiggins nowhere expressly quantifies the size of the spillovers and how these spillovers differ when ownership of the resource is unitary or highly fractured. The present paper contributes to the understanding of the economics of oil and natural gas production by showing that unitary ownership does significantly abate the race to extract as previous theoretical models have predicted.

Regulation is also important to consider when measuring possible interference between leases. Regulation of hydrocarbon production in Texas is overseen by the Texas Railroad Commission and comes in three flavors: command and control, taxes and production quotas. Of the command and control regulations, well-spacing regulations and regulations on the inclination of drilling (slant and horizontal drilling) are the most relevant in addressing issues of common pool production. Slant drilling is prohibited without special permission. Additionally, the statewide spacing rule disallows the drilling of wells within 467' of a property line or within

1200' of an existing well.[16] Although well-spacing guidelines have the virtue of easy verification and enforcement, one-size-fits-all regulations are not flexible enough to account for the heterogeneity in permeabilities and flow dynamics. Owing to local geologic conditions wells 100' apart may communicate less than wells 3000' apart in more permeable rock. Optimal well-spacing guidelines should account for the local geologic parameters, and assign well spacing accordingly. By measuring the effect of neighbor's production on own production, this paper can provide evidence as to the efficacy of spacing regulations.

Well-spacing exceptions may be granted to protect ownership rights, or to prevent resource waste. In the former case, a producer would be allowed to drill closer to a property line if drilling according to regulation would result in substantial portion of the resource underlying the lease to be captured by neighboring producers. In the latter case, exception may be granted if the oil could not otherwise be recovered. Yet these two goals frequently conflict when production tracts are small, as is the case in Slaughter field. Until the decision *Halbouty vs. Texas Railroad Commission* (1962) small lease holders were given a greater production allowable, to cover the costs of drilling plus a reasonable profit, even at the expense of neighbor's production (Lowe 2003). The alternative to well-spacing exceptions, preferred by most states, is forced pooling.

Monthly quotas on production are assigned in Texas as a percentage of a maximum allowable production for the well. Maximum allowable production is based on the depth of the well, and the lease size. In Texas, natural gas is subject to a royalty at of 7.5% of the value of production, while oil is subject to a lower royalty of 4.6% of the value. While these taxes are not consistent with Pigouvian taxation to address the problems of common pool production (Dasgupta and Heal 1979),

[16] Texas Administrative Code, Title XVI, Part 1, Chapter 3

economic theory predicts that these royalties slow the rate of extraction (Gamponia and Mendelsohn 1985).

The interaction between regulation, contracting, geology and firm production decisions determines the nature of the spillover. This paper provides empirical evidence as to how these complex interactions play out on Slaughter field. I find, after controlling for secondary injection, that regulation and contracting have not been fully successful in securing property rights. Indeed the empirical model uncovers evidence consistent with a race to extract.

Methods

The goal of the chapter is to estimate the impact of neighbor's production, y_j, on own production, y_i, for a cross-section of leases indexed $i, j = 1, 2, 3......, N$. Doing this via OLS regression would result in biased parameter estimates because of simultaneity. The problem is that production at j is not predetermined: lease i affects the production of lease j, while the production of lease j simultaneously affects production at lease i. Adding further difficulty is that production on a patch is partially determined by unobserved geological characteristics such as porosity and permeability, and these unobserved variables are likely to be correlated through space. This makes it difficult to tease out what part of production is impacted by neighbors production, and what part of production is the result of correlated but unobserved geological characteristics. Disentangling genuine spatial dependence from autocorrelation is necessary for achieving the goals of this chapter. Fortunately, Kelejian and Prucha (1998) and Kelejian and Prucha (1999) develop a computationally feasible generalized method of moments procedure for controlling for spatial dependence and spatial autocorrelation. The canonical model I estimate is

$$
\begin{aligned}
y &= X\beta + \lambda W y + u \\
u &= \rho M u + \epsilon
\end{aligned} \tag{9}
$$

where y is an $N \times 1$ vector of the dependent variable, X is an $N \times k$ matrix of the k independent variables, W and M are $N \times N$ spatial weighting matrix, β is a $k \times 1$ vector of regression parameters, λ and ρ are scalar spatial parameters, u is an $n \times 1$ vector of regression disturbances, and finally, ϵ are i.i.d. innovations. Full technical assumptions necessary for estimation of the model, as well as the moment condition exploited for estimation, can be found in Kelejian and Prucha (1998) Kelejian and Prucha (1999); however, two assumptions are important to understand the intuition of the model. First, the contribution of nearby producing leases (a, b, c) on the production at lease i are assumed to be a linear function of production, some weighting function (in our case a function of distance) and vector of spillover parameters, $(\lambda_a, \lambda_b, \lambda_c)$, such that $y_i = \lambda_a w(i,a) y_a + \lambda_b w(i,b) y_b + \lambda_c w(i,c) y_c$. In order to estimate the model, we assume that $\lambda_a = \lambda_b = \lambda_c = \lambda$ (otherwise there would be N parameters and N observations). The second assumption is that $|\lambda| < 1$, which insures that spatial spillovers are non-explosive. Similar assumptions hold for the structure of the errors.

Execution of this Generalized Spatial Two-Stage Least Squares (GS2SLS) model requires three stages. In the first stage, to control for simultaneity in production decision, Wy is instrumented for by $H = (X, WX, W^2X, \ldots)$. The implicit exclusion restriction is that a neighbor's X affects your production only through how the neighbor's X affects her own production. Identification comes through the spatial structure—the interaction between X variables and the weighting matrix—and so there is no excluded instrument. This first stage generates consistent results for β and λ, but these estimates are inefficient because the information available in the structure of the autocorrelated errors has yet to be exploited. In the second stage, residuals, $\tilde{u}$, from the first stage are plugged into a moment condition to estimate spatial autocorrelation parameter, ρ, and the innovation variance, σ_ϵ^2. In the last stage, the structure of the autocorrelation is exploited to arrive at more efficient estimates of β and λ.

Interpretation of the spatial parameters depends on the choice of weighting matrix, W. The choice of weight matrix, in turn, is defined by the conceptual framework one uses to interpret the spatial data. There are two possibilities: viewing the data as a lattice of discrete spatial connections, or viewing the data as sample points from a continuous surface (Anselin 2002). In the former case, w_{ij}, representing whether i is connected with j takes on discrete values, 0 or 1. The drawback is that defining the connections can be arbitrary. When the observations are viewed as a sample from a continuous surface, w_{ij} often takes the value of the (inverse) distance between observation i and j. When distance weighting is used, the spatial autoregressive parameter has the potential to be interpreted as a reservoir specific transfer coefficient, which reveals geologic information on reservoir permeability, porosity, and viscosity[17]. Of course strategic interaction between agents will result in biased estimates of the transfer coefficient. To see this consider the primary model of the paper,

$$\begin{aligned} y &= X\beta + \lambda_F Fy + \lambda_U Uy + u \\ u &= \rho Mu + \epsilon. \end{aligned} \tag{10}$$

Here M represents an inverse distance weight matrix controlling for autocorrelation in unobservables. The goal is to see if there is a difference in estimated spillover coefficients between leases that have the same owners and leases with different owners. Weight matrix W from the previous specification is broken up into two

[17]Well reaction functions specified by Theis and Darcy flow equations have been used in physical-economic models of water recovery. See Savage and Brozovic (2011) Brozovic et al. (2006) Pfeiffer and Lin (2009).

separate weight matrices, F and U ("F" for "friendly", "U" for "unfriendly"). Weights in F take on inverse distance weights only when leases i and j have the same operator; conversely, weights in U take on values when leases i and j have different operators. Estimates of friendly (λ_F) and unfriendly spillovers (λ_U) can then be estimated and compared. Without strategic interaction, spillover parameters should be identical and negative, the result of the cone of depression caused by production. With strategic interaction, the estimates of spillovers should be biased upwards and should diverge with $\lambda_F < \lambda_U$. The divergence occurs because the rights to the resource *in situ* are less secure when competing operators own nearby leases. In fact, we may even see a race to extract, which would manifest itself as apparently positive spillovers in production, $\lambda_U > 0$. The positive spillovers come from lease-owners shifting their production profile towards the present. Discussion and results from this model can be found in Section 5.5.

To identify the shift in the production profile due to security of ownership, a variety of cross-sections per field must be analyzed. Field age and ownership structure of the fields affects the degree and nature of spillovers. Early in the life of a highly decentralized field, the measured autoregressive parameter should be positive, reflecting a race to extract. Later in the life of the decentralized field, the spillover may decline toward zero as damage to the reservoir halts flow between wells. Conversely, in a highly concentrated area, the autoregressive parameter might be zero over the entire life of the reservoir as a result of effective management.

Data

Data for analysis is provided by HPDI Corporation, which collects, compiles and publishes oil and natural gas production data for 31 US States, 4 Canadian provinces and the Federal offshore areas in the Gulf of Mexico and the Pacific. Previous research has indicated that Texas is a state where common pool problems can be substantial (Libecap and Wiggins 1984) (Libecap and Wiggins 1985) (Wiggins and Libecap 1985). Not coincidently, Texas also has the most extensive data available, with time series for production, injection and well tests going back even before 1960. Data on leases size come from W-1 drilling permits, public information made available by the Texas Railroad Commission.

The focus of the analysis is Slaughter field in West Texas, located near the Texas-New Mexico border. Wells are mapped in Figure 6. This field has a variety of characteristics making it a worthy focus or research. First, it ranks in the top 20 fields in the US country in terms of 2009 proved reserves for oil (EIA 2009), and is therefore of economic interest. Second, in order to quantify the importance of unitization and ownership concentration, it is necessary to have within field variation of ownership concentration.[18] A local Herfindahl concentration index mapped for Slaughter field in Figure 7 shows the field has variation in ownership concentration.

Slaughter Field was discovered in 1937. My earliest data only goes back to 1955. Figure 8 shows the evolution of the number of productive wells on the field over time. Figure 9 shows the field aggregated history of production. Peak oil production occurred in the middle 1970s. Also evident is a sharp drop in natural gas production between December 2004 and January 2005. This is apparently due to the weighty tails of the distribution. When looking at a figure of logged average well production, no such drop off in production is evident(Figure 10). A large part of the decline in gas production at the end of 2004 can be attributed to a significant decline in gas production on the Slaughter Estate Unit and Central Mallet Unit. These units, under the operation of Occidental Permian Ltd., were the subject of legal controversy. Carbon dioxide and hydrogen sulfide injections on the Slaughter

[18] Because the underlying permeabilities and porosities will differ across fields, the estimated spillover parameter will not be comparable across fields, and so cross-field variation in ownership can not be exploited to demonstrate the impact of ownership concentration on the spillover.

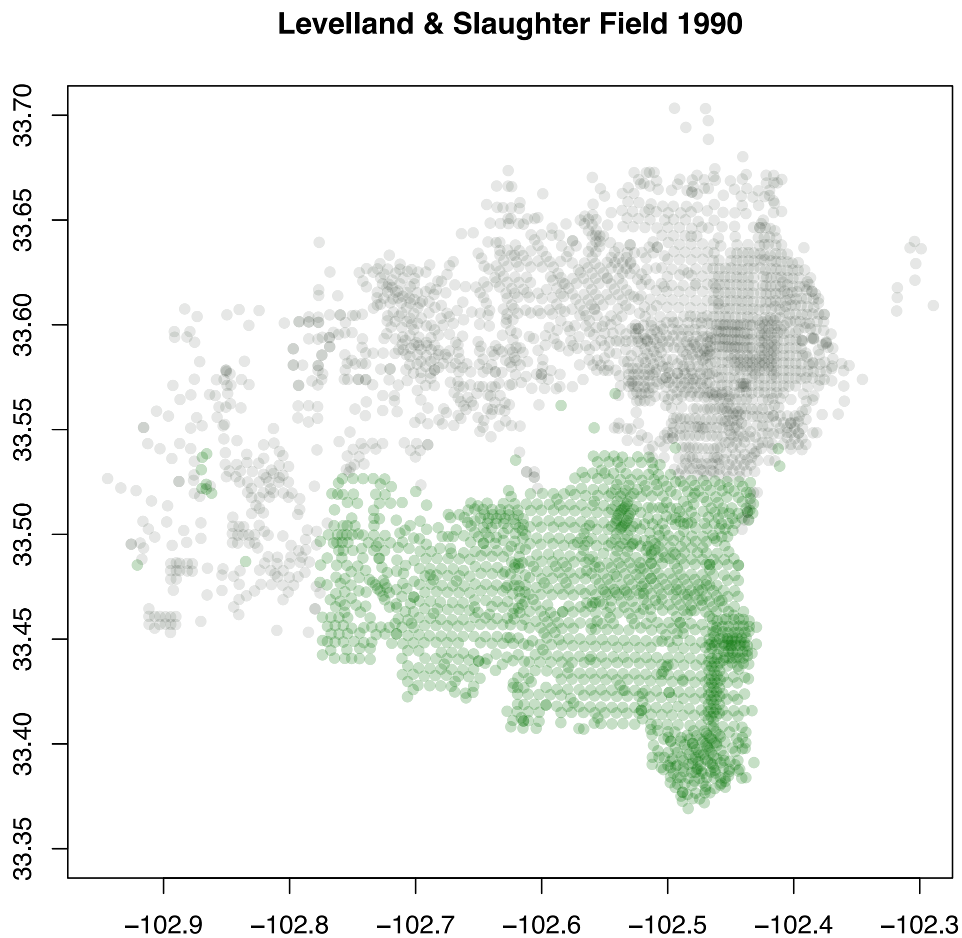


Figure 6: Well locations. Slaughter field is in green. To the north in gray is Levelland field which is geologically similar, but separated from Slaughter by an anhydrite salt dome.

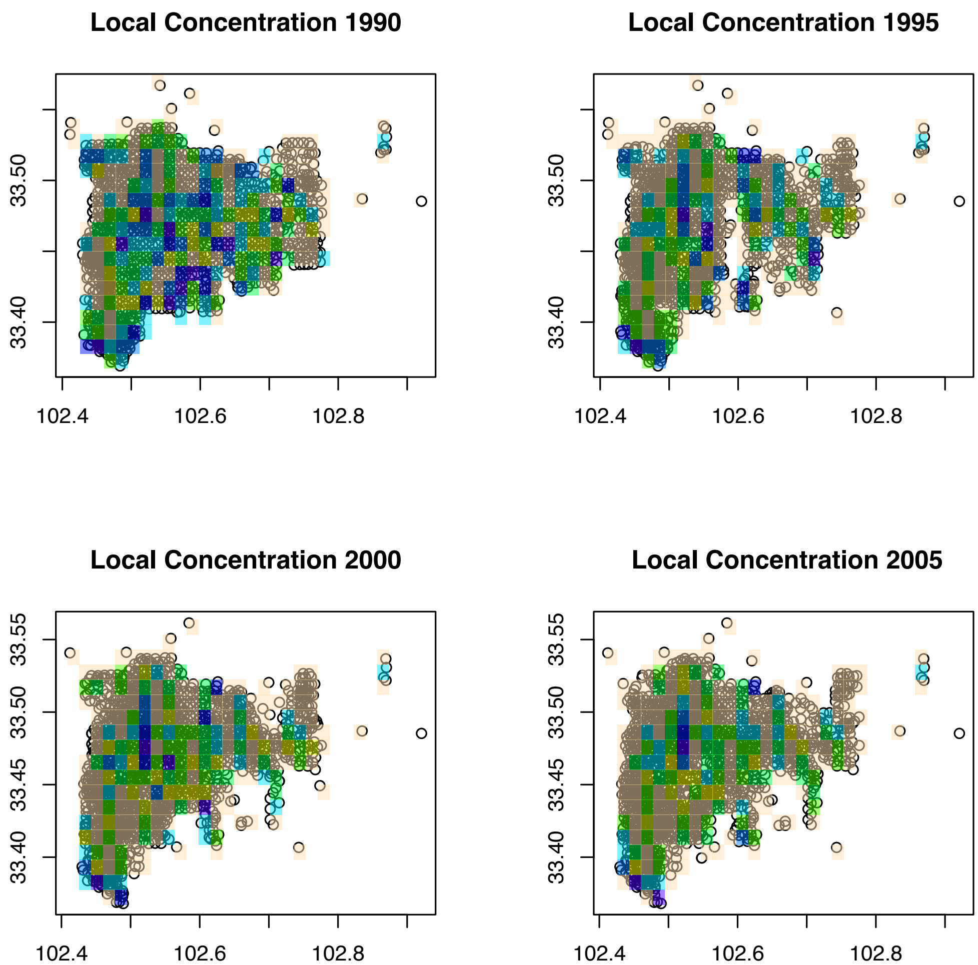


Figure 7: Local Herfindahl concentration index. Herfindahl concentration index computed for each cell in a 30 x 30 grid. Lighter colors indicate higher ownership concentration.

Estate Unit–which have aided in the recovery of oil– contaminated gas in the reservoir making it difficult to process. Occidental Permian, however, owns the gas processing plant, and the Texas Supreme Court held in *Helen Jones Foundation vs. Occidental Permian, 2011* that Occidental used the increased costs of processing as cover to avoid paying royalty owners.

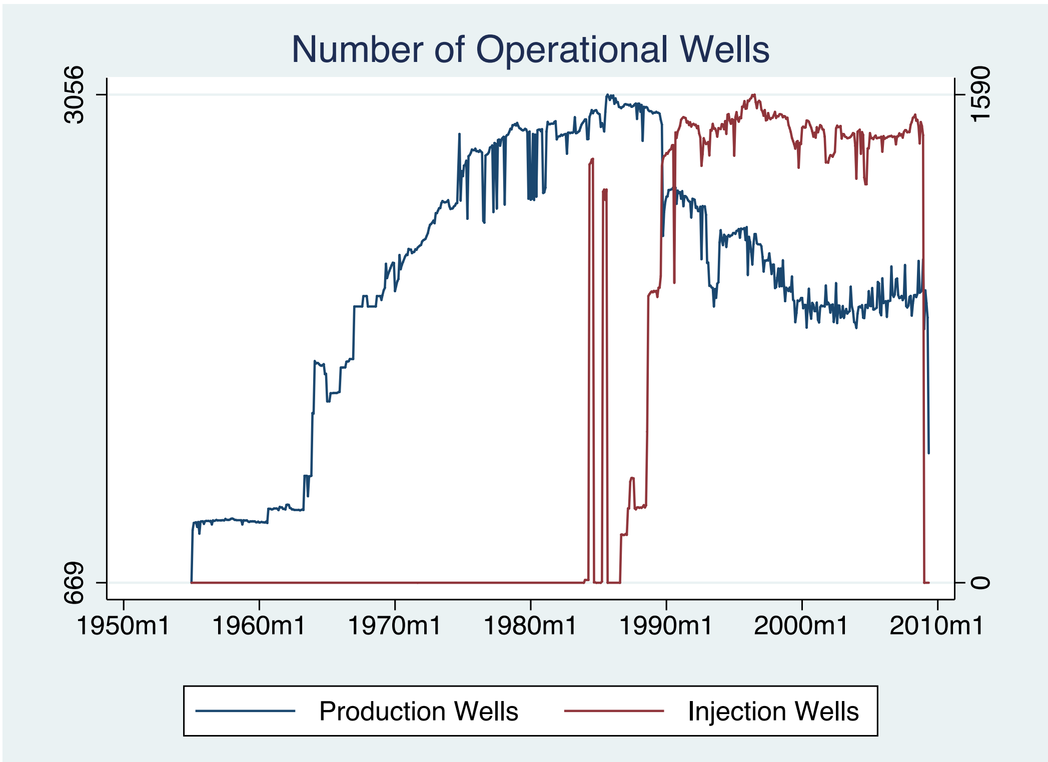


Figure 8: Total wells in production.

Summary statistics for Slaughter field can be found in Table 9, the bottom panel displays summary statistics for the subset of the data used for regression analysis. Production data is at the lease-level, while injection data is at the well level. Data

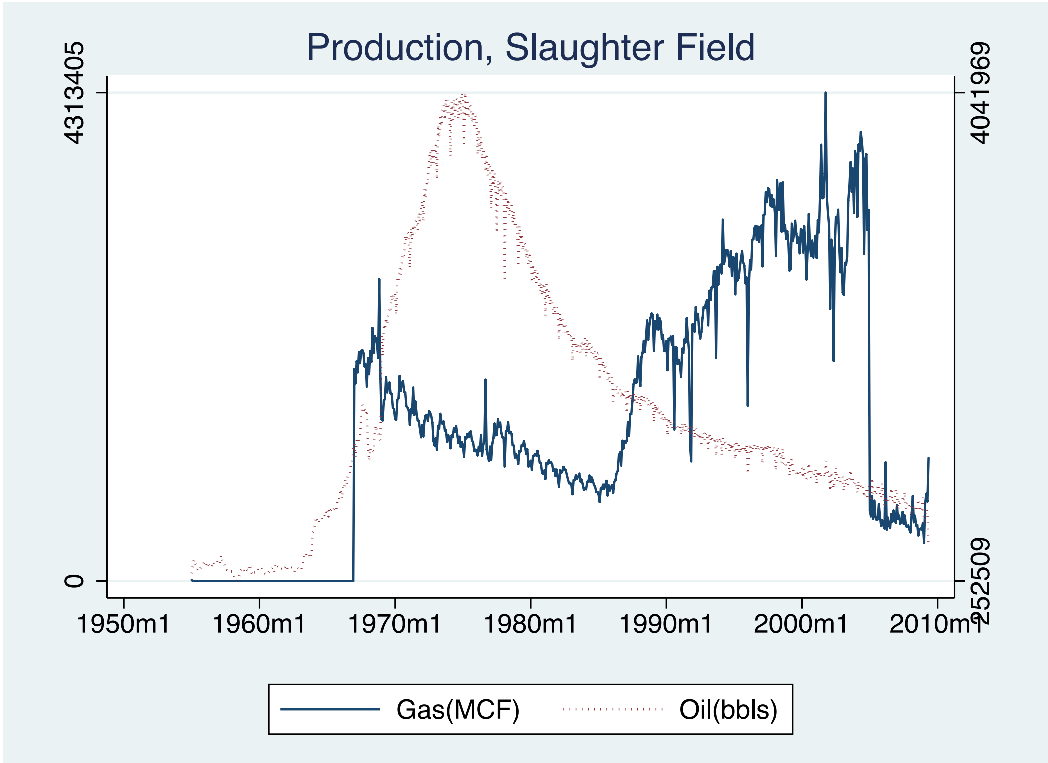


Figure 9: Aggregate production.

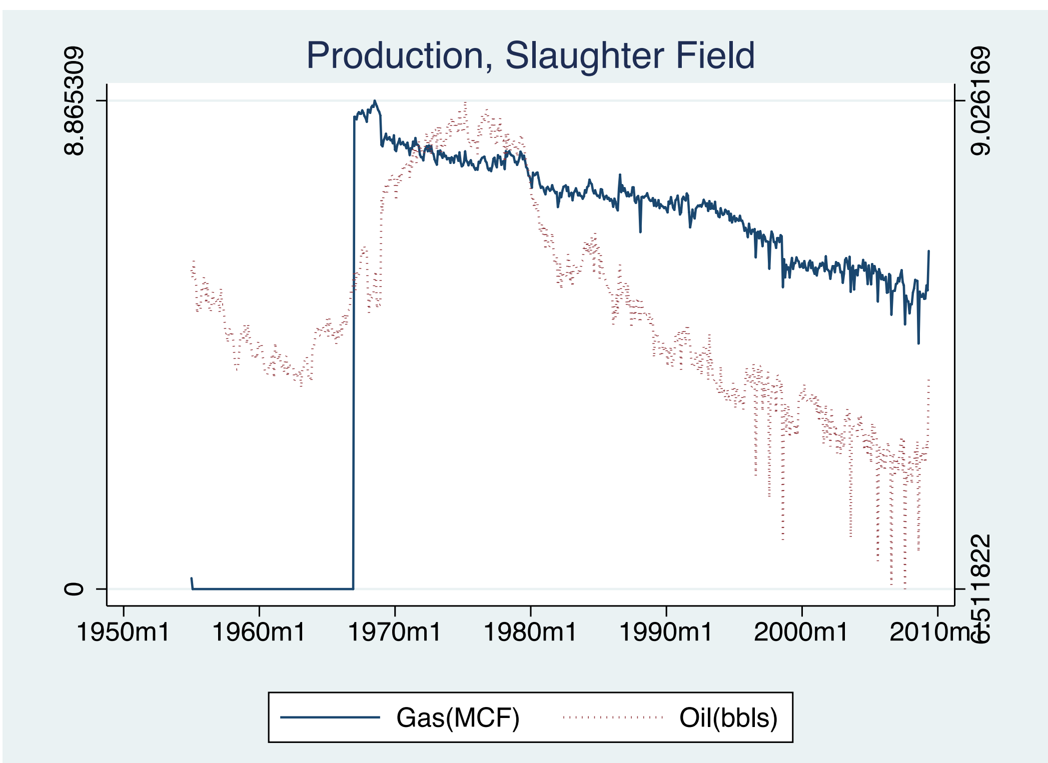


Figure 10: Log of well average production.

for production spans January 1955 to May 2009, injection data becomes extensive beginning in the late 1980s. The regression sample represents a subsample of production data for the month of January in 1990, 1995, 2000, and 2005.

Table 9: Summary statistics: Slaughter field

Full Sample variable	Obs.	Mean	Std. Dev.	Min	Max
log gas	61427	6.00	3.75	0	14.3
log oil	61427	8.07	2.19	0	13.5
log water	61427	7.81	5.30	0	15.1
GOR	61427	-2.06	3.07	-12.53	9.9
age	61427	6324.72	7100.95	0	39870.0
age sq	61427	9.04e+07	1.75e+08	0	1.59e+09
well count	61427	28.04	48.21	1	652.0
well count sq.	61427	3110.75	10426.54	1	425104.0
total depth	430431	608.42	1721.06	0	12384.0
lease acreage	421938	3507.56	2803.63	0	8684.3
Regression Sample					
log gas	387	6.37	3.16	0	14.0
log oil	387	7.54	2.25	0	12.5
log water	387	8.01	4.76	0	14.3
GOR	387	-1.17	2.15	-9.25	5.4
age	387	8384.34	7521.19	0	36862.0
age sq	387	1.27e+08	2.16e+08	0	1.36e+09
well count	387	23.39	41.65	1	243.0
well count sq.	387	2276.80	7588.24	1	59049.0
total depth	6125	306.90	1290.07	0	10700.0
lease acreage	6055	3673.79	2834.06	0	8684.3

Notes: This table reports summary statistics tests for the HPDI data available for Slaughter field in the top panel, and for the regression sample in the bottom panel. The regression sample represents a subsample of production data for the month of January in 1990, 1995, 2000, and 2005. Note that injection observations far outnumber production observations because injection is reported at the well level. In all regressions, injection is summed over the past year within a half mile of each production observation. Log gas, log oil, and log water are log of one plus monthly lease level gas, oil and water production, respectively. Water and oil production are reported in log barrels, while gas is reported in log thousand cubic feet. GOR is gas to oil ratio, which is also logged. Age is the age of the most recent well drilled on the lease. Well count is the number of active producing wells on the lease. Total depth is the total depth of the most recently completed well on the lease. Lease acreage is the only variable not provided by HPDI, it is taken from the Texas Railroad Commission and measures the area of the lease.

Results

This section presents results from cross sectional and panel models of Slaughter field, with and without injection. Cross sectional models are valuable in measuring how the spillover parameter evolves with time, but are limited in the sense that they cannot account for unobserved fixed effects specific to location. Panel models allow for unbiased estimates when spatial fixed effects are present. Both cross-sectional and panel models are run with and without injection. The goal of the various models is to demonstrate a race to extract when ownership is comparatively less secure–that is, well operators should increase their rate of extraction leaving less in the ground for the future. As a descriptive exercise I first run specifications of my statistical model with a local concentration index with a half-mile radius ($herf$) as the variable of interest. These results are presented in table 10 and include lease-level fixed effects.

Table 10: Regression on local Herfindahl index

	liq OLS	GLS	gas OLS	GLS
constant	0.000	0.126	0.000	-0.087
	(0.016)	(0.064)	(0.057)	(0.134)
wcnt	0.013	0.018	0.052	-0.006
	(0.003)	(0.033)	(0.013)	(0.099)
wcnt2	-0.006	0.015	-0.020	0.021
	(0.001)	(0.015)	(0.006)	(0.044)
age	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.592	0.618	0.935	1.608
	(0.038)	(0.131)	(0.130)	(0.415)
ginj	-0.010	-0.045	-0.026	0.060
	(0.002)	(0.018)	(0.009)	(0.048)
winj	0.007	0.036	0.004	-0.026
	(0.002)	(0.019)	(0.008)	(0.051)
herf	-1.469	-4.357	-2.845	-19.149
	(0.571)	(2.942)	(1.939)	(9.580)
moran	10990.750		1351.197	
ρ		0.037		0.039
		(0.001)		(0.001)

Notes: OLS columns represent standard OLS regression. GLS columns weight by inverse distance to account for spatial autocorrelation. *herf* is the coefficient for a local Herfindahl concentration index with a half-mile bandwidth.

Moran-I statistics for both oil and natural gas show significant evidence of spatial autocorrelation.[19] Because the productivity of a lease is driven by geological characteristics, which are unobserved in the dataset but likely to be highly correlated over space, it is expected that the residuals are positively correlated in space. According to OLS specifications for both oil and natural gas, drilling another well on the lease tends to increase production, but at a diminishing rate (the square term is insignificant). The affect of the number of wells on the lease is estimated less precisely in the GLS specification–the standard errors are larger and the coefficients are not distinguishable from zero. The age of the latest well on the lease does not

[19]The Moran statistics are calculated with the same inverse distance weight matrix used in the statistical models of the next sections.

meaningfully impact lease-level production. Water injection enhances oil recovery, but does not have a statistically noticeable affect on natural gas recovery. Gas injection is negatively correlated with oil recovery, and has little effect on gas recovery when controlling for spatial autocorrelation in the GLS specification. It appears then, that gas injection is not very successful *across* leases in secondary recovery in Slaughter field.

The variable of interest in the specification is the local Herfindahl concentration index. As the concentration index falls, the rights to the resource *in situ* become less secure, and the producer should extract at a higher rate. We, therefore, expect a negative correlation between concentration and lease-level extraction, and indeed this is what table 10 indicates. It can be argued that concentration is endogenous because naturally more productive areas face fiercer competition and thus lower concentration.[20] By explicitly accounting for how the unobserved productivity is correlated in space through the GLS specification, I hope to attenuate some of these problems; and yet concentration is a choice variable for the economic agents involved. Therefore in the rest of the paper, I use a different strategy to uncover a race to extract. Treating the structure of ownership as predetermined, I investigate how neighbor's pumping affects own extraction. With no strategic interaction, the estimated effect should be negative. A race to extract will manifest itself as a positive correlation between neighbor's extraction and own extraction.

[20]It could just as easily be the case that it is the more productive areas that are monopolized by one owner; but this bias works in my favor, with the parameter for concentration taking a lower bound to the true value.

Cross-sectional Model: Single Inverse Distance Weight Matrix, No Injection

The initial cross sectional model I estimate is

$$y_i = \lambda * \sum w_{ij} y_j + \alpha + \beta_1 depth_i + \beta_2 wcnt_i + \beta_3 wcnt_i^2 + \beta_4 age_i + \beta_5 wtr_i + \epsilon_i \quad (11)$$

where the weights are given by the inverse of distance between leases i and j. The spillover parameter of interest is λ; estimates can be found in tables 11-14. The dependent variables I consider are the log of month lease-level oil and natural gas production. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal.

Table 11: Oil cross-section 1990s

	1990				1995			
	OLS	GLS	2SLS	GS2SLS	OLS	GLS	2SLS	GS2SLS
constant	3.347	3.473	3.353	2.381	2.339	2.241	2.318	2.102
	(0.105)	(1.503)	(0.474)	(1.072)	(0.107)	(1.567)	(0.493)	(0.994)
depth	0.000	-0.000	0.000	0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wcnt	0.060	0.056	0.060	0.066	0.053	0.052	0.054	0.065
	(0.002)	(0.074)	(0.008)	(0.017)	(0.002)	(0.035)	(0.008)	(0.014)
wcnt2	-0.018	-0.013	-0.018	-0.021	-0.018	-0.019	-0.019	-0.022
	(0.001)	(0.037)	(0.004)	(0.009)	(0.001)	(0.016)	(0.004)	(0.007)
age	-0.000	-0.000	-0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.166	0.210	0.166	0.227	0.279	0.370	0.278	0.246
	(0.011)	(0.147)	(0.023)	(0.036)	(0.011)	(0.188)	(0.031)	(0.050)
λ	0.002		0.002	0.002	0.002		0.002	0.002
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
moran	133.63				-18.67			
ρ		0.070		0.068		0.068		0.064
		(0.004)		(0.004)		(0.004)		(0.003)

Notes: The table presents results for cross-sectional regressions with logged lease level oil production as the dependent variable. The columns entitled "OLS" present results for equation 11. The "GLS" column omits the variables for spatial dependence, but allows errors to be correlated according to the same inverse distance weight matrix. The column "2SLS" is estimated according to equation 11, but with y_j instrumented with $w_{ij} * X_j$, with $X_j = \{depth_j, wcnt_j, wcnt_j^2, age_j, wtr_j\}$. Finally, the "GS2SLS" column presents results from the model given by equation 9. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal. λ is the parameter estimate for the spatial autoregressive lag; ρ is the parameter for the spatial autoregressive error. The model is estimated using a single inverse-distance spatial weighting matrix.Standard errors are reported on the line underneath coefficient estimates.

Table 12: Oil cross-section 2000s

	2000				2005			
	OLS	GLS	2SLS	GS2SLS	OLS	GLS	2SLS	GS2SLS
constant	3.621	3.660	3.608	3.712	3.092	4.095	3.071	4.192
	(0.073)	(2.191)	(0.369)	(1.313)	(0.094)	(1.583)	(0.448)	(1.416)
depth	-0.000	0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wcnt	0.063	0.064	0.063	0.068	0.060	0.064	0.061	0.065
	(0.001)	(0.030)	(0.006)	(0.015)	(0.002)	(0.048)	(0.008)	(0.017)
wcnt2	-0.020	-0.022	-0.020	-0.022	-0.020	-0.021	-0.020	-0.022
	(0.000)	(0.016)	(0.003)	(0.008)	(0.001)	(0.021)	(0.004)	(0.009)
age	0.000	-0.000	0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.171	0.214	0.170	0.161	0.210	0.227	0.209	0.157
	(0.007)	(0.084)	(0.024)	(0.062)	(0.009)	(0.161)	(0.030)	(0.064)
λ	0.001		0.001	0.001	0.001		0.001	0.002
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
moran	-81.27				-143.92			
ρ		0.103		0.099		0.106		0.093
		(0.009)		(0.008)		(0.011)		(0.009)

Notes: The table presents results for cross-sectional regressions with logged lease level oil production as the dependent variable. The columns entitled "OLS" present results for equation 11. The "GLS" column omits the variables for spatial dependence, but allows errors to be correlated according to the same inverse distance weight matrix. The column "2SLS" is estimated according to equation 11, but with y_j instrumented with $w_{ij} * X_j$, with $X_j = \{depth_j, wcnt_j, wcnt_j^2, age_j, wtr_j\}$. Finally, the "GS2SLS" column presents results from the model given by equation 9. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal. λ is the parameter estimate for the spatial autoregressive lag; ρ is the parameter for the spatial autoregressive error. The model is estimated using a single inverse-distance spatial weighting matrix. Standard errors are reported on the line underneath coefficient estimates.

Table 13: Gas cross-section 1990s

	1990				1995			
	OLS	GLS	2SLS	GS2SLS	OLS	GLS	2SLS	GS2SLS
constant	5.088	6.047	5.335	5.335	4.1651	7.669	4.182	4.105
	(0.171)	(2.486)	(0.856)	(0.856)	(0.165)	(3.189)	(0.818)	(1.140)
depth	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wcnt	0.059	0.055	0.058	0.058	0.044	0.087	0.043	0.036
	(0.003)	(0.062)	(0.013)	(0.013)	(0.003)	(0.059)	(0.012)	(0.014)
wcnt2	-0.014	-0.012	-0.013	-0.013	-0.011	-0.031	-0.011	-0.009
	(0.002)	(0.031)	(0.006)	(0.006)	(0.002)	(0.029)	(0.006)	(0.007)
age	0.000	0.000	0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.104	0.109	0.105	0.105	0.366	0.128	0.367	0.457
	(0.019)	(0.182)	(0.038)	(0.038)	(0.017)	(0.152)	(0.049)	(0.049)
λ	0.000		0.000	0.000	-0.000		-0.000	-0.001
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
moran	886.068				230.246			
ρ		0.000		0.000		0.039		0.039
		(0.001)		(0.001)		(0.002)		(0.002)

Notes: The table presents results for cross-sectional regressions with logged lease level gas production as the dependent variable. The columns entitled "OLS" present results for equation 11. The "GLS" column omits the variables for spatial dependence, but allows errors to be correlated according to the same inverse distance weight matrix. The column "2SLS" is estimated according to equation 11, but with y_j instrumented with $w_{ij} * X_j$, with $X_j = \{depth_j, wcnt_j, wcnt_j^2, age_j, wtr_j\}$. Finally, the "GS2SLS" column presents results from the model given by equation 9. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal. λ is the parameter estimate for the spatial autoregressive lag; ρ is the parameter for the spatial autoregressive error. The model is estimated using a single inverse-distance spatial weighting matrix.Standard errors are reported on the line underneath coefficient estimates.

Table 14: Gas cross-section 2000s

	2000				2005			
	OLS	GLS	2SLS	GS2SLS	OLS	GLS	2SLS	GS2SLS
constant	2.919	6.834	3.094	1.644	2.679	1.449	2.684	2.023
	(0.215)	(6.812)	(1.134)	(1.352)	(0.207)	(4.525)	(1.024)	(1.362)
depth	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wcnt	0.064	0.165	0.061	0.046	0.058	0.157	0.058	0.048
	(0.004)	(0.113)	(0.018)	(0.019)	(0.004)	(0.097)	(0.018)	(0.019)
wcnt2	-0.017	-0.053	-0.016	-0.010	-0.018	-0.062	-0.018	-0.015
	(0.002)	(0.053)	(0.009)	(0.009)	(0.002)	(0.050)	(0.009)	(0.010)
age	0.000	-0.000	0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.323	0.054	0.333	0.446	0.330	0.576	0.331	0.437
	(0.023)	(0.361)	(0.074)	(0.070)	(0.021)	(0.332)	(0.069)	(0.069)
λ	0.001		0.000	0.000	0.001		0.001	0.000
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
moran	385.388				-23.270			
ρ		0.046		0.062		0.062		0.064
		(0.004)		(0.004)		(0.009)		(0.009)

Notes: The table presents results for cross-sectional regressions with logged lease level gas production as the dependent variable. The columns entitled "OLS" present results for equation 11. The "GLS" column omits the variables for spatial dependence, but allows errors to be correlated according to the same inverse distance weight matrix. The column "2SLS" is estimated according to equation 11, but with y_j instrumented with $w_{ij} * X_j$, with $X_j = \{depth_j, wcnt_j, wcnt_j^2, age_j, wtr_j\}$. Finally, the "GS2SLS" column presents results from the model given by equation 9. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal. λ is the parameter estimate for the spatial autoregressive lag; ρ is the parameter for the spatial autoregressive error. The model is estimated using a single inverse-distance spatial weighting matrix. Standard errors are reported on the line underneath coefficient estimates.

Tables 11-14 present regression results of the model for the month of January in 1990, 1995, 2000, and 2005. The columns entitled "OLS" present results for equation 11. The "GLS" column omits the variables for spatial dependence, but allows errors to be correlated according to the same inverse distance weight matrix. The column "2SLS" is estimated according to equation 11, but with y_j instrumented with $w_{ij} * X_j$, with $X_j = \{depth_j, wcnt_j, wcnt_j^2, age_j, wtr_j\}$. Finally, the "GS2SLS" column presents results from the model given by equation 9.

There are two parameters of particular interest in Table 11 and Table 12 . λ, the parameter for the spatial autoregressive lag, conflates the geophysical and strategic effects of neighbor's extraction. As mentioned earlier, the geophysical effects are expected to be wholly negative: own extraction causes a cone of depression to extend out from the well and causes oil to migrate from nearby leases, reducing production at other nearby leases. Everyone knows this, and so neighbors react by extracting at a higher rate to counteract the affects of the nearby pumping. In this case the strategic effect is positive and may be large enough to countervail the negative geophysical effects. The other parameter, ρ, measures the spatial autocorrelation of the errors. *A priori*, it is expected that $\rho > 0$ because it captures unobserved geological variables which are positively correlated through space.

Conditional on inverse distance weighting, I find consistent evidence of positive spatial autocorrelation in errors. Assuming that unobserved geological productivity is spatially correlated, these estimates are immanently reasonable. The parameter for the spatial autoregressive lag (λ) is also positive and significant across specifications. This indicates an increase in neighbors' production (or a diminution in the distance to neighbor's lease) results in an increase in own oil production, which is consistent with a race to extract.

The story for lease level natural gas production is not as clear cut. Estimates for the coefficient on spatial autoregressive lag (λ) and spatial autoregressive error (ρ)

are not as close across specifications or cross-sections. There is no evidence of autocorrelation of errors in 1990; the coefficient is positive and significant thereafter. The preferred specification for measuring λ is the GS2SLS estimate, which is negative and significant in 1995 but statistically non-distinguishable from zero in all other years.

Cross sectional models may be flawed because it is unlikely that the independent variables used in estimation are truly exogenous. Take for example the variable *wcnt*. It is easy to imagine that $E(wcnt_i * \epsilon_i) \neq 0$: if a lease is especially productive, or the lease owner expects that it will be, then more wells may be drilled. It is possible to control for these unobserved lease-level time invariant productivity differences by pooling the data and estimating the model adding fixed effects.

Cross-sectional Model: Simultaneous Inverse Distance Weight Matrices, With Injection

In this model the weight matrices differ according to whether plots are operated by the same owners or different owners. Weights are given as the inverse distance when two plots have a common operator (the "friendly" weight matrix); similarly they are also give as the inverse distance when two plots have different operators (the "unfriendly" weight matrix). Results from these regressions are presented in tables 15-18 for Slaughter field.

Table 15: Oil cross-section 1990s, simultaneous weighting with injection

	1990				1995			
	OLS	GLS	2SLS	GS2LS	OLS	GLS	2SLS	GS2LS
constant	1.960	5.021	1.943	2.218	0.664	6.206	0.664	-1.313
	(0.102)	(66.366)	(1.013)	(1.709)	(0.105)	(3.627)	(0.974)	(1.919)
depth	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wcnt	0.058	0.104	0.058	0.042	0.051	0.052	0.051	0.050
	(0.002)	(0.118)	(0.008)	(0.009)	(0.002)	(0.049)	(0.008)	(0.012)
wcnt2	-0.018	-0.043	-0.018	-0.010	-0.018	-0.016	-0.018	-0.016
	(0.001)	(0.060)	(0.004)	(0.004)	(0.001)	(0.023)	(0.004)	(0.005)
age	0.000	-0.000	0.000	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.151	-0.141	0.151	0.174	0.266	0.323	0.267	0.205
	(0.011)	(0.242)	(0.023)	(0.027)	(0.011)	(0.214)	(0.033)	(0.034)
ginj	-0.012	0.130	-0.013	0.021	-0.012	0.029	-0.014	-0.091
	(0.006)	(0.999)	(0.031)	(0.090)	(0.006)	(0.085)	(0.029)	(0.101)
winj	0.122	-0.421	0.122	0.117	0.130	-0.205	0.129	0.334
	(0.005)	(4.355)	(0.072)	(0.161)	(0.005)	(0.272)	(0.072)	(0.181)
λ_F	0.003		0.003	0.002	0.001		0.001	0.001
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
λ_U	0.001		0.001	0.000	0.001		0.001	0.001
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
Moran	-196.074				-384.907			
ρ		-0.234		0.219		0.118		0.128
		(0.035)		(0.039)		(0.024)		(0.011)

Notes: The table presents results for cross-sectional regressions with logged lease level oil production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. λ_F is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator; λ_U is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators. The canonical model is given by equation 12. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal; *winj* is water injected within a half-mile of the lease in the past year; *ginj* is the gas injected within a half mile of the producing lease in the past year; ρ is the parameter for the spatial autoregressive error. Standard errors are reported on the line underneath coefficient estimates.

Table 16: Oil cross-section 2000s, simultaneous weighting with injection

	2000				2005			
	OLS	GLS	2SLS	GS2LS	OLS	GLS	2SLS	GS2LS
constant	3.309	6.126	3.310	4.651	3.019	3.661	3.024	3.406
	(0.070)	(22.653)	(0.714)	(1.028)	(0.092)	(25.114)	(0.934)	(1.654)
depth	0.000	-0.000	0.000	-0.000	0.000	-0.000	0.000	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
wcnt	0.062	0.074	0.062	0.075	0.060	0.063	0.060	0.062
	(0.001)	(0.198)	(0.006)	(0.009)	(0.002)	(0.196)	(0.008)	(0.013)
wcnt2	-0.020	-0.024	-0.020	-0.025	-0.020	-0.015	-0.020	-0.020
	(0.000)	(0.098)	(0.003)	(0.004)	(0.001)	(0.097)	(0.004)	(0.006)
age	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.157	0.046	0.158	0.160	0.201	0.029	0.200	0.170
	(0.007)	(0.619)	(0.025)	(0.026)	(0.009)	(0.667)	(0.031)	(0.030)
ginj	0.036	0.056	0.035	0.062	0.044	0.008	0.043	0.021
	(0.004)	(0.638)	(0.019)	(0.040)	(0.005)	(0.458)	(0.025)	(0.070)
winj	0.013	-0.156	0.012	-0.102	-0.008	0.026	-0.009	0.019
	(0.003)	(1.626)	(0.050)	(0.093)	(0.005)	(1.574)	(0.065)	(0.139)
λ_F	0.001		0.001	0.001	0.001		0.001	0.001
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
λ_U	0.001		0.001	0.001	0.001		0.001	0.001
	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
Moran	-300.429				-238.674			
ρ		0.252		0.225		0.198		0.182
		(0.038)		(0.028)		(0.022)		(0.044)

Notes: The table presents results for cross-sectional regressions with logged lease level oil production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. λ_F is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator; λ_U is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators. The canonical model is given by equation 12. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal; *winj* is water injected within a half-mile of the lease in the past year; *ginj* is the gas injected within a half mile of the producing lease in the past year; ρ is the parameter for the spatial autoregressive error. Standard errors are reported on the line underneath coefficient estimates.

Table 17: Gas cross-section 1990s, simultaneous weighting with injection

	1990				1995			
	OLS	GLS	2SLS	GS2LS	OLS	GLS	2SLS	GS2LS
constant	8.981	9.260	8.995	9.489	6.756	11.821	6.758	2.866
	(0.163)	(12.000)	(1.700)	(2.020)	(0.156)	(11.600)	(1.490)	(3.320)
depth	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wcnt	0.061	0.090	0.061	0.047	0.042	0.072	0.041	0.032
	(0.003)	(0.063)	(0.013)	(0.015)	(0.003)	(0.063)	(0.012)	(0.014)
wcnt2	-0.016	-0.029	-0.015	-0.009	-0.009	-0.025	-0.009	-0.006
	(0.001)	(0.036)	(0.006)	(0.007)	(0.001)	(0.029)	(0.006)	(0.007)
age	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.096	0.071	0.098	0.156	0.349	0.240	0.348	0.493
	(0.018)	(0.098)	(0.038)	(0.042)	(0.016)	(0.206)	(0.050)	(0.055)
ginj	0.133	0.010	0.133	0.122	0.122	0.154	0.122	-0.041
	(0.009)	(0.189)	(0.052)	(0.082)	(0.009)	(0.128)	(0.046)	(0.156)
winj	-0.255	-0.259	-0.258	-0.184	-0.172	-0.467	-0.172	0.046
	(0.008)	(0.761)	(0.116)	(0.155)	(0.008)	(0.677)	(0.108)	(0.290)
λ_F	0.002		0.002	0.004	0.000		0.000	0.000
	(0.000)		(0.001)	(0.002)	(0.000)		(0.001)	(0.001)
λ_U	-0.001		-0.000	-0.003	-0.002		-0.002	-0.001
	(0.000)		(0.001)	(0.001)	(0.000)		(0.001)	(0.001)
Moran	372.317				-21.177			
ρ		0.031		0.031		0.054		0.062
		(0.001)		(0.001)		(0.005)		(0.005)

Notes: The table presents results for cross-sectional regressions with logged lease level gas production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. λ_F is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator; λ_U is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators. The canonical model is given by equation 12. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal; *winj* is water injected within a half-mile of the lease in the past year; *ginj* is the gas injected within a half mile of the producing lease in the past year; ρ is the parameter for the spatial autoregressive error. Standard errors are reported on the line underneath coefficient estimates.

Table 18: Gas cross-section 2000s, simultaneous weighting with injection

	2000				2005			
	OLS	GLS	2SLS	GS2LS	OLS	GLS	2SLS	GS2LS
constant	7.864	9.423	7.881	7.290	5.901	4.153	5.894	5.071
	(0.203)	(19.644)	(2.080)	(3.398)	(0.196)	(19.121)	(2.002)	(3.306)
depth	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wcnt	0.068	0.128	0.067	0.057	0.066	0.073	0.066	0.054
	(0.004)	(0.064)	(0.017)	(0.018)	(0.004)	(0.111)	(0.017)	(0.019)
wcnt2	-0.018	-0.052	-0.018	-0.015	-0.020	-0.025	-0.021	-0.018
	(0.002)	(0.039)	(0.009)	(0.009)	(0.002)	(0.051)	(0.009)	(0.010)
age	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.307	0.499	0.311	0.426	0.283	0.279	0.286	0.418
	(0.021)	(0.295)	(0.075)	(0.062)	(0.020)	(0.366)	(0.071)	(0.054)
ginj	0.130	0.212	0.135	0.033	0.138	0.074	0.134	0.059
	(0.012)	(0.145)	(0.057)	(0.126)	(0.011)	(0.191)	(0.055)	(0.135)
winj	-0.354	-0.510	-0.351	-0.385	-0.224	-0.076	-0.229	-0.205
	(0.011)	(1.082)	(0.145)	(0.274)	(0.010)	(1.079)	(0.140)	(0.274)
λ_F	0.003		0.003	0.002	0.003		0.003	0.001
	(0.001)		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)
λ_U	-0.000		-0.001	0.002	-0.001		-0.001	0.001
	(0.000)		(0.001)	(0.001)	(0.000)		(0.001)	(0.001)
Moran	155.359				-142.382			
ρ		0.068		0.080		0.123		0.107
		(0.005)		(0.005)		(0.013)		(0.022)

Notes: The table presents results for cross-sectional regressions with logged lease level gas production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. λ_F is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator; λ_U is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators. The canonical model is given by equation 12. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal; *winj* is water injected within a half-mile of the lease in the past year; *ginj* is the gas injected within a half mile of the producing lease in the past year; ρ is the parameter for the spatial autoregressive error. Standard errors are reported on the line underneath coefficient estimates.

Focusing first on GS2SLS specifications of oil production in Slaughter Field, there does not appear to be a clear pattern in the relationship between λ_F and λ_U; friendly spillovers are positive and significant across years, larger than unfriendly spillovers in 1990 and equal to unfriendly spillovers in 2005. There is evidence that the errors are strongly correlated over space. There is also significant autocorrelation in Slaughter gas production. Friendly spillovers are positive and significant in 1990; no other spillover parameters are significant at the 95% confidence level, and point estimates between friendly spillovers and unfriendly spillovers are close.

Panel Model: Single Inverse Distance Weight Matrix, No Injection

Table 19 presents fixed effects estimated for the pooled data. The fixed effects models for lease-level gas production indicate positive and significant spatial autocorrelation in the time-varying aspect of the errors. There is also evidence of a race to extract in natural gas production, as the estimates for λ are positive and stable, although the preferred GS2SLS estimate is insignificant.

Table 19: Fixed effects panel regressions

	Oil				Gas			
	OLS	GLS	2SLS	GS2SLS	OLS	GLS	2SLS	GS2SLS
constant	3.915	5.401	3.347	6.053	3.139	4.461	2.872	3.183
	(0.059)	(0.517)	(0.212)	(1.902)	(0.096)	(0.680)	(0.348)	(0.992)
wcnt	0.088	0.090	0.093	0.074	0.093	0.026	0.095	0.087
	(0.001)	(0.019)	(0.004)	(0.028)	(0.002)	(0.028)	(0.007)	(0.010)
wcnt2	-0.033	-0.044	-0.034	-0.029	-0.034	-0.002	-0.035	-0.030
	(0.000)	(0.010)	(0.002)	(0.013)	(0.001)	(0.013)	(0.003)	(0.004)
age	0.000	-0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
λ	0.002		0.002	0.000	0.001		0.001	0.001
	(0.000)		(0.000)	(0.001)	(0.000)		(0.000)	(0.001)
moran	5049.479				7359.250			
ρ		0.031		0.031		0.015		0.015
		(0.001)		(0.001)		(0.000)		(0.000)

Notes: The table presents results fixed effects panel regressions with log gas and log oil production as the dependent variable and a single inverse distance spatial weight matrix with coefficient λ. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal; *winj* is water injected within a half-mile of the lease in the past year; *ginj* is the gas injected within a half mile of the producing lease in the past year; ρ is the parameter for the spatial autoregressive error. Standard errors are reported on the line underneath coefficient estimates.

Turning to the oil results, it is apparent that the OLS and 2SLS estimation yield parameter estimates for λ that are positive and significant and very close to previous cross-sectional estimates. The parameter for the spatial autocorrelation in errors (ρ) is a third to half the size of estimates in Tables 11 and 12. This difference cannot be wholly attributed to the importance of lease fixed effects in controlling for spatial autocorrelation,[21] nevertheless, it is expected for the error to attenuate because fixed effects diminishes the unexplained variation in the model. What is striking is the insignificance of λ in the GS2SLS specification when controlling for lease-level fixed effects.

There are two ways to interpret the insignificance. The first is that the spatial dependence in production is a statistical illusion. Cross sectional GS2SLS estimation is not powerful enough to properly distinguish true production spillovers from time invariant differences in lease productivities. But going deeper, fixed effects estimation differences out the variation in ownership structures, which are precisely the effects that I seek to isolate.

Panel Model: Separate Inverse Distance Weights, No Injection

To isolate the spillovers that result from different ownership structures, I estimate the fixed effects models in two specifications which differ in the weight matrix used. In particular, I separately estimate

$$\begin{aligned} y &= X\beta_i + \lambda_i W_i y + u \\ u_i &= \rho_i M u_i + \epsilon_i \end{aligned} \tag{12}$$

where $i = \{F, U\}$ indexes the weight matrix to be used in the specification. M represents the inverse distance weighting matrix used previously. W_F represents the inverse distance weights, but take values only when the neighboring leases are

[21] The models have different regressors. The variable *depth* is time invariant and cannot be used in fixed-effect specification.

produced by a common operator (*i.e.*, the weights take values only for "friendly" leases.) Similarly, W_U takes inverse distance values when neighboring leases do no share a common operator (the leases are said to be "unfriendly") *A priori*, we would expect leases with common owners to more fully account for the spillovers in production , so that $\lambda_F < \lambda_U$. Equation 12 thus provides a testable hypothesis.

Result from estimation of models with weight matrices given according to Equation 12 are given in Tables 20 and 21. Again, the preferred specification is GS2SLS. For oil production, the parameter measuring the spatial spillover among friendly wells is positive and statistically significant. What is surprising is that the parameter for friendly spillover is larger and significantly different than the parameter for unfriendly spillovers (which is not significantly different from 0). Results are qualitatively similar for natural gas production. Combined, these results seem to indicate that leases under common stewardship are more likely to engage in a race to extract.

Table 20: Fixed effects, separate weighting: oil

	Friendly			Unfriendly		
	OLS	2SLS	GS2SLS	OLS	2SLS	GS2SLS
constant	5.617	5.630	5.630	4.353	4.110	5.482
	(0.061)	(0.112)	(0.112)	(0.064)	(0.258)	(1.696)
wcnt	0.078	0.078	0.078	0.082	0.084	0.071
	(0.001)	(0.005)	(0.005)	(0.001)	(0.005)	(0.020)
wcnt2	-0.029	-0.029	-0.029	-0.032	-0.032	-0.030
	(0.000)	(0.002)	(0.002)	(0.000)	(0.002)	(0.009)
age	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
λ	0.004	0.004	0.004	0.001	0.002	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
moran	5163.20			6276.67		
ρ			0.000			0.031
			(0.001)			(0.000)

Notes: The table presents results fixed effects panel regressions with log oil production as the dependent variable and a single inverse distance spatial weight matrix with coefficient λ. In "Friendly" columns this weight matrix takes inverse distance values only when neighboring leases have common operators; "Unfriendly is the opposite;The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal; *winj* is water injected within a half-mile of the lease in the past year; *ginj* is the gas injected within a half mile of the producing lease in the past year; ρ is the parameter for the spatial autoregressive error. Standard errors are reported on the line underneath coefficient estimates.

Table 21: Fixed Effects, Separate weighting: gas

	Friendly			Unfriendly		
	OLS	2SLS	GS2SLS	OLS	2SLS	GS2SLS
constant	3.822	3.883	4.007	4.632	4.319	5.067
	(0.088)	(0.162)	(0.166)	(0.098)	(0.397)	(0.772)
wcnt	0.094	0.093	0.090	0.083	0.085	0.085
	(0.001)	(0.007)	(0.006)	(0.002)	(0.008)	(0.009)
wcnt2	-0.033	-0.033	-0.032	-0.031	-0.032	-0.032
	(0.001)	(0.003)	(0.003)	(0.001)	(0.003)	(0.004)
age	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
λ	0.006	0.005	0.006	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
moran	5395			7552		
ρ			0.015			0.015
			(0.000)			(0.000)

Notes: The table presents results fixed effects panel regressions with log gas production as the dependent variable and a single inverse distance spatial weight matrix with coefficient λ. In "Friendly" columns this weight matrix takes inverse distance values only when neighboring leases have common operators; "Unfriendly is the opposite;The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated. Independent variables are *depth*, the total depth of the most recent well completed on the lease; *wcnt*, the number of active producing wells on the lease (and its square); *age*, the time since the most recent well was completed; *wtr*, the amount of water produced on the lease that month, a proxy for whether the lease is constrained by disposal; *winj* is water injected within a half-mile of the lease in the past year; *ginj* is the gas injected within a half mile of the producing lease in the past year; ρ is the parameter for the spatial autoregressive error. Standard errors are reported on the line underneath coefficient estimates.

One plausible explanation for the unexpected results is that no account has been made for how injection impacts recovery. Injection is more likely and more effective when contiguous leases are controlled by a common operator. If these types of leases are more successful in injection, then recovery across the leases may be highly correlated, contributing to what looks like spatial dependence. Moreover, the impact of injection will not be swept up with lease-level fixed effects because injection varies over time. This can be tested for by adding injection into the model. The variables for injection are defined as the sum of injection that occurred within a half-mile radius of the well within the past year.

Panel Model: Simultaneous Inverse Distance Weights, With Injection

The next model I estimate is

$$\begin{aligned} y &= X\beta + \lambda_F W_F y + \lambda_U W_U y + u \\ u &= \rho M u + \epsilon. \end{aligned} \tag{13}$$

Tables 22 and 23 present estimations controlling for gas and water injection for the spillover parameters λ_F and λ_U for oil and gas, respectively. This model is slightly different from equation 12 in that the effect of spillovers from friendly wells is estimated in the same model as spillovers from unfriendly wells. In the columns labeled "GS2SLS" in both tables 22 and 23 we see evidence of a race to extract, in that the spillover estimate for unfriendly wells comes up as positive and significant, and is larger than the estimate for the spillover from friendly wells. Take for example table 22. The spillovers from nearby leases managed by the same operator is insignificantly different from zero, evidence that the operator is fully accounting for the externality in production. Meanwhile, the spillover from unfriendly wells is significant and positive. After controlling for spatial autocorrelation, this indicates that when a neighboring competing operator increases production, you also tend to increase production–a classic race to extract.

Table 22: Fixed effects, simultaneous weighting: oil

	OLS	GLS	2SLS	GS2SLS
constant	-0.009	-0.014	-0.010	-0.015
	(0.014)	(0.235)	(0.015)	(0.006)
wcnt	0.013	0.140	0.013	0.016
	(0.003)	(0.039)	(0.005)	(0.013)
wcnt2	-0.007	-0.049	-0.007	-0.011
	(0.001)	(0.014)	(0.002)	(0.006)
age	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.542	0.136	0.543	0.335
	(0.033)	(0.138)	(0.036)	(0.051)
ginj	-0.006	0.071	-0.007	-0.023
	(0.002)	(0.020)	(0.005)	(0.008)
winj	0.009	-0.069	0.009	0.014
	(0.002)	(0.012)	(0.004)	(0.007)
λ_F	0.008		0.007	-0.006
	(0.001)		(0.002)	(0.005)
λ_U	0.004		0.004	0.005
	(0.000)		(0.000)	(0.000)
moran	201.341			
ρ		0.120		0.093
		(0.005)		(0.003)

Notes: The table presents results fixed effects panel regressions with logged lease level oil production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. λ_F is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator; λ_U is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated.

Table 23: Fixed effects, simultaneous weighting: gas

	OLS	GLS	2SLS	GS2SLS
constant	-0.013	-0.858	-0.019	-0.009
	(0.057)	(0.293)	(0.057)	(0.019)
wcnt	0.051	0.217	0.051	0.054
	(0.013)	(0.044)	(0.021)	(0.025)
wcnt2	-0.022	-0.036	-0.023	-0.019
	(0.006)	(0.020)	(0.010)	(0.011)
age	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.871	1.354	0.847	0.543
	(0.129)	(0.408)	(0.136)	(0.111)
ginj	-0.023	0.051	-0.022	0.008
	(0.010)	(0.074)	(0.021)	(0.021)
winj	0.011	-0.107	0.010	-0.013
	(0.008)	(0.052)	(0.018)	(0.017)
λ_F	0.007		0.004	0.010
	(0.004)		(0.006)	(0.007)
λ_U	0.004		0.005	0.004
	(0.001)		(0.002)	(0.001)
moran	-685.593			
ρ		0.068		0.045
		(0.003)		(0.002)

Notes: The table presents results fixed effects panel regressions with logged lease level gas production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. λ_F is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator; λ_U is the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated.

Panel Model: Simultaneous Inverse Distance Weights By Well Age, With Injection

We observe that there is a large "positive" spillover when neighboring wells are owned by competing operators– evidence of a race to extract. However, it is possible that it is not ownership per se that drives the results. Well age is an important variable to consider in modeling reservoir dynamics. Young wells are likely to have much more capacity to communicate with neighboring wells, than comparatively older wells, simply because, all things being equal, younger wells will be in higher energy parts of the reservoir and more potential for drawdown. To test for this, I allow the spillover parameter to be vary across wells of different ages, and these parameters are allowed to be different for both friendly and unfriendly operators. The model I estimate is

$$\begin{aligned} y &= X\beta + \textstyle\sum_a (\lambda_{F,a} W_F y_a + \lambda_{U,a} W_U y_a) + u \\ u &= \rho M u + \epsilon. \end{aligned} \tag{14}$$

Production at neighboring wells, y_a has been separated into 4 vectors depending on what age bin production falls. I arbitrarily choose 4 bins, so that each production been of well age represents a quartile, $a \in 1, 2, 3, 4$. We expect that the spillover parameter should decline with as neighboring wells increase in age. The decline occurs for two reasons. First, since the neighboring wells are older, there is likely less capacity at those wells for drawdown because with time the pressure in the surrounding reservoir and at the well face tends toward equilibrium. Second, it is reasonable to expect that the age of wells is correlated across space; therefore, when neighboring wells are older, it is likely that your own well is older, and that you have less capacity to adjust own production, although this effect would be accounted for to some degree by the linear term in age. Results for regressions with oil as the dependent variable are in table 24, while gas results are in table 25.

Table 24: Oil spillover by well age

	OLS	GLS	2SLS	GS2SLS
constant	0.005	0.236	-0.012	0.006
	(0.014)	(0.311)	(0.020)	(0.008)
wcnt	0.013	0.152	0.012	0.021
	(0.003)	(0.044)	(0.005)	(0.016)
wcnt2	-0.007	-0.048	-0.007	-0.015
	(0.001)	(0.017)	(0.002)	(0.007)
age	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.533	0.530	0.541	0.005
	(0.033)	(0.188)	(0.038)	(0.002)
ginj	-0.006	0.059	-0.007	-0.012
	(0.002)	(0.030)	(0.005)	(0.009)
winj	0.009	-0.055	0.007	0.012
	(0.002)	(0.023)	(0.004)	(0.008)
λ_{F1}	0.008		-0.004	0.030
	(0.001)		(0.007)	(0.004)
λ_{F2}	0.071		0.004	-0.000
	(0.057)		(0.000)	(0.000)
λ_{F3}	-0.545		-0.000	0.000
	(0.532)		(0.000)	(0.000)
λ_{F4}	-0.697		-0.000	0.000
	(0.883)		(0.000)	(0.000)
λ_{U1}	0.003		0.005	0.004
	(0.000)		(0.001)	(0.000)
λ_{U2}	0.032		0.010	0.024
	(0.008)		(0.026)	(0.016)
λ_{U3}	-0.060		-0.004	0.000
	(0.208)		(0.001)	(0.000)
λ_{U4}	-0.053		0.012	0.001
	(0.129)		(0.000)	(0.000)
moran	317.985			
ρ		0.125		0.117
		(0.005)		(0.004)

Notes: The table presents results fixed effects panel regressions with logged lease level oil production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. $\lambda_{F1} - \lambda_{F4}$ are the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator, and well the ages of the wells are in the quartile 1-4; $\lambda_{U1} - \lambda_{U4}$ are the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators and well the ages of the wells are in the quartile 1-4, respectively. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated.

Table 25: Gas spillover by well age

	OLS	GLS	2SLS	GS2SLS
constant	0.018	-2.332	-0.011	-0.060
	(0.055)	(1.335)	(0.046)	(0.026)
wcnt	0.049	-0.859	0.050	0.055
	(0.012)	(0.163)	(0.021)	(0.025)
wcnt2	-0.021	0.119	-0.022	-0.016
	(0.006)	(0.071)	(0.009)	(0.011)
age	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
wtr	0.864	-0.371	0.849	0.053
	(0.126)	(1.379)	(0.141)	(0.022)
ginj	-0.025	0.010	-0.022	0.023
	(0.009)	(0.227)	(0.021)	(0.022)
winj	0.014	0.118	0.010	-0.028
	(0.008)	(0.158)	(0.017)	(0.017)
λ_{F1}	0.005		0.001	0.011
	(0.004)		(0.013)	(0.008)
λ_{F2}	0.005		0.013	-0.020
	(0.136)		(0.002)	(0.008)
λ_{F3}	-0.101		0.001	-0.003
	(0.622)		(0.000)	(0.001)
λ_{F4}	-0.037		0.010	-0.021
	(0.145)		(0.007)	(0.010)
λ_{U1}	0.000		0.004	0.008
	(0.001)		(0.002)	(0.002)
λ_{U2}	0.046		0.010	-0.020
	(0.014)		(0.049)	(0.028)
λ_{U3}	0.268		-0.014	-0.016
	(0.243)		(0.005)	(0.006)
λ_{U4}	0.069		0.017	-0.096
	(0.044)		(0.035)	(0.049)
moran	-725.961			
ρ		0.093		0.046
		(0.004)		(0.001)

Notes: The table presents results fixed effects panel regressions with logged lease level gas production as the dependent variable and separate inverse distance spatial weight matrices given according to whether neighboring leases are under common ownership. $\lambda_{F1} - \lambda_{F4}$ are the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have a common operator, and well the ages of the wells are in the quartile 1-4; $\lambda_{U1} - \lambda_{U4}$ are the estimated spatial autoregressive coefficient for the weight matrix that takes inverse distance values when leases have competing operators and well the ages of the wells are in the quartile 1-4, respectively. The column labeled "OLS" presents results when no instrumentation is made for the spatially lagged dependent variable, and no spatial structure is assumed in the errors; "GLS" is a specification, where the errors are assumed to be spatially autocorrelated; the column "2SLS" instruments for endogenous production with WX; "GS2SLS" instruments for endogenous production, and assume the errors to be spatially autocorrelated.

Looking first at results for oil, the Moran-I statistic indicates significant positive spatial autocorrelation in the errors, making inference on the OLS parameters untenable. After controlling for other covariates an, an extra day of production (*age*)

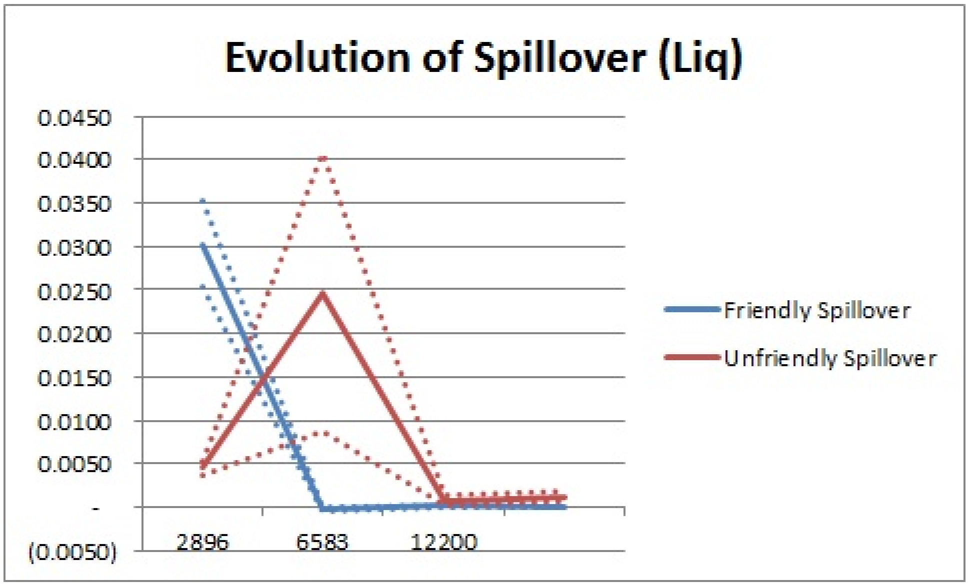


Figure 11: Table 24 plotted.

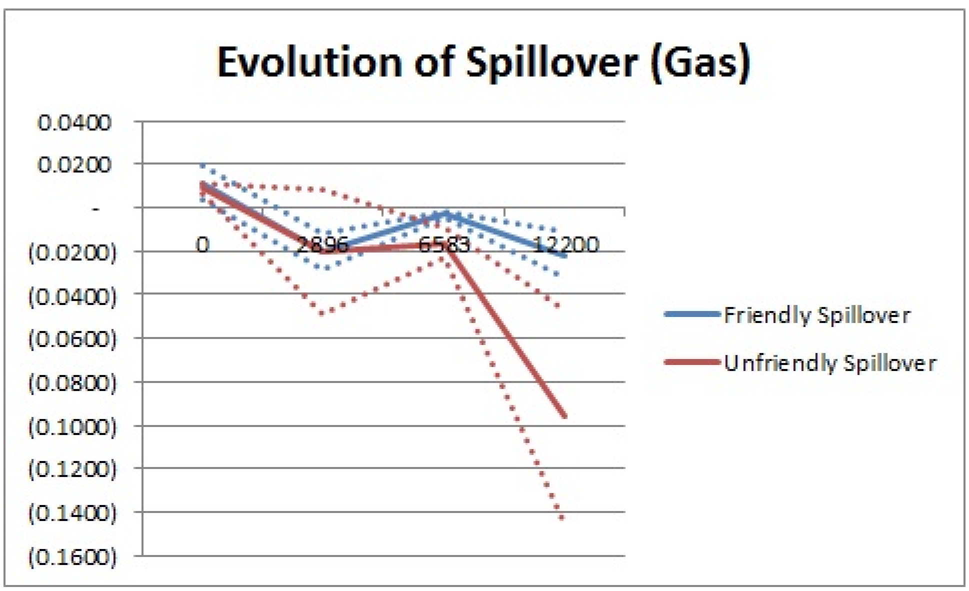


Figure 12: Table 25 plotted.

does not meaningfully affect oil production. The OLS and GS2SLS specifications also indicate that gas injection is negatively correlated with production when controlling for spillovers; however, local gas injection is positive in the GLS specification while local water injection is negative and statistically significant. The instability of the parameter estimates for injection is likely due to to the complex spatial dynamics of the reservoir. A sudden within drop off in production could precipitate a local within increase in injection to compensate; a positive relationship is also easy to explain. The countervailing pressures in injection explain why estimates are not significant in the GS2SLS specification. The specifications for oil do pick up significant positive autocorrelation in oil production between leases.

Results for explanatory variables for natural gas are similar to those for oil. The age of the well does not seem to significantly impact production; the number of wells on the lease increases production, although the negative square term, indicates that this is at a decreasing rate. Injection is generally insignificant. Injection is likely to be even more difficult to identify with gas production, since produced gas can be re-injected into the reservoir. The Moran-I statistic on the OLS regression indicates negative spatial correlation, while the GLS and GS2SLS specification pick up the expected positive spatial autocorrelation in errors.

The parameters for spatial autoregressive lag are graphed in figure 11 for oil and figure 12 for gas. The older the age of the neighboring well the more the spillover parameter should attenuate, and so we see a negative slope in the graphs in both cases. Additionally, there should be more of a race to extract when wells are owned by competing ("unfriendly") operators, and so we would expect that the unfriendly line lies above the friendly line on the graphs. This is generally the case for oil, with the exception being spillovers from wells within the first quartile of age. Unlike oil the spillovers parameters for natural gas are, with the exception of the first period, negative, here the friendly spillover tends to lie above estimates for unfriendly spillovers, and is closer to zero. Pinning down the interpretation of the spillover parameter requires a full spatial dynamic model of joint resource recovery, as well as controls for cumulative recovery, which is not attempted at present.

Conclusion

The spatial interaction between wells is an important consideration in efficiently draining oil an natural gas from expansive underground reservoirs. Yet previous research has shown that the present structure of lease-ownership in Texas impedes efficient field development because rights to the resources *in situ* are not fully delineated. This insecurity perverts economic incentives so that the resources are extracted too quickly, with too much of the rents depleted by costly excess capital. This paper exploits recent advances in spatial econometrics to quantify the production spillovers between leases. Results show evidence of a race to extract across a variety of specifications. The most extensive model show that after controlling for injection and fixed effects, consolidated ownership reduces spillovers and tends to slow the rate of extraction as compared to areas where ownership is highly fractured. These results are directly in line with economic theory.

Chapter V

POWER LAWS IN TEXAS OIL AND NATURAL GAS PRODUCTION

Introduction

Power laws are of growing interest to economists. They are relations of the type $Y = kX^{\alpha}$, where Y and X are two variables, and α is the power-law exponent, also known as the scaling parameter. Power laws are used to describe the distribution of firm sizes (Axtell 2001), the distribution of city sizes (Gabaix and Ioannides 2004), and most famously, the distribution of income and wealth (Atkinson and Piketty 2007), (Pareto 1896). More speculative laws include the distribution of stock price fluctuations and trading volume, as well as the relationship between the number or lines of state regulations and state population, and the number of links to a website and it's popularity (Gabaix 2009). There are also many proposed power laws outside the realm of economics. Most famously, Zipf finds that the frequency a word is used is inversely proportional to its rank (in terms of usage) (Zipf 1949).

In this paper, I examine the distribution of oil and natural gas recovery in Texas and find that recovery is power law distributed. In particular, I find that recovery for a lease of rank r is proportional to $1/r$, *i.e.*, high ranking leases have a disproportionate share of total recovery. In distributional terms, this implies that the probability that a lease recovers more than x barrels of oil is proportional to $1/x$. In particular, $P(Recovery > x) = k/x^{\alpha}$, with $\alpha \simeq 1$ for cumulative oil recovery. For cumulative natural gas recovery, $\alpha \simeq 1.3$. What this means, intuitively, is that the distribution of oil and natural gas wells have extremely fat tails, and that a small percentage of producers are responsible for a great percentage of recovery. These findings have policy relevance in terms of regulation of spills, as "large" spills are only possible for a few leases. Thus to prevent a BP-type spill regulation need not be universal. The same principle applies to royalty enforcement, as a great portion of royalty revenues come from a small percent of leases.

Data

Identifying power laws is very data-intensive. Power law behavior is most apparent in the tails of a distribution, which is also where there are the fewest observations. It, therefore, takes very large datasets to be able to distinguish between different distributions. The HPDI data set, which compiles time series for oil and natural gas production for 31 states as well as the federal offshore areas in the Gulf and the Pacific, is uniquely well-suited for the purpose. In this chapter, I limit my focus to the oil leases and gas wells in Texas (oil is reported at the lease level, gas at the well level). Time series for these data go back to as early as 1934. I focus on cumulative oil production, and cumulative gas production yielding 591,764 observations. Many of these wells are still active.

Moments and summary statistics for the distributions can be found in table 26. The first thing to notice is that both the oil and gas distribution are right-skewed: The mean is significantly higher than the median in both cases, and the sample estimate for skewness is large and positive in both cases. Secondly, and what is very suggestive of the power law behavior of the distribution, is that it spans nine orders of magnitude. The minimum cumulative oil and gas produced is 0, while the maximum in both cases is measured in the billions. While it could be argued that this dispersion is a result of the units chosen, parameter estimates of power law distributions are independent of the units of measure chosen.[22] For gas production 50% of the data lie within the first 4 orders of magnitude, for oil 50 % of observations lie within the first 3 orders of magnitude. The large sample estimate for kurtosis indicates that there is substantial weight in the tails of the distribution.

Table 26: Moments of sample distribution

	gas	oil
mean	533619.6	73111.9
median	10676	266
maximum	2.49e+9	1.18e+9
minimum	0	0
variance	4.11e+13	7.89e+12
skewness	216.79	271.97
kurtosis	66594.53	94687.69

Notes: Statistical moments for Texas hydrocarbon production.

Digression on Moments

The moment generating function for a power law distribution is given by

$$\langle x^m \rangle = \int_{x_{min}}^{\infty} x^m p(x) dx = \frac{\alpha - 1}{\alpha - 1 - m} x_{min}^m, \tag{15}$$

where p(x) and x_{min} are defined in equation 16. The important thing to notice is that for the power law distribution, moments become infinite unless $m < \alpha - 1$. In this chapter, I find evidence that $\alpha < 3$ for both oil and gas. This implies that there is no finite moment beyond the mean. Of course, in any finite sample, it is possible to calculate higher order moments such as variance, skewness and kurtosis. The

[22]Indeed, they are the only family of distributions where the parameters do not depend on the units of measurement, hence, they are known as *scale-free* or *scaling* distributions (**?**).

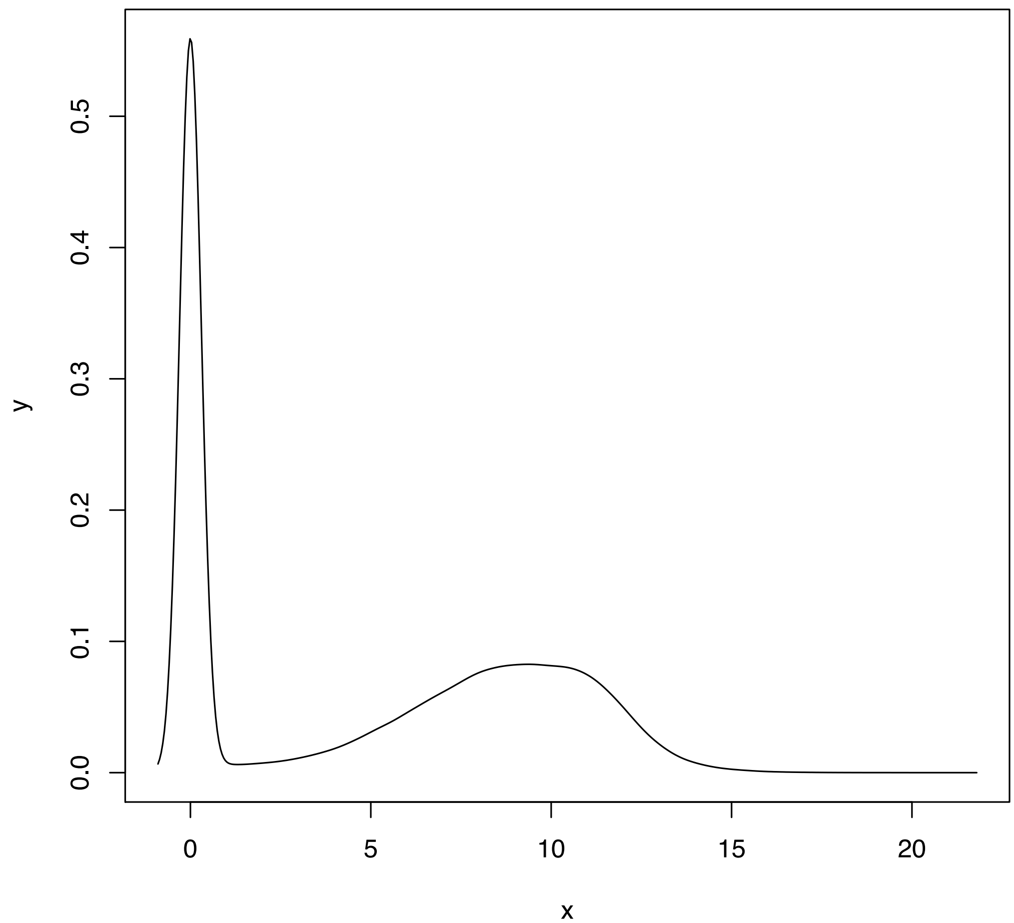


Figure 13: Distribution of Oil Productivities. x is log cumulative production in barrels, y is the probability.

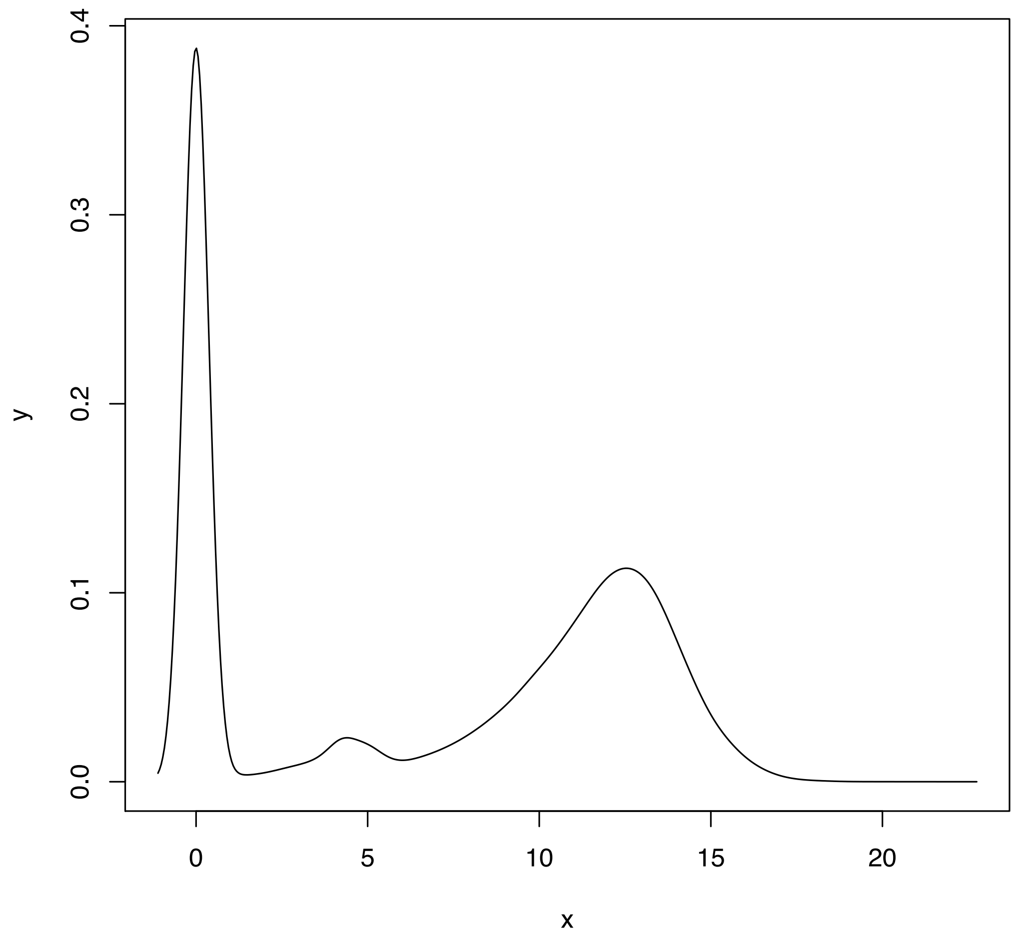


Figure 14: Distribution of Natural Gas Productivities. x is log cumulative production in MCF, y is the probability.

catch is that as the sample size is increased, the sample estimate for moments will increase, never converging to anything.

The reality is that there is a finite amount of oil and gas resources on the planet, a finite number of wells that can be drilled, and so practically speaking, the moments cannot increase without bound. Yet the power law still captures important aspects of the distribution. To make this point more clear I take an example from Newman (2005). The magnitude of flooding is thought to be power law distributed with $\alpha < 2$. In this case there is not even a well-defined mean for the distribution. It is possible to calculate the average flood from the historical data, but this is not particularly useful, because most of the data will be far from that average. The quantiles of the distribution are informative, however, which is why instead of talking about the average flood, we make reference to the Great Mississippi Flood of 1927.

Methods

To estimate the exponent of the presumed power law distribution of cumulative oil and natural gas recovery, I focus on the right tail of the data. In particular, I choose two samples. In one I limit observations to those in the top 5% as recommended by Gabaix (2009) of cumulative oil and natural gas recovery, resulting in two samples, each with 29,588 observations. In the second sample, I endogenously estimate the threshold using techniques described in the following paragraphs. In this chapter I use three methods for estimating the power law exponent. The most straightforward is by using the method of maximum likelihood.

The probability distribution function for a power law is given by

$$p(x) = \frac{\alpha - 1}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\alpha}, \tag{16}$$

where α is the parameter of interest. The parameter x_{min} is the threshold at which the power law behavior begins. Notice that as $x_{min} -> 0$ the probabilities diverge, and so power laws cannot match real world data over the entirety of the distribution. It is for this reason that power laws are normally relegated to modeling the tails of the distribution. The log likelihood function is then written as

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{n}\left[\ln(\alpha-1) - \ln x_{min} - \alpha \ln \frac{x_i}{x_{min}}\right] \\ &= n\ln(\alpha-1) - n\ln x_{min} - \alpha \sum_{i=1}^{n} \ln \frac{x_i}{x_{min}}. \end{aligned} \tag{17}$$

By taking the derivative of the log likelihood function with respect to α, setting it equal to zero and solving for α, the maximum likelihood estimate is

$$\alpha^{MLE} = 1 + n\left(\sum_{i=1}^{n} \ln \frac{x_i}{x_{min}}\right)^{-1}, \tag{18}$$

of which the standard error is

$$\sigma = \frac{\alpha^{MLE}-1}{\sqrt{n}}. \tag{19}$$

A potential point of contention with MLE estimation is in specifying the threshold parameter, x_{min}. Researchers have traditionally "eyeballed" the data to determine where the power-law behavior begins. Gabaix (2009) recommends limiting the analysis to the 95% quantile, which I follow for one of the data samples I analyze. For the other data sample, I follow the procedure recommended by Clauset et al. (2009) and choose x_{min} endogenously to minimize the Kolmogorov-Smirnov (KS) goodness of fit statistic. The basic idea is to choose x_{min} to minimize the distance between the empirical cumulative distribution, $E(x)$, and the estimated power law cumulative distribution function, $P(x)$. The KS statistic is given as

$$KS = \max_{x \geq x_{min}} |E(x) - P(x)|. \tag{20}$$

The thresholds that minimize this statistic are given in table 28.

Also commonly employed, and asymptotically equivalent to $\alpha^{MLE} - 1$ is Hill's estimator, which is given by

$$\gamma^{Hill} = \frac{(n-2)}{\sum_{i=1}^{n-1}(\ln x_i - \ln x_{min})} \tag{21}$$

The standard error for the Hill estimator is given by $\gamma^{Hill}(n-3)^{-1/2}$. Finally, the power law exponent can be estimated via the following OLS specification.

$$\ln(i) = \alpha - \gamma^{OLS} \ln x_i + \epsilon_i \tag{22}$$

where (i) represents the observation's rank in the distribution, α and γ^{OLS} are the parameters to be estimated, and ϵ_i is the error term. The asymptotic standard error is given by $\gamma^{OLS}(n/2)^{-1/2}$.

γ and α

The literature is divided on reporting of the exponent for the probability distribution function, and the reporting of the exponent for the counter-cumulative distribution function, γ .[23] This section is meant to clarify any misunderstanding. The counter-cumulative distribution function, $P(X > x)$ is given as

$$P(X > x) = \int_x^\infty P(X)dX = \left(\frac{x}{x_{min}}\right)^{-\alpha+1} = \left(\frac{x}{x_{min}}\right)^{-\gamma}. \tag{23}$$

[23] The counter-cumulative distribution function is just (1-CDF).

Results

Results for the sample of the 5 % tail of the distribution are presented in table 27 (The threshold cutoff for oil is 175,375, bbls, the threshold cutoff for gas is 2,051,885 MCF). The MLE and Hill estimator give virtually identical results owing to the large sample size in both the oil and gas samples. For both cumulative oil and gas recovery, the OLS estimates exceed the MLE. Cumulative oil recovery is very close to following Zipf's law (a power law relationship where $\gamma = 1$). The maximum likelihood parameter estimate for oil recovery implies that 82.5% of oil is recovered on 1 % of the leases. For natural gas recovery, the exponent is slightly larger. The divergence between Hill and OLS estimators is also slightly greater. The parameter implies that 1% of wells account for 45.5 % of cumulative gas recovery. For both oil and natural gas, estimates of the scaling parameters imply the distributions have infinite variance.

Table 27: Power law estimates, 5% tail

	gas	oil
α^{MLE}	2.206	2.043
	(0.007)	(0.006)
γ^{Hill}	1.206	1.043
	(0.007)	(0.006)
γ^{OLS}	1.368	1.079
	(0.011)	(0.008)
Observations	29,588	

Notes: Estimates for PL exponent based on 5 % tail. Standard error are in parentheses.

Results for estimates where the sample is chosen endogenously to minimize the Kolmogorov-Smirnov statistic are presented in table 28. The procedure chooses a threshold cutoff further in the tail for both oil and natural gas, and so the sample size diminishes. Compared to the 5% threshold, the estimated exponent for natural gas increases substantially; the same is true for oil recovery, but less pronounced.

Again, parameter estimates indicate that neither distribution has a finite variance. Exponents imply that 1% of gas wells recover 16.9% of gas, while 1% of oil leases are responsible for 70 % of oil recovered.

Table 28: Power law estimates, endogenous threshold

	gas	oil
Threshold	11,599,772 MCF	372,460 bbls
KS	0.010	0.005
α^{MLE}	2.630	2.086
	(0.028)	(0.009)
γ^{Hill}	1.629	1.085
	(0.028)	(0.009)
γ^{OLS}	1.612	1.097
	(0.039)	(0.013)
Observations	3379	13780

Notes: Estimates for power law exponent based on sample with endogenous cutoff. Standard error are in parentheses.

Robustness Tests

Simple parameter estimation is not enough to assert oil and natural gas are power law distributed. Although the data appear linear on a log-log graph, over short enough spans, other distributions such as log-normal can also appear linear. There are two tests that may be implemented, Clauset *et al.* (2008) recommend Kolmogorov -Smirnov testing based on simulated samples; at present I implement a simpler test based on the linear regression proposed by Gabaix and Ibragimov Gabaix and Ibragimov (Gabaix and Ibragimov). Define $x*$ as

$$x* = \frac{cov((\ln x_j)^2, \ln x_j)}{2var(\ln x_j)}, \tag{24}$$

then regress the following equation,

$$\ln(i - \frac{1}{2}) = \alpha + \zeta \ln_i + q(\ln x_i - x*)^2 + \epsilon_i. \tag{25}$$

The test parameter is $\hat{q}/\hat{\zeta}^2$. The null hypothesis, that cumulative oil and natural gas recovery is power law distributed is rejected if $\frac{\hat{q}}{\hat{\zeta}^2} > 1.95(2n)^{-1/2}$. Results of the tests are printed below. The null hypothesis that oil and natural gas are power law distributed–at least in the upper tail–cannot be rejected in any sample.

Table 29: Gaibaix-Ibragimov test of power law

	gas	oil
95% sample		
Test Statistic	1.79e-10	4.11e-10
Threshold	0.008	
Endogenous threshold sample		
Test Statistic	1.56e-08	2.18e-08
Threshold	0.023	0.011

Notes: Gabaix-Ibragimov test statistics. Reject PL distribution if test statistic¿ threshold.

The most convincing evidence of power laws is visual. Graphs of the data can be found below, along with the best fitting distributions from the class of power law, exponential, and log-normal distributions. When the data are graphed in levels, it is difficult to see very much; however, the linear relationship after log transformation of the data and axes is stunning. The best fit exponential and log-normal distributions can not adequately explain fit the data in the tails: these distributions assume that such observations are just too unlikely. Only the power law can fit the explosive randomness in the tails.

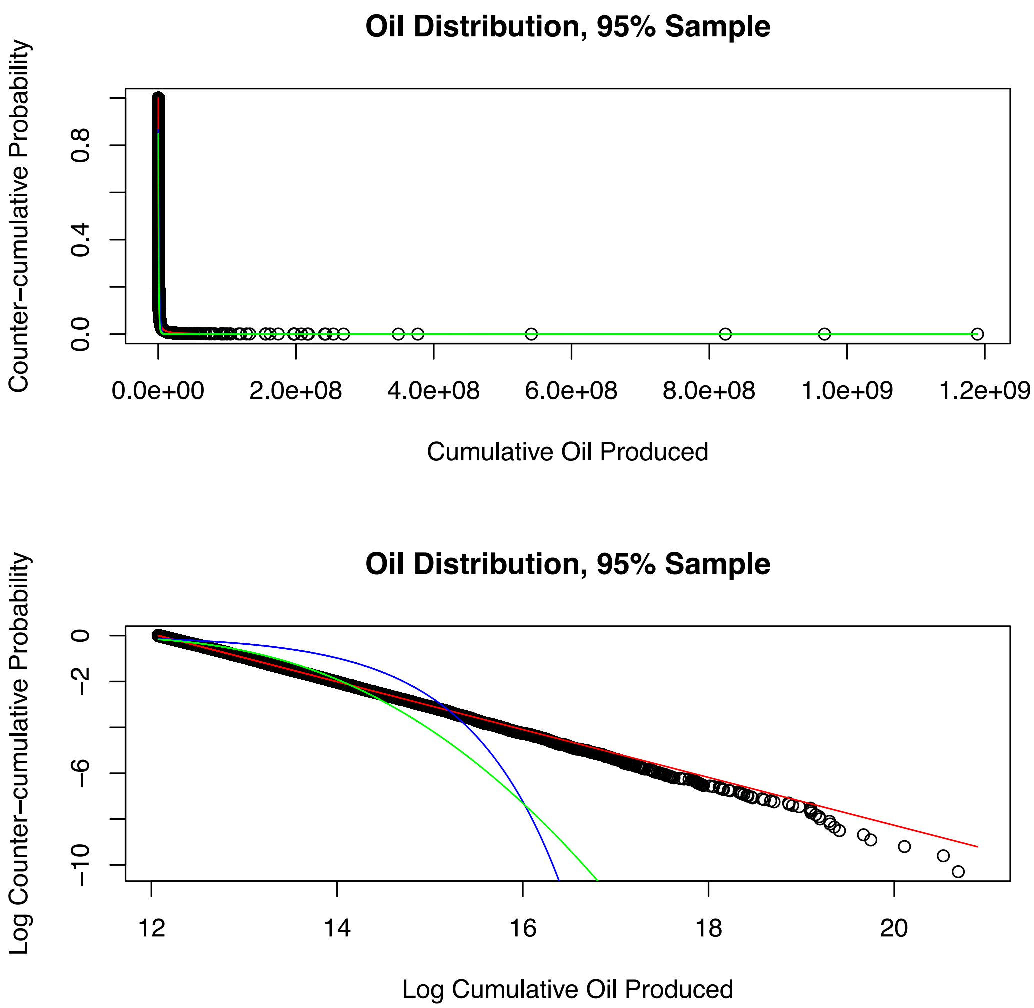
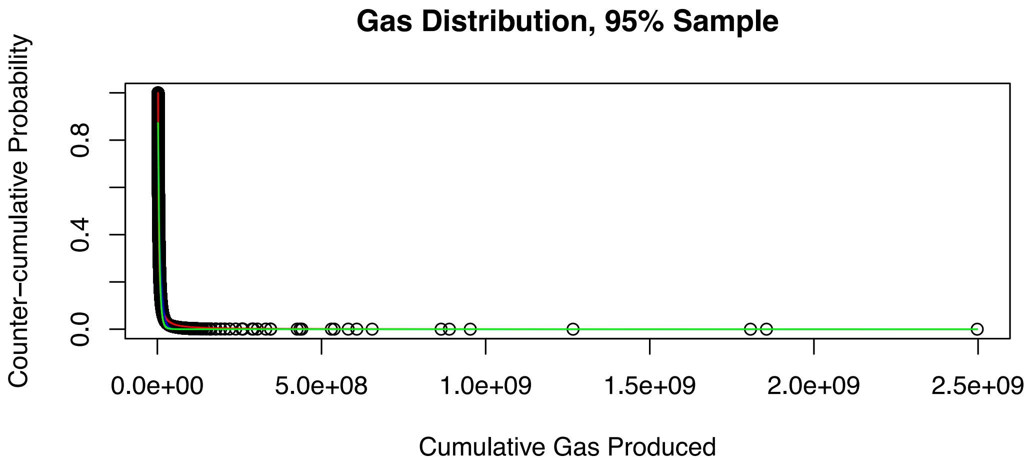


Figure 15: Log cumulative production versus log empirical probability. The data and their empirical probabilities are plotted with circles. The likelihood maximizing power law distribution is in red, the exponential in blue, and the log-normal in green. The data are the 5% tail.



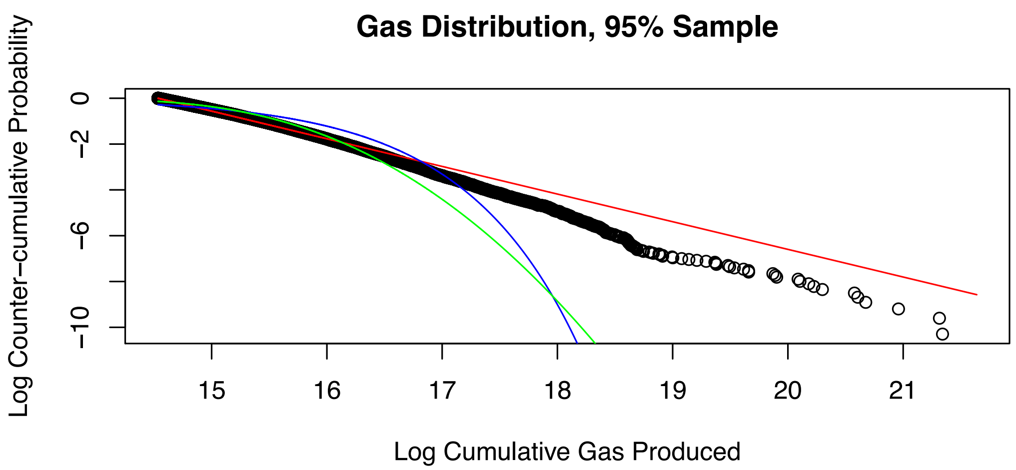


Figure 16: Log cumulative production versus log empirical probability. The data and their empirical probabilities are plotted with circles. The likelihood maximizing power law distribution is in red, the exponential in blue, and the log-normal in green. The data are the 5% tail.

The qualitative evidence provided by the graphs can be made quantitative by implementing a likelihood-ratio type test, as recommended by **?**. The test compares the predicted likelihoods of two competing distributions, favoring the distribution that is more likely. In particular, the test is computed as

$$\mathcal{R} = \sum_{i=1}^{n} \left[\ln p_1(x_i) - \ln p_2(x_i) \right], \tag{26}$$

where $p_1(x)$ and $p_2(x)$ are the probabilities predicted by two distributions.[24] The authors go on to show that $\mathcal{R}$ is normally distributed and give formulas for calculating p-values. I compare the likelihoods computed under the assumption of power law to those under the assumption of exponential and log-normal. These results are presented in table 30. It is readily apparent that the power law distribution has much more explanatory power than the competing distributions across both samples and for both oil and natural gas.

Table 30: Likelihood ratio tests of competing distributions

	PL-Exponential	PL-Log-normal
Gas		
95% sample	18560.26	15585.86
	(0.000)	(0.000)
Endogenous Threshold	2819.377	1441.832
	(0.000)	(0.000)
Oil		
95% sample	30499.41	32302.33
	(0.000)	(0.000)
Endogenous Threshold	13827.24	14751.59
	(0.000)	(0.000)

Notes: Likelihood ratios computed as power law log likelihood-competing distribution log likelihood. Positive numbers indicate the power law distribution is the better fit. P-values for significant differences in likelihoods are in parentheses.

[24] These predicted values are obtained after estimating parameters for the competing distributions via maximum likelihood

Conclusion

This chapter presents strong evidence for power law tails in the distribution of cumulative oil and natural gas production. Leases productivities span many orders of magnitude, log-log graphs of cumulative production demonstrate a striking linear relationship, and quantitative robustness tests indicate the power law distribution to be a good approximation. Of course, given the infinite number of distributions to select from, it is possible to find one that better fits the data. Yet the power distribution illustrates key features in the data (its heavy tails) and does so parsimoniously. I have also shown that a power law distribution fits the data in the tails much better than more commonly used distributions such as exponential and log-normal.

The power law result is significant for both management and regulation, particularly in the case of oil production. By overseeing just 1 % of leases, regulators can monitor nearly 83 % of cumulative production. Similarly for production companies, their profitability is determined not by the vast majority of the leases operated, but by their most productive 1 %. How this distribution should affect managerial decisions is an exciting avenue for further research.

Finally, it is natural to ask what is causing the data to be power law distributed. Power law distributions have been found to result from a broad away of processes, including, optimization problems with a particular set of constraints, random walks, Yule processes, combinations of exponential distributions, and phase transitions–and this is still a very active area of research. A likely explanation is that scaling is common in nature, and therefore the scaling distribution is as well. Indeed we see it in the distribution of galaxies, supernovas, severity of flooding and earthquakes, and, apparently, in the productivity of oil an natural gas leases.

Chapter VI

CONCLUSION

In this dissertation, I have examined data in Texas and Oklahoma, looking for evidence of common pool externalities, which distort production incentives away from the social optimum. To identify the common pool externality, I compare areas where ownership is secure–where leases are operated by a single manager–to areas where *in situ* resource ownership is insecure–areas where there are many competing operators. I find that secure property rights enhance cumulative oil recovery. This result cannot be explained with present economic models that assume the stock of recoverable reserves is fixed. For empirical applications, it is necessary to model the stock of reserves as endogenous to the slope of the extraction profile. Using two very different empirical techniques, I also uncover evidence of distorted production profiles that can explain the difference in cumulative recovery. Both regression discontinuity and the spatial model indicate that lease owners with insecure rights to the resource *in situ* extract at a higher rate conditional on the age of the well. Finally, I demonstrate the the results are economically important. The average well is important from a managerial and regulatory perspective, and so enhancing recovery at the average well by solving the common pool externality is economically important.

Appendix A

APPENDIX TO CHAPTER III

Federal Regulation

Federal regulations, being identical for both Texas and Oklahoma, cannot be the sole contributor to differences in production evident at the border. Federal regulations can, however, magnify or diminish the effects of existing state-level policy differences previously mentioned. In this section, I give an overview of important federal regulation of oil and natural gas production, which, excluding production on federal lands, is implemented mostly through the federal tax code through credits and deductions.

- ENHANCED OIL RECOVERY CREDIT. This tax credit has been a target for repeal in the 2012 fiscal year (FY2012). The tax credit is worth 15% of allowable costs related to secondary injection and is only available in years where oil price is "low" (it has not been effective in recent years). The credit was first introduced in 1990 and was worth 10% of allowable costs.

- CREDIT FOR OIL AND GAS FROM MARGINAL WELLS. The credit is designed to keep high-cost wells in production even when prices are low and has been targeted for repeal in FY2012. Marginal wells are defined as wells that produce less than 15 bbls of oil (or oil equivalent) per day. This credit came into effect as part of American Jobs Creation Act of 2005.

- EXPENSING OF INTANGIBLE DRILLING COSTS. Costs of drilling such as land clearing, surveying, wages, drilling mud, chemicals, cement, etc., can be expensed. This is at present only available to independent oil producers. Major integrated oil companies (i.e, vertically integrated companies) can only expense 30 % of drilling costs over a 60 month period (rather than expensing it in the same year). This was first introduced in 1916, and is a target for repeal in FY2012. Without expensing, drilling costs would be capitalized into the well, and expensed over the lifetime of the well (standard capital depreciation allowance). Taking it all in one year does much to make the well profitable. This measure was eliminated for large companies in 1970s. It does not take too much imagination to see how this measure could exacerbate common pool externalities in Texas.

- TERTIARY INJECTANTS DEDUCTION. Tertiary injectants (injectants used in enhanced recovery operations distinct from secondary flooding) can be fully deducted in the current tax year. This deduction has been targeted for repeal in FY2012.

- PASSIVE LOSS EXCEPTION FOR WORKING INTERESTS IN OIL PROPERTIES. Although not a large item in terms of revenue, the exception permits deduction of losses in oil and gas projects against other income earned. It is targeted for repeal in FY2012.

- PERCENTAGE DEPLETION ALLOWANCE. Independent companies are allowed a 15 % deduction from gross income for depletion of the deposit. This allowance was repealed for major oil companies in 1975. It was first introduced in 1926 and has been targeted for repeal in FY2012.

- MANUFACTURING TAX DEDUCTION. The oil and gas industry is classified as being in the manufacturing sector according to 2005 US Jobs

Creation Act. Companies are allowed a 9 % deduction from net income with a cap given according to employment.

- AMORTIZATION OF GEOLOGICAL AND GEOPHYSICAL PERIOD. This measure concerns costs associated with exploration: Independent oil and gas companies take geological/geophysical expenses in the same year, while major oil companies must amortize the expenses over 7 years (which is less beneficial). This measure is a target for repeal in FY2012.
- SECTION 29 PRODUCTION TAX CREDIT FOR NON-CONVENTIONAL OIL. Originally part of the windfall profits tax, this credit was retained after the repeal of the windfall profits tax. The credit allows a $3 (indexed in 1979 dollars, $6.80 today) credit per barrel of oil equivalent production. This credit is especially beneficial for the production of coal-bed methane.
- 1980 WINDFALL PROFITS TAX. A higher tax rate goes into effect when price climbs above a threshold; the tax was repealed in 1988.
- CERCLA. A $ 0.098 per barrel tax is levied on crude oil received at refineries; the tax expired in 1995.

Other Parametric Specifications

Parametric regression discontinuity results in Chapter 3 are estimates from a fully interacted model of the form:

$$y_{it} = \alpha + \sum_{j=1}^{n} \beta_j dist_i^j + \sum_{j=1}^{n} \beta_j dist_i^j * OK + \tau OK + \gamma_t + \epsilon_{it}. \tag{27}$$

Where $dist_i$ represents distance to the border (in terms of decimal degrees) of well i. To show robustness of results, in this appendix I re-estimate a model of the form:

$$y_{it} = \alpha + \sum_{j=1}^{n} \beta_j latitude_i^j + \tau OK + \gamma_t + \epsilon_{it}. \tag{28}$$

The reasoning for showing this result is that since there is no reason to expect that the Oklahoma treatment should affect latitude, except via the intercept, latitude should not be interacted with treatment. This is to say that latitude affects the dependent variable in the same way in Oklahoma and Texas. These models were run with pooled data, as well as a sample of data limited to the reservoirs straddling the border (”within”), and the sample of reservoirs that do not cross the border (”between”).

As is argued in the paper, we expect the sample of within reservoirs to provide a lower bound on the true treatment effect because Texas production may interfere with the benefit of partial Oklahoma unitization. Indeed, the tables show that between estimates are larger, but also slightly less comparable in terms of exogenous variables.

Table 31: Within and between estimation: sample of young wells

	Pooled		Within			Between			
DEP. VAR.	(1)	(2)	(1)	(2)	(3)	(1)	(2)	(3)	(4)
log gas	-0.904***	-0.927***	-1.174***	-1.215***	-0.235	-0.149			
	(0.345)	(0.347)	(0.410)	(0.411)	(0.586)	(0.539)			
log oil	-0.522	-0.499	-0.729	-0.688		0.277	0.266	1.861**	1.738**
	(0.439)	(0.431)	(0.507)	(0.493)		(0.655)	(0.632)	(0.853)	(0.816)
log rev	-0.824***		-1.019***			-0.308			
	(0.274)		(0.317)			(0.443)			
log cum rev	-0.798***		-0.989***			-0.283			
	(0.275)		(0.322)			(0.450)			
log cum oil	-0.147		-0.38	-0.338		0.971	0.956	2.600*	2.428*
	(0.534)		(0.595)	(0.585)		(1.032)	(1.002)	(1.391)	(1.351)
log cum gas	-0.755*	-0.785*	-1.020**	-1.071**	0.089	-0.011			
	(0.400)	(0.403)	(0.479)	(0.481)	(0.682)	(0.594)			
well depth	234.904	227.17	8.279			1,301.881**	1,292.984**		
	(209.391)	(208.555)	(187.587)			(659.137)	(646.115)		
completion	-4.804		37.048			-164.032			
	(54.679)		(65.397)			(101.060)			
longitude	0.042		0.013	0.007	-0.128	0.15			
	(0.064)		(0.070)	(0.071)	(0.094)	(0.123)			

Notes: The columns represent average treatment effects estimates from OLS regression for different orders of polynomial distance interactions (the model is specified in equation 28). Rows represent different independent variables. Robust standard errors clustered at the lease level are reported in parentheses. Only specifications where the polynomials are jointly significant are reported.

Table 32: Within and between estimation: sample of old wells

	Pooled		Within		Between
DEP. VAR.	(1)	(2)	(1)	(2)	(1)
log gas	-2.299***	-2.207***	-2.289***		-1.961
	(0.456)	(0.455)	(0.561)		(1.407)
log oil	1.452***	1.385***	1.277***	1.189***	3.104***
	(0.319)	(0.320)	(0.394)	(0.389)	(0.893)
log rev	0.711**		0.48		2.299***
	(0.319)		(0.379)		(0.815)
log cum rev	0.323		0.364		-0.208
	(0.310)		(0.387)		(0.965)
log cum oil	2.967***	2.786***	2.860***	2.692***	4.808***
	(0.638)	(0.645)	(0.732)	(0.735)	(1.802)
log cum gas	-0.873		-1.038		-0.701
	(0.618)		(0.759)		(1.580)
well depth	281.35		169.881		1,117.48
	248.287		(243.706)		(1031.974)
completion	-983.098**		-1,118.789**		-81.629
	(386.351)		(472.208)		(1195.052)
longitude	0.333***		0.274***		0.650**
	(0.074)		(0.078)		(0.279)

Notes: The columns represent average treatment effects estimates from OLS regression for different orders of polynomial distance interactions (the model is specified in equation 28). Rows represent different independent variables. Robust standard errors clustered at the lease level are reported in parentheses. Only specifications where the polynomials are jointly significant are reported.

Appendix B

APPENDIX TO CHAPTER IV

The main result of Chapter IV is of the model with separate weighting for the effect of neighboring wells, based on whether those wells have common owners or competing owners, while controlling for other covariates. The results from this specification are flawed because they fail to include dummy variables for time. The tables below include time dummies, which significantly change the results. The estimated spillover parameter for " friendly" and "unfriendly" wells are negative and not statistically different from each other or zero.

Table 33: Simultaneous weight matrices with injection and time dummies: oil

	OLS	GLS	2SLS	GS2SLS
c.1.1.	-0.0097	0.1188	-0.0102	-0.0115
	0.0148	0.2879	0.0153	0.0058
wcnt	0.0135	0.1220	0.0134	0.0219
	0.0034	0.0425	0.0057	0.0129
wcnt2	-0.0076	-0.0422	-0.0076	-0.0151
	0.0016	0.0155	0.0026	0.0059
age	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
wtr	0.5416	0.6595	0.5418	0.3074
	0.0337	0.2499	0.0364	0.0497
ginj	-0.0069	0.0879	-0.0070	-0.0270
	0.0026	0.0305	0.0064	0.0110
winj	0.0103	-0.2609	0.0103	0.0453
	0.0021	0.1247	0.0211	0.0452
λ_F	0.0088		0.0058	-0.4663
	0.0016		0.3131	0.6755
λ_U	0.0041		0.0519	-0.0703
	0.0004		0.0452	0.0715
Moran	177.2908			
ρ		0.1229		0.0938
		0.0053		0.0035

Notes: Results in this table replicate the specification of table 22 in Chapter IV, except with time dummies added to the specification. Oil production is the dependent variable.

Table 34: Simultaneous weight matrices with injection and time dummies: gas

	OLS	GLS	2SLS	GS2SLS
c.1.1.	-0.0125	-0.9334	-0.0188	-0.0102
	0.0566	0.3252	0.0566	0.0174
wcnt	0.0502	0.2354	0.0501	0.0560
	0.0131	0.0522	0.0211	0.0246
wcnt2	-0.0217	-0.0521	-0.0228	-0.0172
	0.0061	0.0258	0.0097	0.0109
age	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0001	0.0000	0.0000
wtr	0.8746	1.2446	0.8489	0.5178
	0.1285	0.4780	0.1360	0.1080
ginj	-0.0314	0.0970	-0.0305	0.0141
	0.0098	0.1183	0.0238	0.0238
winj	0.0701	-0.3051	0.0681	-0.0804
	0.0082	0.2402	0.0781	0.0845
λ_F	0.0068		-0.8763	1.1273
	0.0040		1.1605	1.2704
λ_U	0.0036		0.0254	0.1200
	0.0012		0.1674	0.1627
Moran	-692.7790			
ρ		0.0684		0.0469
		0.0034		0.0019

Notes: Results in this table replicate the specification of table 23 in Chapter IV, except with time dummies added to the specification. Gas is the dependent variable.

This specification is still the favored model because it controls for local injection, and the different incentives for production based on the ownership of nearby wells. The result can be investigated further in future research with different specifications of the friendly and unfriendly weighting matrix. At present both use simple inverse distance, which may put too much weight on wells that are far apart. There is little reason to expect wells to communicate at great distances; even wells that are close would not communicate if there is not a direct line of sight between them. Parameter estimates are of an average spillover based on the assumption that the spatial landscape is homogenous and isotropic. Within reservoir, I argue this assumption is tenable, but it undoubtedly becomes more so as we examine only the wells that are relatively close to one another within reservoir (which the square of inverse distance would achieve). Additional information on the reservoir, such as rock permeabilities, elevations changes, faults and nonconformities would allow even more accurate weighting.

Also for future research, owner dummies can be added to the models. At present the idiosyncrasies of producers fall in the error term, and it is possible that this can bias the estimated spillover parameters.

References

Alley, W. and J. Schefter (1987). External effects of irrigator's pumping decisions, high plains aquifer. *Water Resources Research 23*(7), 1123–1130.

Anselin, L. (2002, July). Under the hood. Issues in the specification and interpretation of spatial regression models. *Agricultural Economics 27*, 247–267.

Atkinson, A. and T. Piketty (2007). *Top Incomes over the Twentieth Century*. Oxford: Oxford UP.

Axtell, R. (2001). Zipf distribution of us firm sizes. *Science 293*, 1818–1820.

Berck, P. (1995). *Handbook of Environmental Economics*, Volume 1, Chapter Emprical Consequences of the Hotelling Principle. Blackwell.

Brown, G. J. (1974, January-February). An optimal program for managing common property resources with congestion externalities. *The Journal of Political Economy 82*(1), 163–173.

Brozovic, N., D. Sunding, and D. Zilberman (2006, July). On the spatial nature of the groundwater pumping externality.

Chermak, J. M. and R. H. Patrick (1995). A well-based cost function and the economics of exhaustible resources: The case of natural gas. *Journal of Environmental Economics and Management 28*, 174–189.

Chermak, J. M. and R. H. Patrick (2001). A microeconometric test of the theory of exhaustible resources. *Journal of Environmental Economics and Management 42*, 82–103.

Clark, C. (1973, August, 17). The economics of overexploitation. *Science 181*(4100), 630–634.

Clauset, A., C. R. Shalizi, and M. Newman (2009). Power law distributions in empirical data. *SIAM Review 51*, 661–703.

Costello, C. and R. T. Deacon (2007). Efficiencey gains from fully dilineating rights in an itq fishery. *Marine Resource Economics 22*, 374–361.

Dake, L. (2001). *The Practice of Reservoir Engineering* (Revised Edition ed.), Volume 36 of *Developments in Petroleum Science*. New York: Elsevier.

Dasgupta, P. and G. Heal (1979). *Economic Theory and Exhaustible Resources.* Cambridge Economic Handbooks. Cambridge University Press.

Davis, G. and R. Cairns (1999). Valuing petroleum reserves using current net price. *Economic Inquiry 37*(2), 295–311.

Demsetz, H. (1967). Toward a theory of property rights. *The American Economic Review 57*(2), 347–359.

Eswaran, M. and T. Lewis (1984, November). Appropriability and the extraction of a common property resource. *Economica 51*(204), 393–400.

Fan, J. (1992, December). Design-adaptive nonparametric regression. *Journal of the American Statistical Association 87*(420), 998–1004.

Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics 1*, 225–293.

Gabaix, X. and R. Ibragimov. A simple ols test of power law behaviour. Working Paper, Harvard University.

Gabaix, X. and Y. Ioannides (2004). The evolution of the city size distributions. In T. J. Henderson V (Ed.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2341–2378. Oxford: Elsevier.

Gamponia, V. and R. Mendelsohn (1985, February). The taxation of exhaustible resources. *The Quarterly Journal of Economics 100*(1), 165–181.

Gordon, H. S. (1954, April). The economic theory of a common-property resource: The fishery. *The Journal of Political Economy 62*(2), 124–142.

Halvorsen, R. and T. Smith (1991). A test of the theory of exhaustible resources. *The Quarterly Journal of Economics 106*(1), 123–140.

Hardin, G. (1968). The tragedy of the commons. *Science 162*(3859), 1243–1248.

Heal, G. and M. Barrow (1980). The relationship between interests rates and metal price movements. *Review of Economic Studies 47*(1), 161–181.

Hendricks, K. and R. Porter (1993). Determinants of the timing and incidence of exploratory drilling on offshore wildcat tracts. NBER Working Paper No. 4605.

Hendricks, K. and R. Porter (1996). The timing and incidence of exploratory drilling on offshore wildcat tracts. *The American Economic Review 83*(3), 388–407.

Hotelling, H. (1931). The economics of exhaustible resources. *The Journal of Political Economy 39*(2), 137–175.

Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*, 615–635.

Kaffine, D. T. and C. Costello (2011). Unitization of spatially connected renewable resources. *The B.E. Journal of Economic Analysis and Policy 11*(1).

Kelejian, H. and I. Prucha (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics 17*(1), 99–121.

Kelejian, H. and I. Prucha (1999, May). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review 40*(2), 509–533.

Kellogg, R. (2010, November). The effect of uncertainty on investment: evidence from Texas oil drilling. NBER Working paper series, No. 16541.

Khalatbari, F. (1977, November). Market imperfections and the optimal rate of depletion of natural resources. *Economica 44*(176), 409–414.

Krautkraemer, J. (1998). Nonrenewable resource scarcity. *The Journal of Economic Literature 36*(4), 2065–2107.

Lee, D. S. and T. Lemieux (2010, June). Regression discontinuity designs in economics. *Journal of Economic Literature 42*(2), 281–355.

Lee, J., J. List, and M. Strazich (2006). Non-renewable resource prices: Deterministic or stochastic trends. *Journal of Environmental Economics and Management 51*, 354–370.

Levhari, D. and L. Mirman (1980). The great fish war: An example using dynamic cournot-nash solution. *The Bell Journal of Economics 11*(1), 322–334.

Lewis, T. R. and R. Schmalensee (1980, November). On oligopolistic markets for nonrenewable natural resources. *The Quarterly Journal of Economics 95*(3), 475–491.

Libecap, G. (1989a). *Contracting for Property Rights*. Political Economy of Institutions and Decisions. New York: Cambridge University Press.

Libecap, G. (1989b, December). The political economy of crude oil cartelization in the United States, 1933-1972. *Journal of Economic History 49*(4), 835–855.

Libecap, G. and J. L. Smith (1993). The self-enforcing provisions of oil and gas unit operating agreements: Theory and evidence.

Libecap, G. D. and J. L. Smith (1999). The self-enforcing provisions of oil and gas unit operating agreements: Theory and evidence. *Journal of Law, Economics, and Organization 12*(2), 526–548.

Libecap, G. D. and S. N. Wiggins (1984, March). Contractual response to the common pool: Prorationing of crude oil production. *The American Economic Review 74*(1), 87–98.

Libecap, G. D. and S. N. Wiggins (1985, August). The influence of private contractual failure on regulation: The case of oil field unitization. *The Journal of Political Economy 93*(4), 690–714.

Lin, C.-Y. C. (2009). Estimating strategic interactions in petroleum exploration. *Energy Economics 31*, 586–594.

Livernois, J. (2008). On the empirical significance of the hotelling rule. *Review of Environmental Economics and Policy*, 1–20.

Long, N. V. (1974, February). International borrowing for resource extraction. *International Economic Review 15*(1), 168–183.

Long, N. V. (1975). Resource extraction under the uncertainty about possible nationalization. *Journal of Economic Theory 10*, 42–53.

Lowe, J. S. (2003). *Oil and Gas Law in a Nutshell* (4th ed.). Nutshell Series. Thomson West.

Ludwig, J. and D. L. Miller (2007, February). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics 122*(1), 159–208.

Lueck, D. (1995, October). The rule of first possession and the design of the law. *Journal of Law and Economics 38*(2), 393–436.

Lueck, D. and P. Schenewerk (1996). An economic analysis of unitized and non-unitized production. *Proceedings of the 1996 Society of Petroleum Engineers Annual Technical Conference*, 67–76.

Miller, M. and C. Upton (1985). A test of the hotelling valuation principle. *The Journal of Political Economy 93*(1), 1–25.

Neumann, S. (1972). Theory of flow in unconfined aquifers considering delayed response in water table. *Water Resources Research 8*(4), 1031–1045.

Newman, M. (2005, September-October). Power laws, pareto distributions and zipf's law. *Contemporary Physics 46*(5), 323–351.

Nind, T. (1981). *Principles of Oil Well Production* (2nd ed.). St. Louis: McGraw-Hill.

Pareto, V. (1896). *Cours D'Economie Politique*. Geneva: Droz.

Pfeiffer, L. and C.-Y. C. Lin (2009, May). Groundwater pumping and spatial externalities in agriculture.

Pindyck, R. (1978). The optimal exploration and production of non-renewable resources. *The Journal of Political Economy* (86), 841–861.

Pindyck, R. (1980). Uncertainty and exhaustible resource markets. *The Journal of Political Economy 88*(6), 1203–1225.

Pindyck, R. (1982, December). Jointly produced exhaustible resources. *Journal of Environmental Economics and Management 9*(4), 291–303.

Reinganum, J. and N. Stokey (1985, February). Oligopoly extraction of a common property natural resource: The importance of the period of commitment in dynamic games. *International Economic Review 26*(1), 161–173.

Savage, J. and N. Brozovic (2011, June). Spatial externalities and strategic behaviour in groundwater pumping. AERE Summer Conference.

Scott, A. (1955, April). The fishery: The objectives of sole ownership. *The Journal of Political Economy 63*(2), 116–124.

Stiglitz, J. (1976, September). Monopoly and the rate of extraction. *The American Economic Review 66*(4), 655–661.

USGS (2011, January). National assessment of oil and gas fact sheet: Assessment of undiscovered oil and gas resources of the Anadarko basin province of Oklahoma, Kansas, Texas, and Colorado, 2010. Technical report, US Geological Survey.

Vousden, N. (1973). Basic theoretical issues of resource depletion. *Journal of Economic Theory 6*, 126–143.

Weitzman, M. (1974). Free access vs private property as alternative systems for managing common property. *Journal of Economic Theory 8*, 225–234.

Wiggins, S. N. and G. D. Libecap (1985, June). Oil field unitization: Contractual failure in the presence of imperfect information. *The American Economic Review 75*(3), 368–385.

Yatchew, A. (1998, February). Nonparametric techniques in economics. *Journal of Economic Literature 36*(2), 669–721.

Yergin, D. (2008). *The Prize: The Epic Quest for Oil, Money and Power.* New York: Free Press.

Yuan, L. (2002, February). Divide and conquer: Multiple leasing in common pool oil fields. *The Canadian Journal of Economics 35*(1), 36–51.

Zipf, G. (1949). *Human Behavior and the Principle of Least Effort.* Cambridge, MA: Addison-Wesley.

VITA

Andrew Travis Balthrop was born in Jackson, Mississippi, USA on December 8, 1982. He holds a Bachelor of Arts in International Studies and Economics from the University of Mississippi.

Andrew began Georgia State's doctoral program in 2006, with a focus in Environmental Economics. He has worked as a graduate research assistant to Dr. Laura O. Taylor, Dr. H. Spencer Banzhaf, and Dr. Kurt E. Schnier and as a graduate teaching assistant to Dr. Shelby D. Frost (Principles of Microeconomics, Spring 2011). He was the sole instructor of Environmental Economics and Policy (May, 2011). He served as research intern to the Federal Reserve Bank of Atlanta from 2008-2011, and as a graduate fellow at the Property and Environment Research Center in Bozeman, Montana.

Andrew received is Ph.D. in Economics from Georgia State University in May, 2012. He accepted an assistant professorship at the American University of Sharjah.