NBER WORKING PAPER SERIES

REAL OPTION EXERCISE:
EMPIRICAL EVIDENCE

Paul Décaire
Erik P. Gilje
Jérôme P. Taillard

Working Paper 25624
http://www.nber.org/papers/w25624

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2019

We thank seminar and conference participants at Babson College, ITAM, USF, Wharton, 2018 Texas Finance Festival, 2018 IDC Finance Conference, 2018 Northeastern University Finance Conference, 2018 European Finance Association Conference, 2019 American Finance Association, as well as Jules van Binsbergen, Jonathan Berk, Gustavo Grullon, John Hund, Evgeny Lyandres, Ohad Kadan, Michael Roberts, Sheridan Titman, and Pavel Zryumov for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Real Option Exercise: Empirical Evidence
Paul Décaire, Erik P. Gilje, and Jérôme P. Taillard
NBER Working Paper No. 25624
February 2019
JEL No. G30,G31,G32

ABSTRACT

We study when and why firms exercise real options. Using detailed project-level investment data, we find that the likelihood that a firm exercises a real option is strongly related to peer exercise behavior. Peer exercise decisions are as important in explaining exercise behavior as variables commonly associated with standard real option theories, such as volatility. We identify peer effects using localized exogenous variation in peer project exercise decisions and find evidence consistent with information externalities being important for exercise behavior.

Paul Décaire
University of Pennsylvania
3620 Locust Walk, Suite 2400
Philadelphia, PA, 19104
pdecaire@wharton.upenn.edu

Erik P. Gilje
The Wharton School
University of Pennsylvania
3620 Locust Walk
Philadelphia, PA 19104
and NBER
gilje@wharton.upenn.edu

Jérôme P. Taillard
Babson College
323 Tomasso Hall
Babson Park, MA 02457
United States
jtaillard1@babson.edu

An online appendix is available at http://www.nber.org/data-appendix/w25624

Introduction

Every investment decision made by a firm is both a decision about which capital project to pursue as well as when to pursue it. The flexibility associated with the timing of investment decisions has value to the firm; this value is commonly referred to as real option value (Myers (1977)). Real options are a central component of models of the macro economy (Bernanke (1983)), and their exercise has received ample attention in the corporate finance theory literature (e.g., Dixit and Pindyck (1994), Kellogg (2014)). Moreover, existing corporate finance theories hypothesize the importance of peer exercise decisions and information revelation in determining exercise behavior.[1] However, despite the importance of real options, micro-level empirical evidence on exercise behavior remains limited.[2] In this study, we provide novel evidence on the real option exercise behavior of firms; and directly assess the role that peer effects and information externalities can have on exercise decisions.

Characterizing firms' real option exercise behavior is empirically challenging. First, detailed data on the timing flexibility associated with capital projects is typically unavailable. Second, in order to understand a firm's exercise behavior, one would need data on both the projects that a firm decides to undertake, as well as those it decides *not* to pursue. This level of disclosure is often not available. Third, being able to observe key inputs that might drive option exercise decisions is necessary in order to characterize exercise behavior; these would include expected project cash flows, costs, and volatility of project cash flows. Fourth, in a competitive setting where peer firms' exercise behavior can have an influence, one needs to be able to precisely measure the actions taken on peer firm projects in order to gauge their potential impact. Lastly, one needs to develop an empirical framework to appropriately identify the effect of peer behavior and mitigate potential endogeneity concerns.

This study focuses on a setting which allows us to make significant progress on each

[1] See Grenadier (1996), Grenadier (1999), Grenadier (2002), Novy-Marx (2007), Grenadier and Wang (2005), Grenadier and Malenko (2011), and Scharfstein and Stein (1990).

[2] Kellogg (2014) studies oil drilling activity and finds that oil price volatility affects investment decisions in a manner consistent with real option models. However, the study does not assess the importance of information externalities across firms as it focuses on fields operated by a single firm. Moel and Tufano (2002) study mine opening and closing decisions relative to what real option theories would imply, however, their setting is also not conducive to assessing the importance of peer effects and information externalities.

of these challenges. We analyze \$107.9 billion in capital projects composed of exercised and unexercised natural gas infill well drilling projects in major shale developments in North America. First, the institutional structure of this setting allows us to have clear visibility into the timing flexibility firms have in making drilling decisions. Second, due to the institutional structure of lease contract terms we are also able to observe both exercised and unexercised options at any given point in time. Third, because the key determinant of project cash flow is the price of natural gas, a commodity whose expected price and implied volatility are readily observable to the econometrician from financial derivatives, we have the inputs necessary to characterize investment behavior. Fourth, due to the regulatory environment of the shale fields in our setting, we are able to observe and precisely measure neighboring activity from peers.[3] Lastly, we develop an empirical framework which uses novel quasi-exogenous variation in peer activity to mitigate some of the challenges in identifying peer effects.

Our empirical design to assess the exercise behavior of firms is based on a duration analysis using a hazard model. The objective of using this empirical framework is to compute how different factors affect the probability of exercising an option at time t, conditional on the option having not been exercised up to time t. The data in our sample is conducive to this type of analysis because each option has a well-defined starting point, we can clearly observe when an option is exercised, and we have detailed data on how covariates vary during and up to the time of exercise. This empirical specification is consistent with others that have modeled drilling decisions (Kellogg (2014)).

We find that the likelihood that a firm exercises its real option is strongly related to peer exercise behavior. Specifically, a one standard deviation increase in adjacent peer project exercise activity is linked with between a 10.9% and 38.2% increase in exercise likelihood. These magnitudes imply that peer behavior can be as economically important as baseline real option inputs, such as commodity prices and volatility, in determining exercise decisions.[4] We show that our baseline peer effect result holds after mitigating endogeneity concerns linked

[3]This is a key distinction from Kellogg (2014) who focuses on single operated fields, where there is only one firm operating in each area.

[4]As in Kellogg (2014), we find that commodity price and implied volatility are linked with exercise decisions in our setting.

with peer exercise decisions as well as across a series of robustness tests.

Corporate finance theory provides a rich set of extensions to baseline real option models highlighting the importance of peer behavior and information revelation for exercise decisions (e.g., Grenadier (1999), Grenadier and Wang (2007), Grenadier and Malenko (2011), Novy-Marx (2007)). Our empirical framework is well suited to assess these theories. In most other settings, even the task of defining the set of peers can be a challenge.[5] In our setting, geographical proximity of *real options* to one another provides a natural way to define peer sets. Specifically, we are able to precisely observe how firms respond to adjacent competitor project exercise decisions because our data is granular enough that we can observe the specific drilling units (real options) a firm has, as well as the adjacent drilling units operated by competitors. The grid pattern of drilling units in the shale fields in our setting are such that every six square mile township is divided in 36 sections and for each section in our sample, we have eight adjacent sections to it (see Figure 3). We can take advantage of the significant variation in neighboring activity to evaluate two possible channels through which peer exercise could affect exercise decisions.

First, as Grenadier (1996) highlights, firms may face a common pool problem, in which case they may decide to exercise early because the common pool of resources could be drained by neighboring competitors and hence unravel any option value to wait. However, this phenomenon is unlikely to explain exercise behavior because shale rock lies deep underground and traps hydrocarbons tightly. It is only under very intense pressure (hydraulic fracturing or "fracking") that the highly non-permeable rock releases hydrocarbons, with minimal impact on neighboring non-fracked shale rock. If shale gas were a significant common pool, one would likely see only a few wells being drilled to extract natural gas, which is in sharp contrast to the dense drilling that one actually observes in shale gas extraction (see Figure 3).

Second, we evaluate the role that competitor exercise behavior has in providing potentially important information externalities. As Grenadier (1999) points out, information revelation

[5] In a broad cross-section of firms, defining peer sets, often through industry classification, can be challenging (see Hoberg and Phillips (2016)). Defining geographic proximity at the firm level represents another challenge; for instance, headquarter location (easily observable) might act as a poor proxy for where firm operations actually take place.

through real option exercise decisions is a key dimension through which real option exercise behavior differs from financial option exercise behavior.[6] However, micro-level empirical evidence attempting to quantify the potential importance of information revelation remains limited. We find direct evidence that information externalities linked with peer behavior are important. Specifically, we find that firm exercise activity is most strongly linked to peer exercise decisions when peers have more experience in drilling shale infill projects. Firms with the most experience in a field are higher up the learning curve in terms of how to extract natural gas, so the information revealed from their exercise is likely more valuable.

What is the nature of the information firms obtain from adjacent exercise activity? Adjacent exercise activity could inform a firm on how to better extract reserves from its own project. Specifically, adjacent exercised projects reveal detailed information on the "target" depths at which the formation was drilled, which helps firms target their own drilling prospects better. Further, public disclosures require information to be disclosed on the mix of fracking chemicals and techniques applied to drill and complete a well; this information can then be used by peer firms to determine which approach will allow them to extract natural gas most efficiently from their own reservoir (e.g., Covert (2015)).[7] Lastly, adjacent exercise activity by peer firms could also be a reflection of some private information about rock quality a firm has which is not yet publicly known, so that observing a peer firm exercise could cause a firm to update positively on the rock quality of a project. All of these reasons highlight how neighboring peer exercise activity can lead to economically important information externalities that can result in upward project NPV revisions.

A central concern when evaluating the effect of peer exercise decisions is endogeneity. For example, common characteristics (e.g., shared geology or technology) may be driving the exercise behavior of both the firm and its neighboring competitors. This common unobserved factor is a well-established source of endogeneity that leads to the reflection problem (Manski (1993)). To mitigate this endogeneity concern we develop novel quasi-exogenous variation in peer firm exercise activity.

[6] A notable potential exception is the exercise of financial options by insiders or executives.
[7] See fracfocus.org for examples of the types of disclosures that are made public.

Our primary identification strategy relies on the idea that beyond the Net Present Value (NPV) of a project, the relative rank of a given project in a firm's portfolio of capital projects may also matter for investment exercise decisions.[8] Therefore, two peer firms with adjacent projects of similar NPVs could undertake exercise decisions differently due to the relative rank of their project within each firm's portfolio of projects. For each real option in our sample, we construct the average relative rank percentile of adjacent projects within the peer firms' portfolio of projects at each point in time. We use this variable to instrument for adjacent peer project exercise activity. We find evidence, using both instrumented and reduced form versions of this measure of quasi-exogenous variation in peer exercise activity, that the adjacent exercise behavior of peer firms affects the exercise behavior of a firm.

The identification assumption of our empirical design is that the relative rank of the NPV of an adjacent real option in a peer firm's portfolio affects a firm's own exercise decision only through its effect on the likelihood that the peer firm will exercise that adjacent option, and not through another channel. While this assumption is not directly testable, we can provide several pieces of evidence that support it. First, if a common characteristic affected both the relative rank of a peer firm's real option as well as the exercise of a firm's own real option, then the exclusion restriction would be violated. In such circumstances, one might expect highly ranked projects by different firms would tend to cluster in the same area and we show that this is not the case. Specifically, we show that after controlling for local geography fixed effects, which essentially controls for the absolute (but not relative) NPV of a project, the relative rank of adjacent projects owned by peer firms is uncorrelated with the relative rank of a given project within a firm's own portfolio.[9] Second, we show that our results hold when we limit our sample only to the real options with low relative rank within a firm's portfolio, while its peer firms' adjacent projects' relative rank is high. Third, we find that firms still respond to peer exercise decisions on units that are directly adjacent

[8] It is well established that firms cannot pursue all positive NPV projects at the same time due to operational, labor, or capital constraints. Hence project ranking is a commonly used tool to select only the most profitable projects (see Berk and DeMarzo (2014) as an example).

[9] The relationship is not statistically significant, and, if anything, is slightly negative. Further, throughout all specifications, we control directly for the absolute quality of peer firm projects by using the production from the first well of each adjacent peer units as a proxy for the NPV of the peers' adjacent infill wells.

to theirs, even after controlling for peer exercise decisions on projects elsewhere. Fourth, we find that firms' response to adjacent peer exercise decisions is concentrated around the activity from peers with substantial experience in extracting shale in the area of interest. Taken together, these tests make significant progress in addressing the primary endogeneity concerns in measuring responses to peer real option exercise decisions, and set a high bar for alternative explanations. Specifically, an alternative explanation would need to reconcile why the relative NPV rank of a given project in a peer firm portfolio would have a direct effect on a firm's exercise decision for a reason other than peer exercise activity, when that relative rank is uncorrelated with any metric that is linked with the absolute NPV of a project ex ante.

As a final set of analysis, we estimate the optimal stopping (exercise) time based on baseline real option models (e.g., Paddock et al. (1988), Dixit and Pindyck (1994)). After incorporating all of the detailed granular inputs our setting affords into these baseline models, we find that differences exist between actual exercise behavior and predicted exercise behavior. However, we find that the predictions from the baseline model are closer to the actual observed behavior once we take into account information externalities due to adjacent peer exercise decisions. Specifically, if we model beliefs about the value of unexercised infill options to be a function of both the production of the first well on a drilling unit and the adjacent peer exercise activity, we find that exercise decisions are significantly closer to those predicted by theory.

By analyzing peer effects and social learning in the context of real option exercise behavior, our study contributes to two important strands of the literature. First, we contribute to the real option literature by empirically evaluating the importance of a broad set of theories, which hypothesize that information revelation and externalities may be an important component of exercise decisions (Grenadier (1996), Grenadier (1999), Grenadier (2002), Novy-Marx (2007), Grenadier and Wang (2005) and Grenadier and Malenko (2011)). In particular, we show that peer exercise is important relative to the predictions from standard real option models (e.g., Dixit and Pindyck (1994), and Kellogg (2014)). To understand why this may be the case, we focus on a setting where we can directly identify peer effects and the role of information externalities in option exercise behavior (Grenadier (1999)). Using a hazard model framework we show that information externalities from peer effects can have economic effects on the same order of magnitude as natural gas prices and volatility. Second, our novel micro-level evidence of the effect of peer activity on option exercise helps us contribute to the literature on learning from peers. That literature documents that peer effects are important for a variety of corporate decisions, such as those on investment policy (Foucault and Fresard (2014), Bustamante and Fresard (2017)), capital structure policy (Leary and Roberts (2014)), and dividend policy (Greenan (2018)). The economics literature provides evidence on social learning and the adoption of new technologies (e.g., Foster and Rosenzweig (1995), Thompson and Thompson (2001), Conley and Hudry (2010), and Covert (2015)). Covert (2015) in particular relates to this study because it documents social learning on decisions related to what technology to use to drill and complete wells. The evidence Covert (2015) provides is precisely the type of information externality that can make social learning important for real option exercise decisions. However, similar to Covert (2015), much of the existing literature related to social learning is focused on how firms learn and invest (see Conley and Hudry (2010), Covert (2015)). Our contribution is to show that this peer learning also has an important impact on the *timing* of investment decisions, that is within a real options context, peer learning affects *when* firms invest.

The paper proceeds as follows. In Section 1, we provide the institutional background on the natural gas industry and our setting. In Section 2 we discuss our data, and in Section 3 we report our main empirical results based on survival analysis. In Section 4 we compare actual vs. predicted exercise time where predicted exercise time is derived from optimal stopping time theory. Section 5 concludes.

1 Real Options in the Context of Shale Drilling

1.1 Project Overview: Natural Gas Shale Drilling

Our setting exploits the institutional features of natural gas shale development to study real option exercise behavior of firms. To extract shale natural gas, firms must first drill a well with a horizontal leg into shale rock (typically more than a mile below the surface), then complete the well by hydraulically fracturing ("fracking") it. The process of drilling a well may take a few days to a few weeks, while fracking is done as a separate process after drilling, and takes another few days. Both drilling and fracking entail substantial upfront capital costs of \$4.7 million per well on average in our sample. Once a well is completed, it produces natural gas, and declines over time. The critical features determining the profitability of the cash flows are natural gas prices and the volume extracted. Costs include lease operating costs and royalty costs, and typically comprise less than 40% of a well's revenues after the well is drilled. Cash flows are at their highest level at the beginning of a well's life, then decline over time as pressure from the well declines. Once a well starts producing there is little that a firm can do to cause the production to go up or down outside of a wells natural decline without risking damage to a well. Figure 1 plots the cash flows and capital expenditures associated with drilling a well (see Gilje and Taillard (2016) for more details).

1.2 Infill Drilling

One of the key features of our setting is the unique ability to observe the flexibility and maturity that firms have on their investment options. As in Kellogg (2014), we focus on "infill" drilling projects in order to have well defined maturity assumptions. An "infill" project corresponds to the decision to drill additional wells on a drilling unit (section) that a firm already operates. The first (or existing well) on a unit contractually holds the operatorship of the acreage as long as the first well produces; in this case the lease is said to be "held by production" or HBP. A firm has the option to drill additional wells at any point in the future so long as the initial well is still producing. This provides firms with options that have very long maturities as the life of the first well can range anywhere from 20 to 40 years. In all the natural gas shale developments that we study in Oklahoma, a single drilling unit (section) of 640 acres can support up to 8 shale wells (or roughly up to $37.8 million in capital expenditures). With 2,853 units representing up to $107.9 billion in potential capital commitments, the infill options in this study represent capital investments that are economically meaningful, with a significant degree of flexibility on when to exercise these options. Figure 2 plots a timeline of the infill drilling decision.

A key advantage of focusing on infill drilling is that, unlike most studies of investment decisions, we can observe both exercised *and* unexercised options. Indeed, drilling units with only one existing well effectively contain many unexercised options as no additional (infill) wells have been drilled in the unit yet. Our study focuses on the timing of the first infill well in a unit. It is important to note that a firm could delay the exercise of the second, third, and follow-up infill wells. However, we find that 90.2% of all infill wells are drilled concurrently to the first infill well. As such, infill drilling does not seem to be exploratory by nature; but rather, a decision to extract significantly more resources from a unit that has been held by production with the first well up to that point.

1.3 Measuring Peer Activity

The ability to analyze firm's investment responses to competitors' actions is a key novelty of our study. We focus on the development of major natural gas fields across multiple operators, a setting where information and other externalities may be more relevant. This is a key distinction from Kellogg (2014) who focuses on single operated fields, where there is only one firm drilling a field.

The regulatory and land environment in Oklahoma lends itself well to further our understanding of how firms might react to adjacent drilling activity. Specifically, every drilling unit in our setting conforms to Jeffersonian survey, and lies on a grid system with squares that are one mile by one mile. Every 6 by 6 group of squares (36 "units" in total) rolls up to a township survey (township level). This is attractive for several reasons. Every drilling unit, by construction, has eight clearly delineated adjacent units. We observe every natural gas well drilled in Oklahoma so we can observe the exact timing and nature of all adjacent activity throughout our sample period. Second, we will use the township survey information to control for potential geography or area specific effects in our econometric specifications. Figure 3 plots the shale drilling activity in a township. The lines represent the horizontal wellbores of shale wells. Sections in the grid are the drilling units, sections with one wellbore have not yet been infill drilled, while sections with multiple wellbores have been infill drilled.

1.4 Real Option Framework

The firm's option to infill drill corresponds to the choice it has to spend capital to further develop its proven natural gas reserves. As noted in the introduction, the timing flexibility related to the investment decision to drill a well on proved reserves can be viewed as an American call option (e.g., Paddock et al. (1988)). Infill drilling maps nicely into the real option framework: The capital needed to develop the reserves can be viewed as the strike price of the option. The value of the reserves after capital has been expended, that is, the producing proved developed reserves, corresponds to the underlying asset. The timing flexibility a firm has to infill drill can be viewed as the time to maturity. Because the first well on the section holds by production (HBP) the section as long as it is economically viable, the option to infill drill has a long maturity attached to it; at least 20 years on average. And as the decision to infill drill (exercise the option) can be made at any time over this period, it can be viewed as an American call option. The cash flow volatility of infill wells corresponds to the volatility of the underlying asset used in standard option pricing model. Firms in our setting all produce the same commodity, natural gas, and the market provides indicators of expected futures prices and volatility, both of which can be used as inputs for an option pricing model, along with other inputs described in more details in Section 4.

1.5 Optimal Exercise Time and Peer Effects

It is well established that American call options on dividend paying underlying assets have an optimal exercise time that can occur prior to maturity. As Dixit and Pindyck (1994) point out dividends can be viewed as either explicit or implicit in the context of real options, and broadly speaking can be viewed as the benefit a firm obtains from exercising an option sooner rather than later. In our setting, a straightforward way of viewing the cost a firm incurs by waiting is that future cash flows get discounted by a firm's cost of capital. The longer a firm waits to exercise, the more discounting will be applied to the underlying cash flows generated by the well. Conversely, waiting (delaying drilling) confers the ability to drill in future states of the world that exhibit higher natural gas prices. Therefore one can view early exercise as the result of a tradeoff between the value of early exercise from having to discount cash flows less relative to delaying the exercise in order to get better natural gas pricing in the future.[10]

All else equal, higher cash flow volatility tends to result in delayed investment, due to the increased prospects of higher cash flows, while a higher cost of capital tends to result in investment occurring sooner. The classic derivations of the optimal stopping time (see Section 4 for more details) lead to a trigger rule, whereby a trigger value can be computed such that it is optimal to exercise the option when the value of the underlying asset (natural gas reserves) exceeds the trigger value from below for the first time. When natural gas prices rise, it is more likely that the value of the underlying asset will exceed the trigger value. Hence, commodity price increases will lead to earlier exercise of the real option all else equal.

Natural gas prices and natural gas price volatility have clear predictions as to how they might affect exercise based on a standard options framework, with volatility being negatively correlated with exercise (more valuable to delay when volatility is high) and natural gas prices being positively correlated with likelihood of exercise. We also include information on nominal interest rates in our initial tests. Typically a decrease in interest rates decreases the discount rate and hence makes projects more valuable and hence more likely to be undertaken.

[10]As we will see in Section 4, in our context, a firm's cost of capital will correspond to the dividend rate on a stock.

However, in the context of real options, the effect of interest rates is more ambiguous because a decrease in interest rates makes waiting more appealing, as cash flows in the future are valued more today.[11]

Assessing how peer effects alter option exercise behavior is the central focus of this study. There are a broad set of theoretical papers that claim that informational spillovers from peer activity can be of first order importance in understanding real option exercise behavior. The mechanism underpinning these peer effects relate to the information content that is revealed by the exercise of infill drill options on the eight adjacent drilling units (see Figure 6). Specifically, the more infill wells being drilled nearby, the more information there is on the depths and porosity of the formation, which will in turn inform a firm on how to most efficiently extract natural gas from its own infill wells. Additionally, public disclosures require information to be disclosed on particular chemical mixes and techniques of hydraulic fracturing of "fracking" a well (see Covert (2015)). This reveals information on techniques that might work well for fracking a particular reservoir as well as those that might not work as well.[12] It is important to note that, even seeing a negative outcome in terms of production in an adjacent section, that is knowing which "fracking" techniques do not work, will allow a firm to learn how to better extract from its own section. Lastly, adjacent exercise activity by peer firms could also be a reflection of some private information about rock quality a peer firm has which is not yet publicly known; as such, observing adjacent exercise may lead a firm to update positively on the rock quality of a project. Grenadier (1999)'s develops a theoretical framework of real option exercise to assess the potential impact of information externalities from peer exercise activity. All of the reasons listed above justify why we could see positive information externalities from neighboring activity in our setting and thus validate the use of our setting to empirically assess Grenadier (1999)'s main prediction that peer exercise activity will lead firms to exercise early. Within the context of a classic Dixit and Pindyck (1994) framework, the information externalities from peer effects result in higher project

[11]The effect depends somewhat on whether a movement in interest rates (r) will have a commensurate impact on the firm's cost of capital (δ). See Section 5.4 of Dixit and Pindyck (1994) for a more detailed discussion on the topic.

[12]See fracfocus.org for examples of the types of disclosures that are made public.

NPV, pushing firms closer to the optimal "trigger" rule, all else equal.

2 Data

2.1 Construction of Panel for Hazard Model

Our sample period begins in January 2005 and ends in December 2016. We construct a panel of all units (sections) in Oklahoma with one horizontal natural gas well in production.[13] This first well confers the operator the option to infill drill the unit with additional wells as described above. The number of these outstanding available options gradually increases over the sample period as shown in Figure 4C. By the end of our sample in 2016, there is a total of 2,853 infill drilling options, 680 of which have been exercised (~24%). The number of firms (operators) corresponds to 159. Table 1 reports the summary statistics for the panel we use in the hazard model. In total our data is composed of wells in 442 townships across every natural gas shale development in Oklahoma.

Our empirical analysis is based on the panel data of exercise decisions to infill drill on sections that are held by production with the existing well (first drilled) on the section. The unit of observation in this panel is at the drilling unit-month level, in total there are 162,905 drilling unit-months prior to exercise in our sample. To test some of the key predictions of the real option framework outlined in the previous section, we include the 18-month natural gas futures price from Bloomberg and 18 month implied volatility of natural gas prices as in Kellogg (2014). We also include the 5 year nominal risk free rate on U.S. Treasury bond to capture the impact of interest rate movements. All these variables are computed at the monthly frequency.

To proxy for the expected value of the reserves that will be unlocked by exercising the option to infill drill, we compute the present value of future cash flows generated by the infill well using the futures curve for pricing, and an expected production profile based off the

[13] Oklahoma contains both oil and gas. We focus only on wells that are designated as natural gas shale wells on their drilling and completion reports, meaning the primary economic rationale for drilling the well is the recovery of natural gas, not oil. Therefore, natural gas prices and natural gas price volatility are directly related to the investment decision to drill a well in our sample.

unit's first horizontal well's production in its first year.[14] Production data is reported by the Oklahoma Corporation Commission and Oklahoma Tax Commission at the well level. Lastly, we compute the expected drilling and completion costs for each well based on regulatory disclosures of well costs filed with the Oklahoma Corporation commission.

The final set of variables relate to adjacent activity from the firm itself (own) and its peers (competitors). Recall that each section can have up to eight neighboring infill options exercised. We find that on average, over the entire sample period, there are 0.34 adjacent options exercised by its peers and 0.40 by itself. Throughout our regression specifications, to aid the economic interpretations in the tables, we standardize all variables related to adjacent activity (adjacent peer exercise, adjacent firm exercise, and associated relative ranking variables) to have a mean of zero and standard deviation of one. This scaling does not affect the statistical significance of any variables, but does provide an attractive economic interpretation of these variables such that the Hazard Impact factors relate to a one standard deviation change relative to the mean. Table 1, also highlights that the medians are at zero reflecting the fact that many units do not have any infill wells during our sample period, the standard deviations do signal heterogeneity in neighboring activity. We exploit this heterogeneity in our main econometric specifications. To address potential endogeneity concerns, we also compute the ranking of each infill well based on the portfolio of options an operator has at any given point in time. This variable can only be computed on a subset of observations (103,451) and is defined as the relative rank of an infill option based on the quality of the first horizontal well drilled on a drilling unit at a given point in time (see Section 3 for details).

The key event that we use to determine whether an option is exercised is the "spud date" of the first infill well. This is the date when drilling capital expenditure is initiated and the drilling of a second well in the section begins and is directly observable from regulatory

[14] There are many potential ways to model for expected production of a well. We settled on the simplest specification based on the 1st well in the drilling unit. Our results are robust to modeling different types of technological improvements over time. Using the simple approach, we find that using the 1st well's production explains (R-squared) 64% of the variation in the 2nd well's production in the drilling unit (the first infill well exercised).

filings from the Oklahoma Corporation Commission. From this data we know precisely the date, time, firm, location (drilling unit) of the infill exercise decision. Figure 4A plots the number of options exercised over time, while Figure 4B plots the amount of time firms wait to exercise an option for the subset of options that are exercised. Because an option only becomes available to exercise after the first well has been drilled on a drilling unit, the number of options during the sample period is not the same across time. Figure 4C plots the number of options over time, as well as the number of options exercised at any given point in time.

3 Results

3.1 Peer Effects and Option Exercise

To assess the factors that might affect real option exercise behavior, we perform a duration analysis based on hazard functions. The objective of using a hazard function is that it allows us to compute the probability of exercising an option, within an interval, conditional on having not exercised the option up to the time of the interval. Specifically, the hazard function is defined as:

$$h(t) = \lim_{s \to 0} \frac{Pr(t \leq T < t + s | T \geq t)}{s}$$

We parametrize the hazard function using a commonly-used semi-parametric approach:

$$h(t) = h_0(t) \exp(\beta_1 NGPrice_t + \beta_2 NGVol_t + \beta_3 DrillingCosts_t + \beta_4 IntRate_t$$

$$+\beta_5 FirstWellProd_i + \beta_6 AdjExercisedOwn_{i,t} + \beta_7 AdjExercisedPeer_{i,t})$$

This parametrization corresponds to the well-established Cox Proportional Hazard Model, whereby the unit of observation is at the drilling unit-month level. This empirical design determines the factors that make it more (or less) likely that the option to drill the 1st infill well on a unit (section) gets exercised. Once an option is exercised on a drilling unit it is dropped from our sample. Specifically, our duration model specification models the infill drilling decision as a "single-spell" dataset, whereby each individual unit (section) enters the dataset when the 1st well in the section is drilled and exits either when the 1st infill well is exercised (drilled) or is (right) censored if no infill wells are exercised prior to the end of our sample period.[15]

We cluster standard errors at the township level in every specification and provide further robustness tests in terms of econometric specifications in the Appendix. A useful baseline when conducting hazard analysis is to plot the survival function; this allows us to observe the rate at which options are being exercised in the sample, we do this in Figure 5. The plot begins at 1 and then declines as time passes (in months) and options are exercised (and no longer survive). By the end of the time period available 23.8% of all options are exercised. Having established this baseline hazard rate, we can then assess which covariates may cause a shift up or down in the curve in Figure 5, that is, what are the factors that might lead firms to exercise options sooner or later.

The focus of our study is on how neighboring peer project activity affects the baseline hazard rate. To do this, we test the effect of neighboring peer activity on the decision to exercise by calculating the number of adjacent drilling sections (as many as eight) that have infill options exercised by peer firms at each point in time. We include this new variable as well as a measure of the firm's own adjacent activity in the parametrization of the hazard function. To provide context for this peer effect, we include in our baseline specifications the same set of variables as those found in Kellogg (2014). These include natural gas prices, natural gas volatility, drilling costs and interest rates. Recall from Section 1 that standard option theory makes prediction on these variables. For instance, as higher volatility makes the option to delay more valuable, all else equal an increase in volatility should push firms to delay investment. By including volatility of natural gas as a covariate ($NGVol_t$), we can

[15] 90.2% of all infill wells are exercised (drilled) concurrently with the first infill well. That is, when firms exercise their first real option to do infill drilling, they typically exercise many infill options at once. Because infill options tend to get exercised together, modeling the time to exercise of the first infill well is capturing the main economic decision associated with infill real option exercise for the firms in our sample; this modeling decision also allows us to maintain a tractable modeling framework.

assess whether this theoretical relationship holds in the data.

Results are shown in Table 2. We find a strong positive relationship between the likelihood of exercising and peer real option exercise activity. To facilitate the interpretation of the adjacent real option exercise variables, we standardize the variables to have mean of zero and standard deviation of one, so that each coefficient/Hazard Impact factor can be interpreted as a one standard deviation change relative to the mean. Specifically, a one standard deviation increase in adjacent peer infill exercise activity increases the likelihood that a firm will exercise its infill option by between 10.9% and 38.1% depending on the specification. This result is supportive of Grenadier (1999)'s main prediction that information externalities play an important role in the exercise decisions of firms.

As in Kellogg (2014) we find that natural gas prices and natural gas volatility affect real option exercise decisions. Namely, we find that higher volatility reduces the hazard rate (the rate at which options are exercised). Conversely, natural gas prices ($NGPrice_t$) have a positive effect on the hazard rate, as an increase in the natural gas price increases the value (NPV) of the project and makes the option to delay less valuable. In economic terms, based on the Hazard Impact percentage in specification (1) of Table 2, we find that a one standard deviation increase in natural gas price volatility decreases the likelihood of exercising an option by 14.0% (-3.23*4.32) relative to the baseline hazard rate. Alternatively, a one standard deviation increase in the price of natural gas increases the likelihood of exercise by 26.1% (14.77*1.77) relative to the baseline hazard rate. These results hold across the three specifications of Table 2. They suggest that firms' behavior is directionally consistent with these key predictors of option exercise activity. Furthermore, these magnitudes provide important context for our peer effect results. Specifically, peer effects have an economic impact on the same order of magnitude as baseline real option model inputs such as natural gas price and volatility.

Lastly, we also control for the quality of the first horizontal well drilled in the unit as well as the estimated cost of the infill well in specifications (2) and (3) of Table 2. The intuition behind the first of these controls is that the first well is an indicator of the quality of the geology in an area: the more it produces, the higher the NPV of the additional infill projects, and hence the more likely the option to infill drill will be exercised. Results in Table 2 support this hypothesis. Specifically, a one standard deviation increase in the quality of the first well results in an 88.5% (51.48*1.72) increase in the likelihood of exercise. Drilling costs will vary over time; for instance wages for qualified workers were rising over our sample period (e.g., Bartik et al. (2018)). These time-varying costs could affect option exercise behavior by changing the strike price over time, so controlling for time-varying drilling costs is also important. Results from Table 2 show no significant impact of drilling costs on the likelihood of exercising early, similar to Kellogg (2014)'s finding.

3.2 Endogeneity: Peer Effects and Option Exercise

A potential concern with the interpretation of Table 2 is that the correlation between a firm's exercise behavior and its competitors' adjacent exercise activity cannot necessarily be attributed to a *reaction* to adjacent activity (Manski (1993)). For example, a common factor, such as shared technology or similar reserve quality, could affect both the adjacent competitors' decisions to exercise as well as a firm's own decision to exercise. To address this concern, we need to identify the exogenous component of adjacent exercise activity.

3.2.1 Defining the Instrument for Peer Activity

For the construction of our measure of exogenous variation of peer activity, we start from the observation that firms typically face operational, labor, or capital constraints and thus are unlikely to undertake all positive NPV projects at once. As such, they make decisions to invest not just based on the absolute NPV of a project, but also on the relative NPV or rank of a project in a firm's portfolio of capital projects.

The measure we construct can best be illustrated with an example. Figure 7 shows the real options of three firms. Firm A has two separate drilling units, each of which is adjacent to drilling units owned by Firm B and Firm C. Now assume that the NPV of Firm A's infill projects and the infill project adjacent to it, owned by peers, is $1 million. However, let's also assume that Firm B has a portfolio of four additional real options with NPVs, if exercised today, of \$2 million, \$3 million, \$4 million, and \$5 million respectively. Alternatively Firm C has a portfolio of real options with an NPV, if exercised today, of \$0.90 million, \$0.50 million, \$0.30 million, and \$0.20 million. All firms have positive NPV projects, but for Firm B the project adjacent to Firm A is ranked fifth among its portfolio of projects, while for Firm C it is ranked first. Now assuming that these firms face some operational, labor, or capital constraints, and firms can only undertake one project at a given point in time.[16] Based on the rankings of these projects, we would expect Firm B will not exercise its project next to Firm A, while Firm C will, even though the projects have the same absolute NPV. When Firm C exercises, Firm A benefits from the information on how to complete the well, and information on the depths of the zone to target, while it has no new information for its project next to Firm B. Therefore Firm A benefits from an information externality not due to any shared or common characteristic of the specific real option in question, but due to the ranking within the existing portfolio of the other real options that Firm C has. The identification assumption is that the rankings of the projects in firm B and Firm C's portfolios is exogenous relative to the investment opportunities that Firm A has. We offer several tests in the next section to document that the project value of a given firm's option is unrelated to the relative ranking of the adjacent options owned by peer firms.

In Table 3, we empirically test whether rank ordering matters in option exercise decisions. The variable we construct is the relative percentile of each infill project in a firm's portfolio. Our rank ordering is based on the production of the first horizontal well on a drilling unit.[17] For every month in the sample, for every firm, we rank the total number of natural gas infill real options the firm has across the entire state of Oklahoma as of that point in time, and then map that rank ordering to percentiles.[18] So, for example, if a firm has 20 real options

[16]Our analysis assumes all projects have the same investment cost at a given point in time, a reasonable assumption in our sample as Gilje and Taillard (2016) provide evidence that investment cost does not vary significantly across firms in a given region for shale gas development.

[17]We assessed the potential of several alternative measures for project ranking, including adjusting the production of the first well by its vintage. We found that the unadjusted first well production had the highest explanatory power over infill production, relative to any alternatives. Additionally, we find no variation in the explanatory power of the first well production for infill productivity based on whether the well was drilled early on or later in the shale development.

[18]Ideally one would want to have visibility to all real options a firm has, including those outside of Oklahoma. However, despite being limited to all infill options in Oklahoma we still obtain strong statistical

in its portfolio, the number one well would be in the 95th percentile (1-1/20). As can be seen in Table 3, the higher the percentile rank in a firm's portfolio, the more likely it is that the project is exercised. To ease the interpretation of the relative rank percentile coefficients, the data has been normalized to have mean of 0 and standard deviation of 1. Therefore, based on the different specifications found in Table 3, for a one standard deviation increase in percentile, a firm is between 65.8% and 84.9% more likely to exercise an option.

3.2.2 Instrumental variable approach

Table 4 Panel A reports both the first stage and second stage estimations, where Adjacent Peer Exercise Activity, defined as the number of infill options exercised by peers adjacent to the drilling unit i at month t, is the variable that is instrumented.[19] The instrument we construct is the average relative percentile of all adjacent drilling units owned by peer firms as of month t based on the relative rank of each adjacent infill project in a peer's portfolio of projects. The relative ranking of each infill project will fluctuate over time; for example, if a peer firm adds real options with strong first wells elsewhere, then the relative percentile will go down. If it adds real options with relatively poor first wells elsewhere, then the relative percentile will improve. We include all control variables from the second stage of our model in the first stage. The first stage regression is given by:

$$\begin{aligned} \#AdjacentExercisedOptionsPeer_{i,t} &= \\ &\beta_1 AvgRelativeRankPercAdjacentPeerProjects_{i,t} \\ &+Controls + TownshipFE + \varepsilon_{i,t} \end{aligned}$$

The second stage is given by the Cox proportional hazard model whereby the covariates are comprised of our instrumented variable for neighboring peer activity from the first stage, as well as a series of additional control variables. We correct for the estimation error in the

power from using the rankings of real options in our sample (see Table 3). This is consistent with the notion that drilling decisions are typically made at the play/regional level and as such the portfolio ranking within Oklahoma seems appropriate.

[19]Table 4 has fewer observations than Table 3 because we can only use our instrument once some adjacent peer infill options exist: if a firm's real option to infill has no adjacent infill options then there is no relative rank from an adjacent peer that can be used to construct the instrument.

first stage in our Cox two-stage IV model by bootstrapping the standard errors (MacKinnon (2002)). The appropriateness of this approach has been supported in recent literature (see Tchetgen et al. (2015)).[20]

As can be seen across the different first stage specifications at the bottom of Table 4 Panel A, the average relative rank of the adjacent real option peer projects has high predictive power for the adjacent peer exercise activity. In our second stage estimations, we control directly for the absolute NPV of adjacent peer infill projects by including the average production from the first (pre-infill) well of adjacent infill peer options as a control. The underlying assumption of this instrument is that the only dimension through which it affects our key dependent variable of interest, the exercise decision of a firm, is through the exercise behavior of peers. We provide a number of tests supporting this assumption in Section 3.2.4. Among the control variables, the only one that loses significance in the instrumental approach (relative to Table 2) is the implied volatility of natural gas prices. We directly test whether our instrument is correlated with implied volatility. The correlation between implied volatility and our instrument is slightly negative, -0.0245, but not statistically different from zero. Further, while the coefficient does lose its statistical significance, it remains firmly in the general range of the baseline estimates. Given the economic channel through which the instrument affects peer activity, this evidence does not suggest that our instrument is operating through any effect on the implied volatility.

Overall the results from Table 4 suggest that the economic interpretation from Table 2 still holds when we use an exogenous source of variation in adjacent activity driven by the relative rank of projects in peers' portfolios. For ease of economic interpretation for our key variable of interest, we report the coefficient on the standardized variable, so each coefficient/Hazard Impact factor can be interpreted as a one standard deviation change relative to the mean. As such, a one standard deviation increase in our instrument leads to between a 79.1% and 94.0% increase in the likelihood of exercising the option to infill drill.

[20]We document the robustness of our main two-stage models by estimating both IV Probit and IV 2SLS models on our data and obtain similar results to our main Cox model tests, see Appendix Tables 1, 2, 3, and 4 and related discussion in Section 3.2.3.

3.2.3 Robustness Tests

We first report the reduced form results in Table 4 Panel B for robustness. This regression is still subject to the exclusion restriction, which in our case means that the relative ranks of adjacent projects only affect a firm's decision to exercise via the relative rank's effect on adjacent peer project exercise decisions. By not instrumenting we lose the economic interpretation of the coefficient on the number of adjacent peer exercised options, but maintain the overall intuition of the result reported in Table 4 Panel A: firms' exercise decisions are affected when a project has plausibly exogenous exposure to a variable that affects adjacent exercise behavior (relative rank percentile of adjacent peer projects (β_6)).

We retain the Cox model as the primary specification in the paper because we are studying the motivation behind the decision to exercise real options, and this decision is dynamic by nature: firms have to decide in each period whether to exercise or not, conditional on not having exercised until then. A natural econometric specification for this is the duration model (as in Kellogg (2014)). The hazard function allows us to approximate the probability of exercising the option, conditional on having not exercised until then. This modeling has been used in other contexts in corporate finance (e.g., Leary and Roberts (2014)) and has several advantages. One of the main advantages in the context of our study being that the hazard function can easily be made to depend on time-varying variables and has a natural interpretation.

Linear probability models and probit specifications both face several drawbacks. First, even though the decision to exercise is binary, a linear specification implicitly assumes that the outcome variable can be non-binary and even negative. This is one drawback of using the linear probability model. Second, both the linear and probit models are not well suited to capture the dynamic nature of the decision to exercise. Even for probit (or logit) models that accommodate for the binary nature of the left-hand side variable, these modeling approaches aim to explain the proportion of exercised options across the entire sample at any given point in time, which is different from what the hazard models capture in terms of the variables that influence the probability of exercise at time t, conditional on not having been exercised up to that time. Third, the censoring of the data is another impediment to implementing traditional methods such as linear probability models or probit regressions. In our setting, the censoring bias is caused by the fact that we only observe the data until the end of the sample (right censoring); for firms that do not exercise prior to the end of the sample period, we only know that they did not exercise their option until that point in time. Although the linear and probit specifications do not have a natural way of handling this right censoring issue, the maximum likelihood estimations (MLE) of Cox hazard models are well suited to handle this specific type of right censoring (see 20.3.2 of Wooldridge (2002)).

That being said, estimating models using the IV-2SLS (two stage least squares) and IV-probit frameworks is informative in assessing the robustness of our estimates to the choice of estimation model. As such, we perform two other specifications for the IV approach based on an IV-probit and IV-2SLS specification for which the statistical properties are well established. Namely, in Appendix Table 1, we run an IV-probit specification where the second stage is a probit modeling of the exercise decision instead of a duration model. The coefficient on the instrumented adjacent drilling activity of peers is positive and significant. Appendix Table 2 provide the results for the IV-2SLS specification. Again, we find a positive and significant loading on the instrumented adjacent peer activity variable.

Throughout all our main specifications, we have clustered the standard errors at the township level. In Appendix Table 5 and 6, we re-run Table 4 Panel A and B, but this time allowing for clustering at the township and year level (double clustering). Our results remain robust to the double clustering approach.[21] The double clustering results typically yield smaller standard errors (i.e. higher t/z statistics) than one way clustering by township, hence to be conservative we report township clustering for our main results.[22] Taken together, the evidence in this section suggests that our primary findings are robust across several different econometric specifications.

[21]Appendix Table 3 and 4 also provide further support for the results found in the context of the IV-probit and IV-2SLS specification when clustering of standard errors at the township and year level (double clustering).

[22]We document in Appendix Table 7 that our main results are robust to including a control for the first well being drilled (“purchasing an infill option) and in Appendix Table 8 we document that our main results are robust to including operator fixed effects..

3.2.4 Internal Validity

In this subsection, we undertake several falsification tests to assess the validity of the instrument we outline above. While the exclusion restriction cannot be tested directly, we can assess the plausibility of some potential explanations which would invalidate our instrument.

One potential explanation which might be problematic for our instrument would be if all firms had similar locations for their high percentile wells. For example, if all firms had their 90th percentile wells in one township, and their 80th percentile wells in another, such clustering would render inference problematic. Although our main tests include specifications with township fixed effects and township level clustering, which would control for an overall township effect; if there is clustering within townships of high percentile groups in some areas and low percentile groups in other areas it would be problematic as one could argue the instrument might proxy for the absolute value of the NPV of a project and not just the relative NPV of a project. We also control directly for production from adjacent peer wells, which should alleviate this concern to some extent. Nonetheless, we can also directly assess the impact of this possibility when we regress the relative rank of a real option in a firm's portfolio on the relative rank of the real options owned by peers that are adjacent to it at a given point in time, as in the below regression.

$$\begin{aligned} RelativeRankPercOwnProject_{i,t} = &\\ &\beta_1 AvgRelativeRankPercAdjacentPeerProjects_{i,t} \\ &+TownshipFE + \varepsilon_{i,t} \end{aligned}$$

The unit of observation is at the drilling unit i, month t level, and we estimate the OLS regression in Table 5. As can be seen the coefficient is neither statistically nor economically significant, suggesting that once township fixed effects are controlled for (as they are in our main specifications in Table 4), there is no correlation between the percentile rank of a given real option and the average percentile ranks from adjacent peer firms' surrounding real options. This test provides evidence against the idea that all firms have their 90th percentile wells clustered together somewhere, and their 80th percentile wells clustered somewhere else in a way that would confound our tests.

Conceptually, this makes sense as prior to any wells being drilled firms go out and lease drilling acreage when not much information is known about the natural gas resource. Firms thus end up with different portfolios which can be quite dispersed in terms of their potential (see Figure 8); this is the variation that is being exploited with our instrument.

An alternative way to test whether the clustering of relative project quality is driving our results is to look at situations where a real option is ranked low in a given firm's relative percentile rank (below median) while the adjacent real options are ranked highly based on peer relative rank (above median). We report results on this subsample of real options with highly dispersed relative rankings in specifications (1) and (2) of Table 6, and as can be seen from the table, our main result holds.[23] Overall we find magnitudes higher in these tests than our baseline regressions, which is consistent with the idea that information externalities become more important when relative ranks are more dispersed.

Another potential concern with our identification is whether a firm exercises its real option because of the action of a competitor (adjacent exercise) or a characteristic of an adjacent competitor as described in Manski (1993).[24] For example, one might imagine that a competitor exercising their option on an adjacent drilling unit might also be pursuing significant drilling activity (exercising other real options) elsewhere in the play, which might signal, for instance, an overall improvement in extraction technology going forward. In this case, a firm and its competitor are both deciding to exercise options that are adjacent to each other, but it is not because the firm is responding to information externalities from the competitor's actions taken on the neighboring drilling unit, but rather, due to the general activity of the competitor taking place both nearby and elsewhere.

To assess empirically whether our main coefficient of interest for peer effects is affected by such characteristics, we look at competitors with adjacent drilling units and test whether their drilling activity *outside* of the township also bears an influence on a firm's decision to exercise. We include this measure as an additional explanatory variable ("Play" activity) in

[23]Township fixed effects for this model are not well identified due to the dramatically reduced sample size, much of the sample is absorbed by township fixed effects.

[24]Leary and Roberts (2014) articulate this issue in detail as it relates to their capital structure analysis.

our hazard regression in Table 7. We find that our main coefficient of interest for peer exercise activity is unaffected by the inclusion of this control variable. Furthermore, we also find no consistent direction in the effect of the "Play" activity variable across model specifications. Overall, this evidence supports the view that firms are influenced by peers' activity when it occurs on the drilling units directly adjacent to them, consistent with the information channel hypothesized.

3.3 The Information Content of Adjacent Exercise Activity

After having established that firms react to neighboring exercise activity when making their own exercise decisions, we set out to investigate the possible channels behind this result. To do so, we re-estimate the hazard model from Table 2 with adjacent exercise activity as an explanatory variable, but this time, we decompose the adjacent exercise activity by competitor type. In particular, we define experienced and inexperienced competitors as those with above (respectively below) median drilling activity in the shale play at the time of exercise.

In an information transmission framework where agents do not have perfect information on the value of their drilling prospects, operators will look for informational cues from more experienced operators about the drilling opportunities in and around their own prospects (e.g., Grenadier (1999)). Moreover, the type of information disclosed via well completion and fracking reports is likely more useful when performed by more experienced firms that are higher up the learning curve in a given resource development. Under this hypothesis, we would expect firms to react strongly to adjacent exercise behavior from experienced operators.

In Table 8, we show the results of our empirical decomposition of neighboring activity. We standardize both of our inexperienced and experienced adjacent activity variables so that we can more readily make a direct comparison between the two coefficients. Specifically we normalize these variables to have a mean of zero and a standard deviation of one. We find that firms exhibit a strong reaction to the adjacent exercise activity of experienced competitors. The economic magnitudes are similar to Table 2 results. These results are supportive of

Grenadier (1999), whereby operators make specific inferences from their competitors' exercise of real options. In particular, their exercise behavior is influenced by the exercise activity of experienced operators, and thus experienced operators seem to be creating positive informational spillovers when exercising their real options.

4 Real Option Modeling

In this section of the paper, we aim to relate the observed exercise behavior to the optimal exercise behavior predicted by real options models. Our data provides us with the unique ability to compute the inputs a firm would have if it were to follow real option decision rules following the classic real options models of Paddock et al. (1988) and Dixit and Pindyck (1994). We calibrate these models to our data in order to derive optimal exercise thresholds, i.e. conditions to be satisfied if firms are to exercise in an optimal manner. We then adjust the framework assuming that firms benefit from some positive information externality from adjacent peer real option exercise activity and compare both approaches to the exercise behavior we observe in the data. In Appendix, we provide results that extend those of this section by calibrating the dynamic discrete choice model of Rust (1987), which was first applied to the oil and gas industry in Kellogg (2014). Our main conclusions remain unchanged.

4.1 Real Option Pricing and Optimal Exercise Time

As mentioned earlier, the option to expend capital in order to develop shale reserves (infill drilling) corresponds to a real option. Firms can decide when to exercise these real options and a large body of work has been developed to establish both the pricing of these real options as well as their optimal exercise (stopping) time. We provide the resulting standard pricing formula as well as the optimal "trigger" rule below.[25]

It can be shown under rather general assumptions, that the value of an option $F(V)$ is

[25] The derivations can be found in Paddock et al. (1988) and Chapters 5.2 (p. 140-143) and 12.1 (p. 396-403) in Dixit and Pindyck (1994).

given by the following closed-form solution:

$$F(V) = AV^{\beta_1}$$

where

$$\beta_1 = \frac{1}{2} - \frac{(r-\delta)}{\sigma_P^2} - \sqrt{\left[\frac{(r-\delta)}{\sigma_P^2} - \frac{1}{2}\right]^2 + \frac{2r}{\sigma_P^2}}$$

and

$$A = \frac{(\beta_1 - 1)^{\beta_1 - 1}}{(\beta_1)^{\beta_1} I^{\beta_1 - 1}}$$

Whereby r is the risk free rate, δ is the dividend rate of the project and σ_P is the volatility of the underlying project value.

The optimal time to exercise a real option is similar to the optimal time to exercise an American call option. It can be shown that the optimal stopping time is given by a trigger rule under very general assumptions.[26] Specifically, there exists V^*, a trigger value, such that for when the underlying asset value crosses V^* from below for the first time, it is optimal to exercise. Defining I as the drilling costs of the well, the trigger value is given by:

$$V^* = \frac{\beta_1}{\beta_1 - 1} I$$

When exercising at the optimal threshold, the firm gets $V^* - I$ where I is the infill drilling capital expenditure. In the NPV framework, the NPV rule would state that one should drill as soon as $V = I$ or equivalently $\frac{V}{I} = 1$. Given that $\beta_1 > 1$, $V^* > I$, or equivalently $\frac{V}{I} > 1$, i.e. there is a wedge between the NPV rule and the optimal exercise rule. Given the option value to delay, the value of the underlying asset needs to exceed the investment cost (and in some cases by a large margin) before it becomes optimal to exercise.

[26] Certain conditions need to be met for this trigger value to exist and be unique (see Kellogg (2014)).

4.2 Data Sources for Real Option Input Variables

An attractive feature of our setting is that we are able to obtain all of the inputs needed to compute the real option decision model for firms within the standard real option framework. The critical components of this data include data on drilling costs, cash flows, natural gas prices, and natural gas implied volatility.

To obtain the cost of each well we collect data from the Oklahoma Corporation Commission pooling regulatory documents. This data is disclosed by all firms who initiate the drilling of the first well in a drilling unit, and is used by other firms with ownership stakes in the drilling unit who are deciding whether to participate in the well or not. Drilling costs fluctuate due to the supply and demand for drilling and completion services, and vary little across operators and geography within a shale basin at any given point in time (Gilje and Taillard (2016)), but they do vary substantially over time. We have detailed data on 996 wells to estimate the costs of the infill wells in our sample.

Cash flows from wells are based on production, prices, lease operating costs, and royalty costs. Production data at a monthly frequency on every well in our sample is available from the Oklahoma Corporation Commission and Oklahoma Tax Commission. We also use this detailed month-level data to derive the depletion rate, ω, a key parameter in standard real option models. Natural gas prices and the implied volatility of natural gas are obtained from Bloomberg data on natural gas price futures contracts. As in our main hazard model specifications, we use the 18 month futures price of natural gas and 18 month implied volatility of natural gas as estimates of the overall gas prices and implied volatility of natural gas over the life of a well (consistent with Kellogg (2014)). Lease operating costs are based on estimates obtained from 10-Ks. Specifically, we collected data on lease operating costs from public firms in our sample, and found that on average during our sample time period lease operating costs were 21.6% of cash flow. Royalty estimates are based on royalty percentages obtained from DrillingInfo on 322,340 oil and gas leases signed in Oklahoma, the sensitivities we report encompass a range that is covered by 87.7% of the royalty terms in the sample.

Taken together, the institutional environment in Oklahoma allows us to calculate each of the key inputs that are needed to compute the trigger value (V^*) and thus compare the exercise decisions made in our sample relative to those predicted by standard real option models such as those of Paddock et al. (1988).

4.3 Calculation of Real Option Exercise Thresholds

For every drilling unit, we need to determine first the expected value of the developed reserves that can be tapped by the new infill well (V), the optimal exercise threshold (V^*), and the value of the undeveloped reserves, i.e. the value of the option $F(V)$. We rely on the closed-form formulas derived from option theory in Section 4.1.

To obtain the expected value of the well's developed reserves (V), we rely on a set of commonly used assumptions. First, we make the simplifying assumption that the 18 month futures price of natural gas (P) can be used to compute the value of the stream of cash flow over the entire life of the well and that firms' discount their cash flows at the discount rate μ. Second, the net profit per mcf is obtained by taking into account the operational cost (ϕ), the royalty rate (ρ), the accounting depreciation rate (θ) and the corporate tax rate (τ). Then, we define the net profit (per mcf) as: $\Pi = P[(1-\phi-\rho)-\tau(1-\phi-\rho-\theta)]$. Finally, we assume that the well's reserves are being depleted at the exponential rate, ω, which enable us to model the value for one mcf of developed reserves as $V_0 = E_0 \int_0^\infty \omega \Pi e^{-(\omega+\mu)t} d_t = \frac{\omega\Pi}{\mu+\omega}$.

4.3.1 Estimates of the Model Parameters

In the baseline scenario, we set the discount rate at 10%, in line with the SEC guidelines in valuing reserves and recent empirical work estimates.[27] Then, we estimated the reserves' depletion rate, ω. From the exponential depletion rate formula of the reserves, we have that the monthly production at a given point in time t is equal to $Prod_t = \omega R_0 e^{-\omega t}$, where R_0 is the initial available reserves. Since we only observe the monthly production, but not the initial available reserves, we compute the ratio of monthly production such that $\frac{Prod_t}{Prod_{t-1}} = e^{-\omega}$. As such, for each well we empirically estimated ω from the ratio of monthly productions. In

[27] See Kellogg (2014). In the sensitivity section, we run the calculation using discounting rate ranging from 7.5% to 12.5% on an annual basis.

our sample, the average well has an annual depletion rate of 27%. In a subsequent section, we run a set of sensitivity analysis varying the ω parameters from 25% to 29%, roughly representing the 90% confidence interval of the depletion rate distribution. For the royalty rate, we obtain the lease data from all the wells in Oklahoma, and the median royalty rate is equal to 18.75%.[28] Finally, we set the depreciation rate to 40% and the effective tax rate at 0%.[29]

4.3.2 Developed Reserves Value Calculation

We obtained above the value of developed reserves per mcf: V_0. To obtain the expected value of the total reserves accessible by the infill well, V, we need to multiply the expected total production of the developed reserves by the value of developed reserves per mcf ($B_0 * V_0$). For that, we need to compute the expected amount of natural gas produced by the infill well (B_0). We compute the expected total reserves of the infill well in three steps. First, using realized data from past infill wells drilled in Oklahoma, we regress the first year of production of the second well on the first year of production of the first well for each section. Second, we use the estimated regression coefficient to obtain a *prediction* of the infill well's first year of production ($Production_{t=1}$). Finally, under the assumption of geometric decline curves, we obtain the expected total reserves tapped by the infill well by computing $B_0 = \frac{Production_{t=1}}{\omega}$.

4.3.3 Optimal Threshold Calculation

To obtain the optimal threshold value of each wells, we need an estimate of the well drilling cost (I). To allow for time varying drilling costs, we obtained well level drilling cost and estimated the expected drilling cost for every month of our sample. We then compute the optimal threshold value (V^*) using the results derived in the previous sections. The computation of V^* depends on δ, the implicit dividend a firm generates from a project. Dixit and Pindyck (1994) show that δ equals a firms risk adjusted cost of capital (μ) minus the

[28]In our sample, the average royalty rate is 19.05%, but the industry standard is 18.75%, and 79% of the lease data has a royalty rate of 18.75%.

[29]During the covered period, natural gas exploration firms benefited from multiple generous deduction and tax credit. It enabled them to virtually pay no cash taxes.

expected appreciation of the project (α), $\delta = \mu - \alpha$. The intuition behind this result is that the effect of discounting can be offset by the expected appreciation of the asset. For the purpose of our study we assume that expected appreciation of the asset (its drift) is zero. This baseline assumption is reasonable given that the natural gas futures curve is relatively flat throughout our sample. In this case δ simplifies to a firm's cost of capital. From the definition of V^*, the higher the cost of capital, the smaller the wedge between the NPV rule and the optimal trigger rule. We explore a wide range for δ in the next section.

4.4 Exercise Behavior: Actual vs. Predicted

In this subsection we calculate the real option decision rules that firms would have if they followed the behavior predicted by real option theory and compare this predicted exercise behavior with their actual exercise behavior described earlier in the paper. Over the period of interest, there are a total of 2,853 potential infill well real options available. Of these infill well options, 680 are exercised. The objective of this section is to assess whether at the time of exercise, the project expected values (V) were above their trigger value (V^*) implied by the standard real option model of Dixit and Pindyck (1994), as outlined above.

4.4.1 Actual vs. Predicted: Full sample

In our baseline assumption case shown in Panel A of Table 9, we find that infill projects have an average NPV of \$1.92 million at the time of exercise. The distribution of NPVs at exercise time is shown in Figure 9 and clearly shows that a majority of infill wells are positive NPV projects at the time they are exercised. However, the estimated optimal threshold value V^* at exercise is higher than the estimated present value of the oil reserves V. For our baseline assumption case, Table 9 shows that firms forgo on average \$0.42 million (\$0.42 = \$7.08 - \$6.66), with a median forgone value standing at more than twice that number.[30] Figure 10 plots the distribution of forgone value at exercise time. The histogram clearly shows that the majority of the wells are exercised when V minus V^* is negative (i.e. $V < V^*$), reflecting the

[30]In Table 9, we report $V^* - V$ to show positive numbers for the estimated forgone values. In the Figures, we compute histograms of $V - V^*$, in which case values below zero represent forgone values.

fact that most exercise decisions result in forgone option value due to early exercise. This conclusion is only reinforced by running a similar exercise with a more advanced model in Appendix, whereby we estimate a dynamic discrete choice model (see Rust (1987)), that also allows for both volatility and drilling costs to be stochastic (see Kellogg (2014)).

Although we have actual project-level data to estimate inputs for our model parameters, it is prudent to assess how sensitive the project values and trigger values are to changes in model parameters. In Table 9 Panel B we report sensitivities across every major parameter in the model: depletion rate, operational costs, discount rate, taxes, and royalties. As expected, the NPV of the average (and median) well goes down as the (1) discount rate, (2) operational costs, (3) tax rate, (4) depletion rate and (5) royalty rate increase. More importantly, this sensitivity exercise informs us on how the forgone value (baseline mean of \$0.42M and median of \$0.87M) changes due to changes in underlying parameters of the model. In each case, both the average and median forgone option value in our sample remain positive and statistically different from zero. It is important to note, however, that these computations do not incorporate any updating from peer exercise behavior.

4.4.2 Actual vs. Predicted: With and Without Adjacent Activity

The previous sections did not consider the potential information externalities generated by adjacent peer exercise activity. Specifically, the infill well's expected production was simply a function of the unit's first well's realized production (see 4.3.2).[31] We now compare firms' second well expectations with their actual realizations. Under rational expectations, second well realizations should be, on average, close to the firms' expectations. The goal of this exercise is to identify the role that adjacent peer exercise decisions may play in firms' expectations on second well recoveries.

We break the sample into two groups: (1) the wells with no adjacent peer activity and (2) the wells with adjacent peer activity.[32] For both groups, we compute the deviations be-

[31] I.e. predicted values for the second well (exercised infill option) are based on estimating second well production on first well production in an OLS framework.

[32] Out of a total of 680 exercised options, we have 635 infill wells (second well in unit) with at least one year of realized production. Out of those 635 infill wells, 214 have adjacent peer exercise activity and 421

tween the realization of the second well and the expectation of the second well based only on information conferred from the first well's production. We find that for second wells with no adjacent activity, forming the expectation based solely on the first well's production leads to deviations that are not statistically different from zero on average. However, for the second wells with adjacent activity, we found that the first well's production realizations do not adequately predict the second well realization. The deviations are positive and statistically different from zero (p-value of 0.066). In other words, adjacent peer activity is associated with significantly higher well productivity, after conditioning for the first well's realized productivity.

Under the assumption that firms form appropriate expectations for their infill wells, such evidence suggests that updating from operators takes into account the information conveyed by peer activity. To incorporate updating of expectations based on adjacent peer activity, expectations now stem from (1) the unit's first well's realized production and (2) the adjacent units' peer exercise activity. We operationalize the updating on this second dimension by using an indicator variable that takes the value of one if there is one or more adjacent infill real options that have been exercised by peer firms.[33] Specifically, we find a statistically and economically significant positive loading on adjacent peer activity when explaining the realized production of second wells based on this augmented set of two variables. The coefficient on the adjacent activity dummy variable is 119,323, which can be interpreted as firms revising up production on the second well by 14.4% relative to the average forecasted production based only on the first well's production if adjacent peer activity occurs. This effect is statistically significant at the 1.9% level.

This result is consistent with one of Grenadier (1999)'s main assertion that real option exercising from peers conveys an informative signal. Namely, units with more adjacent real options activity are more likely to hold greater reserves. It is also consistent with the findings from the broader literature that documents the importance of peers and "social learning" in

have no adjacent peer activity.

[33]I.e. predicted values for the second well in the unit (exercised infill option) are based on estimating second well production on first well production and an indicator variable for adjacent peer exercise activity in an OLS framework.

technological adoption (see, for instance, Griliches (1957), Foster and Rosenzweig (1995), Thompson and Thompson (2001), Conley and Hudry (2010), and Stoyanov and Zubanov (2012)). Specific to the oil and gas industry, Covert (2015) shows that there is some degree of technological sharing across peers in shale drilling techniques (e.g., optimal mix of sand and water used in fracking). This finding could also be at work in our context as firms learn how to improve extraction from reserves by observing how peers drill wells in leases adjacent to theirs.

The next logical step in our analysis is to assess whether incorporating information from adjacent peer activity makes the decisions to exercise closer to those predicted by theory. To do so, we compute V minus V^* under the two different information sets, one information set that relies on the first well's production exclusively and one information set which incorporates both the first well's production and an indicator for adjacent peer exercise activity. Recall that the optimal trigger threshold V^* is invariant to productivity expectations of the infill well. However, the expected discounted value of the developed reserves of the infill well, V, depends on its expected productivity. Figure 11 Panel A (respectively Panel B) plots the histogram of V minus V^* for the subset of infill wells with (respectively without) adjacent peer exercise activity.[34]

Panel A of Figure 11 reveals significant differences between the distributions of V minus V^* across the two different information sets used to form expectations. This difference can be explained by the fact that V is revised upwards under the information set that takes into account adjacent peer activity. Comparing the proportions of options exercised too early (i.e. $V < V^*$ at the time of exercise), we find that 57% of infill wells are exercised too early under the first information set, relative to only 44% when the information set is augmented to take into account adjacent peer exercise activity. These differences are statistically significant at the 1% level. This evidence suggests that updating expectations for the productivity of the infill well based on adjacent peer activity leads to an approximate 20% reduction in the likelihood of exercising too early.[35]

[34]Each bin represent a \$1M interval.

[35]For completeness, and as a falsification, we show Figure 11 Panel B that reveals no meaningful differences between the distributions of V minus V^* across the two different information sets used to form expectations

The results shown in this section allow us to show that through a basic updating framework, incorporating information on adjacent peer exercise decisions helps to explain a portion of the gap between V and V^* using a baseline Dixit and Pindyck (1994) framework. We do not observe the full model that firms use for either updating of beliefs or real option decision making, and there may be important additional components to such models, which we do not include here. However, the objective of our exercise is to demonstrate that under a basic set of assumptions on real option modeling and a plausible framework for updating, adjacent peer exercise activity could play a first order role in explaining the gap between actual and predicted behavior for real option exercise. Overall, this exercise provides useful context for our empirical results in Section 3.

5 Conclusion

In this paper we exploit detailed data on a large set of real options to empirically characterize the option exercise strategies employed by firms. We find that peer exercise behavior via an information revelation channel is as important in explaining exercise activity as standard real option inputs such as commodity prices and volatility. To date, the empirical real options literature has been limited, largely by data constraints. Our paper provides important micro-level evidence on both how real options are exercised, and which channels are important in explaining exercise behavior. Our results provide novel empirical support for the importance of information revelation from competitor exercise behavior in explaining how firms exercise real options.

when looking at the subset of wells without adjacent peer exercise activity. This result should not come as a surprise as we know from above that the big difference in expected productivity comes from observing adjacent activity.

References

Bartik, A., J. Currie, M. Greenstone, and C. R. Knittel (2018). The local economic and welfare consequences of hydraulic fracturing.

Berk, J. and P. DeMarzo (2014). *Corporate Finance*. Pearson.

Bernanke, B. (1983). Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics 98*, 85–106.

Bustamante, M. C. and L. Fresard (2017). Does firm investment respond to peers' investment? *Unpublished Working Paper*.

Conley, T. and C. R. Hudry (2010). Learning about a new technology: Pineapple in ghana. *American Economic Review 100*(1), 35–69.

Covert, T. (2015). Experiental and social learning in firms: The case of hydrauling facturing in the bakken shale. *Unpublished Working Paper*.

Dixit, A. K. and R. S. Pindyck (1994). *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press.

Foster, A. and M. Rosenzweig (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. (103), 1176–1209.

Foucault, T. and L. Fresard (2014). Learning from peers' stock prices and corporate investment. *Journal of Financial Economics 111*, 554–577.

Gilje, E. and J. Taillard (2016). Do public firms invest differently than private firms? Taking cues from the natural gas industry. *Journal of Finance 71*, 1733–1778.

Greenan, G. (2018). Dividend payments as a response to peer influence. *Journal of Financial Economics*.

Grenadier, S. R. (1996). The strategic exercise of options: Development cascades and overbuilding in real estate markets. *Journal of Finance 51*, 1653–79.

Grenadier, S. R. (1999). Information revelation through option exercise. *Review of Financial Studies 12*, 95–129.

Grenadier, S. R. (2002). Option exercise games: An application to the equilibrium investment strategies of firms. *Review of Financial Studies 15*, 691–721.

Grenadier, S. R. and A. Malenko (2011). Real options signaling games with applications to corporate finance. *Review of Financial Studies 24*, 3993–4036.

Grenadier, S. R. and N. Wang (2005). Investment timing, agency, and information. *Journal of Financial Economics*.

Grenadier, S. R. and N. Wang (2007). Investment under uncertainty and time-inconsistent preferences. *Journal of Financial Economics 84*, 2–39.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. (25), 501–522.

Hoberg, G. and G. Phillips (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy 124*, 1423–1465.

Kellogg, R. (2014). The effect of uncertainty on investment: Evidence from texas oil drilling. *American Economic Review 104*(6), 1698–1734.

Leary, M. T. and M. R. Roberts (2014). Do peer firms affect corporate capital structure? *Journal of Finance 69*, 139–178.

MacKinnon, J. (2002). Boostrap inference in econometrics. *Canadian Journal of Econometrics* (35), 615–645.

Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Econonomic Studies 60*, 531–542.

Moel, A. and P. Tufano (2002). When are real options exercised? an empirical study of mine closings. *Review of Financial Studies 15*, 35–64.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics 5*, 147–175.

Novy-Marx, R. (2007). An equilibrium model of investment under uncertainty. *Review of Financial Studies 20*, 1461–1502.

Paddock, J. L., D. R. Siegel, and J. L. Smith (1988). Option valuation of claims on real assets: The case of offshore petroleum leases. *The Quarterly Journal of Economics 103*, 479–508.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 999–1033.

Scharfstein, D. and J. Stein (1990). Herd behavior and investment. *American Economic Review 80*, 465–479.

Stoyanov, A. and N. Zubanov (2012). Productivity spillovers across firms through worker mobility. (4), 168–198.

Tchetgen, E. J. T., S. Walter, S. Vansteelandt, T. Martinussen, and M. Glymour (2015). Instrumental variable estimation in a survival context. *Epidemiology* (26), 402–410.

Thompson, R. and P. Thompson (2001). Learning from experience and learning from others: An exploration of learning and spillovers in wartime shipbuilding. (91), 1350–1368.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT Press.

Appendix

To confirm the robustness of the results presented in Section 4, we calibrated a dynamic discrete choice model (see Rust (1987)), in the spirit of Kellogg (2014).[36] In particular, the model presented in this section incorporates more uncertainty by allowing both volatility and well costs to be stochastic. To properly map our institutional setting into Kellogg's model, we had to recalibrate the entire model.

First, we calibrated the nominal discount rate to follow the industry standard of 10%. Then we used the time series for drilling costs and natural gas futures prices during our sample period to estimate both processes. The average dayrate drilling cost is \$49,638.98, with a range between \$25,850 and \$74,671 in our sample. The average natural gas future price is \$4.56, with a range between \$2.56 to \$12.07 during our sample period. We then directly follow Kellogg (2014)'s methodology and calibrate the process for natural gas futures prices such that the expected drift of the natural gas futures price $\mu(P_t, \sigma_t^2)$ is obtained by the OLS regression: $E[ln(P_{t+1}] - ln(P_t) + \frac{\sigma_t^2}{2} = \kappa_{p0} + \kappa_{p1} P_t + \kappa_{p2} \sigma_t^2$. During our sample period, we find $\kappa_{p0} = 0.01526$, $\kappa_{p1} = -0.00016$ and $\kappa_{p2} = 0.343$. Similarly, to calibrate the dayrate process, we estimate the scaling dayrate-to-natural gas futures price parameter (α) and the correlation parameter (ρ) such that $\alpha = 2.2500$ and $\rho = 0.04816$. Our scaling parameter is drastically higher than the one used in Kellogg (2014). The difference is driven by two factors. First, our sample is comprised of horizontal wells, as opposed to vertical wells. The horizontal well technology is much more expensive. Second, Kellogg's focuses on oil wells as opposed to natural gas wells. So his work refers to oil futures price, which are substantially greater than the natural gas futures price we use. Thus, because the standard deviation is not a normalized measure of dispersion, when we constructed the scaling measure, it is normal to obtain substantially greater magnitude for α. Also, the correlation coefficient we obtain is an order of magnitude smaller than the one presented in Kellogg paper. This can be explained by the different periods covered by both our studies. Finally, for the volatility process calibration, we followed the mean reverting calibration methodology and we estimated

[36] Ryan Kellogg's original matlab code is publicly available on the American Economic Review web site. For technical details on the model properties, see Kellogg (2014).

that the volatility of the volatility process to be $\gamma = 0.08852$.

Just as in the standard real options model from Dixit and Pindyck (1994), the dynamic discrete choice model of Kellogg (2014) yields optimal exercise thresholds for different combinations of expected well production and natural gas futures prices, as a function of the volatility of the underlying commodity. We proceed to replicate Figure 6 of Kellogg (2014) in the context of our study, whereby we estimate the optimal exercise threshold curve under two extreme scenarios of uncertainty: the lower and higher uncertainty bounds measured in our sample (19% and 39%). The result of this exercise is shown in Appendix Figure 1A and 1B.

Appendix Figures 1A/B show that in our sample, most wells are exercised in a price range where the Rust-Kellogg model produces optimal threshold values that can hardly be reconciled with the exercised patterns, even assuming the lowest volatility over our sample period (Appendix Figure 1A). That is, the model fails to provide reasonable trigger estimates in the price range we observe, as the trigger curves go "vertical," i.e. head towards infinity just below the \$10 value per mcf in Appendix Figure 1A. Despite this, it is clear from this exercise that relative to the optimal threshold, the conclusion that firms appear to be exercising early is robust across the Dixit-Pindyck framework and Kellogg-Rust framework. To highlight this, we undertake the next exercise, which is to reproduce those curves under the Dixit-Pindyck framework of Section 4.

Specifically, to compare the results obtained using the Dixit-Pindyck model from those obtained by the Rust model, we generate the optimal exercise thresholds under the Dixit and Pindyck (1994) model for those two scenarios also. In Figures 9, 10A and 10B, we evaluated the optimal threshold for each well at the time of exercise, namely with prices and volatility values given at the time of exercise. To map with the Kellogg approach applied in Appendix Figure 1A and 1B, we plot the curves of expected production needed to trigger drilling (y-axis) for the full range of expected futures prices (x-axis), at the two different volatility levels described above (min and max volatility over our sample period) in Appendix Figures 2A and 2B.

Even under the Dixit-Pindyck model, we find that most options (infill wells) are exercised when they are to the left or below the exercise threshold curves; that is, firms tend to exercise earlier than the optimal thresholds would suggest, even in the case where uncertainty is the lowest (Appendix Figure 2A). Hence even when the option to delay has the lowest value, firms seem to forgo the benefits associated with the option to delay drilling. Taken together, these results suggest that firms appear to exercise earlier than what would be deemed optimal by two of the most well-established real options models in the literature. In Section 4.4, we take a first step in reconciling these findings of early exercise with optimal behavior by distinguishing between exercise with and without observed prior adjacent peer activity. We leave to future research the theoretical modeling of incorporating peer activity to the computations of optimal exercise thresholds.

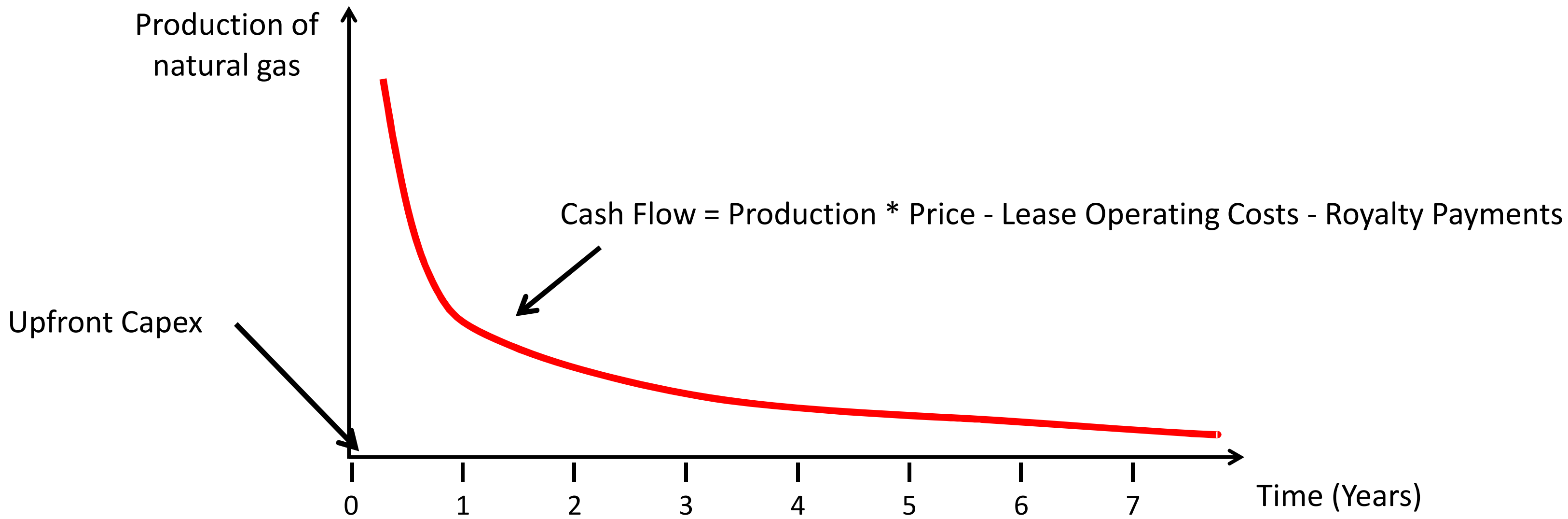


Figure 1: Project Timeline
This figure plots a typical production curve over time for a natural gas well, once production begins. It is based on similar figures found in Lake, Martin, Ramsey, and Titman (2012) as well as company investor presentations.

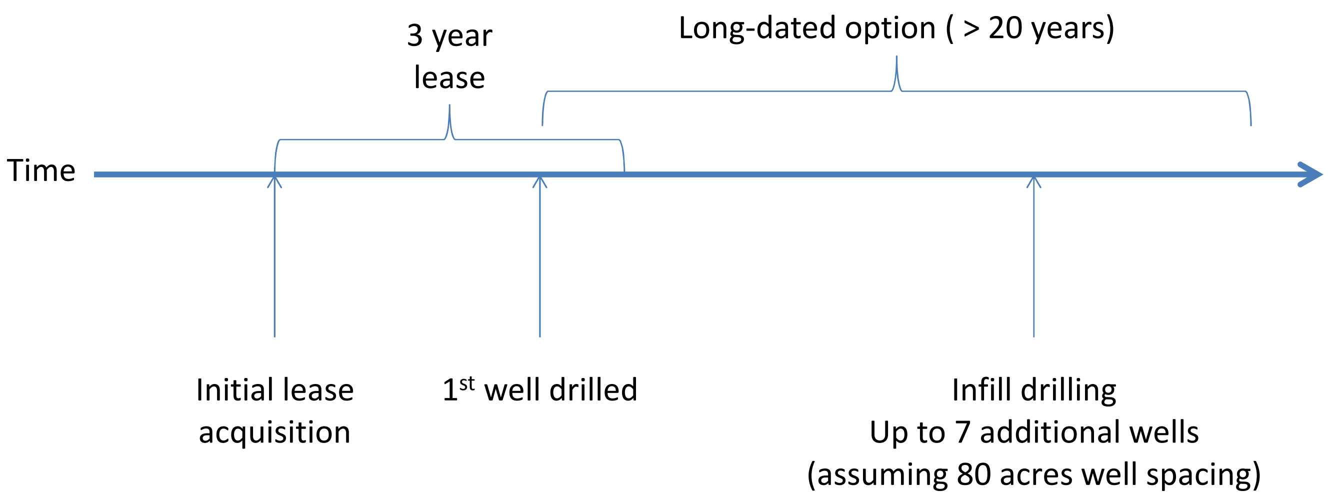


Figure 2: Infill Drilling Option Exercise Timeline
This figure plots the timeline associated with the option to infill drill.

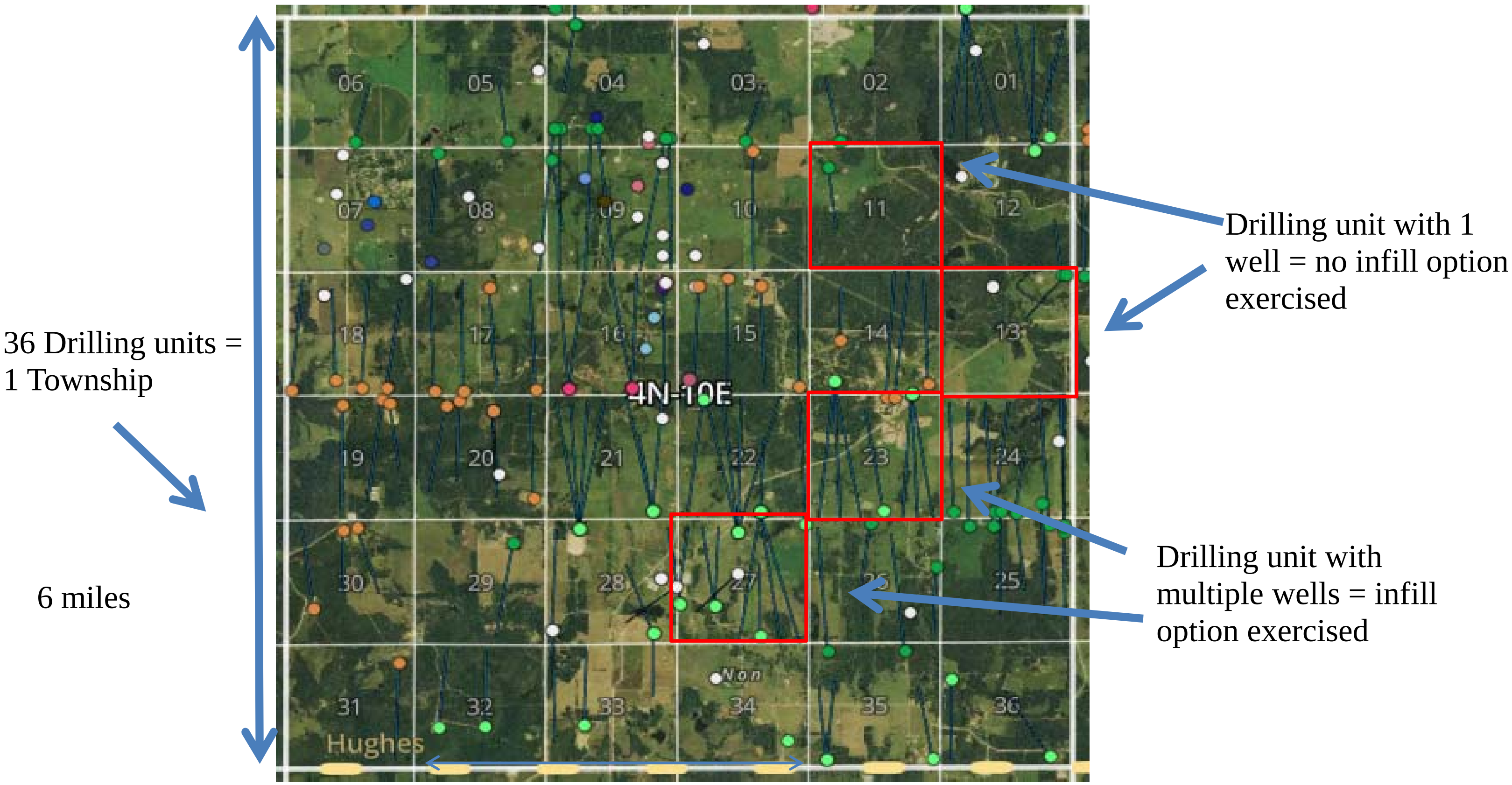


Figure 3: Map of Real Option Exercise Activity

This figure provides a map of drilling activity in one township in the Arkoma Woodford shale. The area covers approximately 36 individual drilling units. The blue lines are the horizontal well-bores of the wells in the drilling units and the multiple horizontal lines in a drilling unit correspond to the real option to "infill" drill having been exercised. In some instances the wellhead (top of the well) may be in a different drilling unit than the horizontal wellbore, in this instance, the well will only drain the reservoir in the drilling unit with the horizontal wellbore. The colors of the wellhead correspond to different companies.

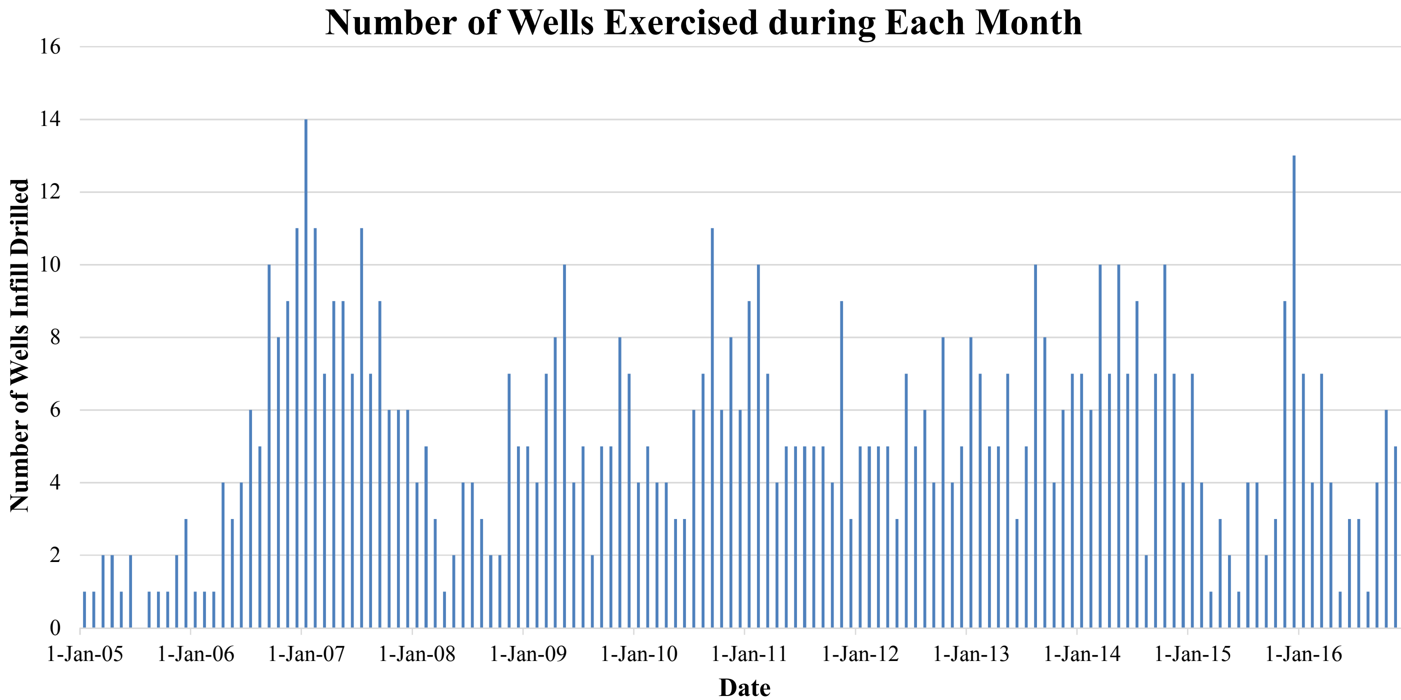


Figure 4A: Number of Infill Well Exercised Over Time

This figure plots the infill drilling exercise activity, measured by the number of infill wells drilled in a given month, for the time period 2005 through 2016.

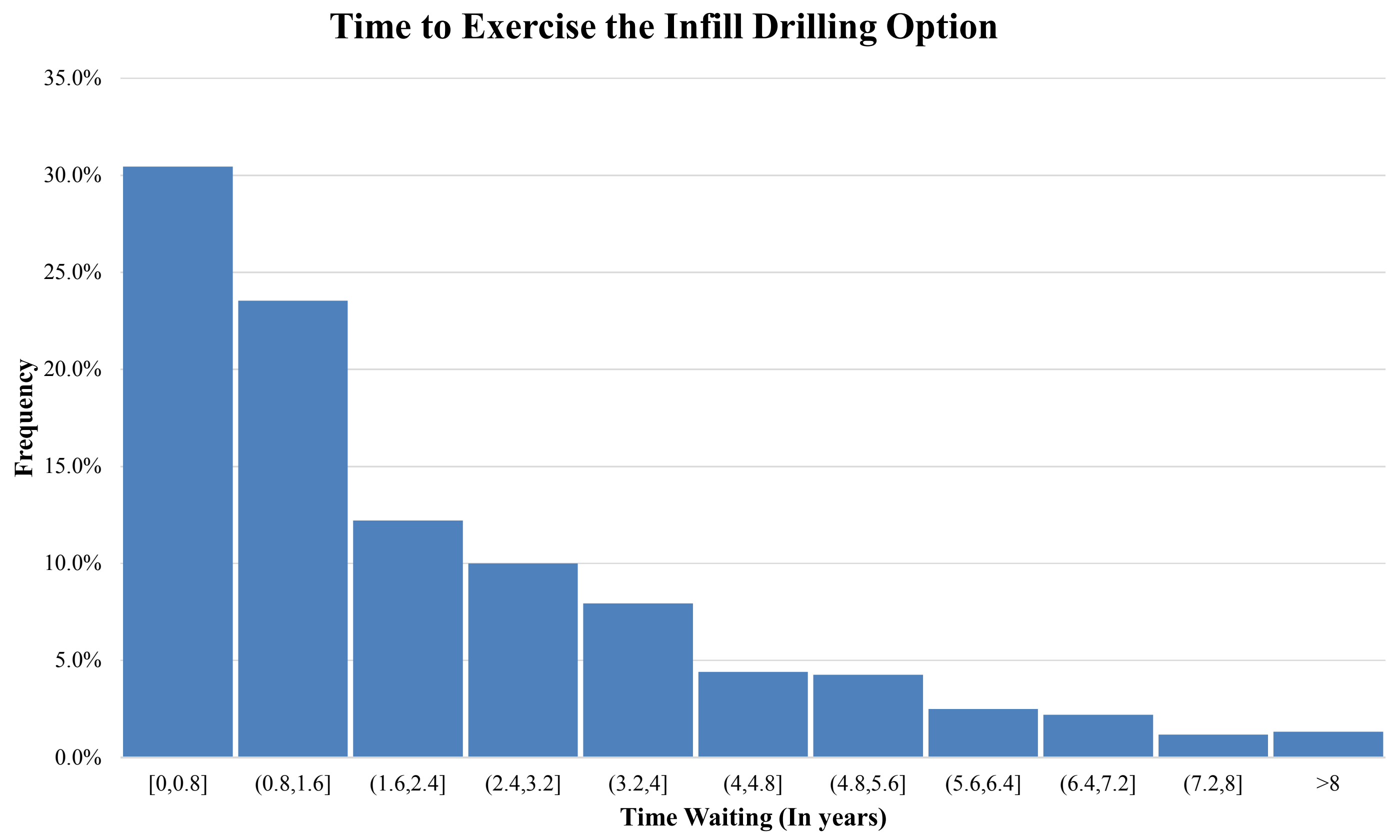


Figure 4B: Time to Exercise the infill Drilling Option

This figure plots the frequency distribution of the time that firms wait before exercising an infill drilling option, for the time period 2005 through 2016.

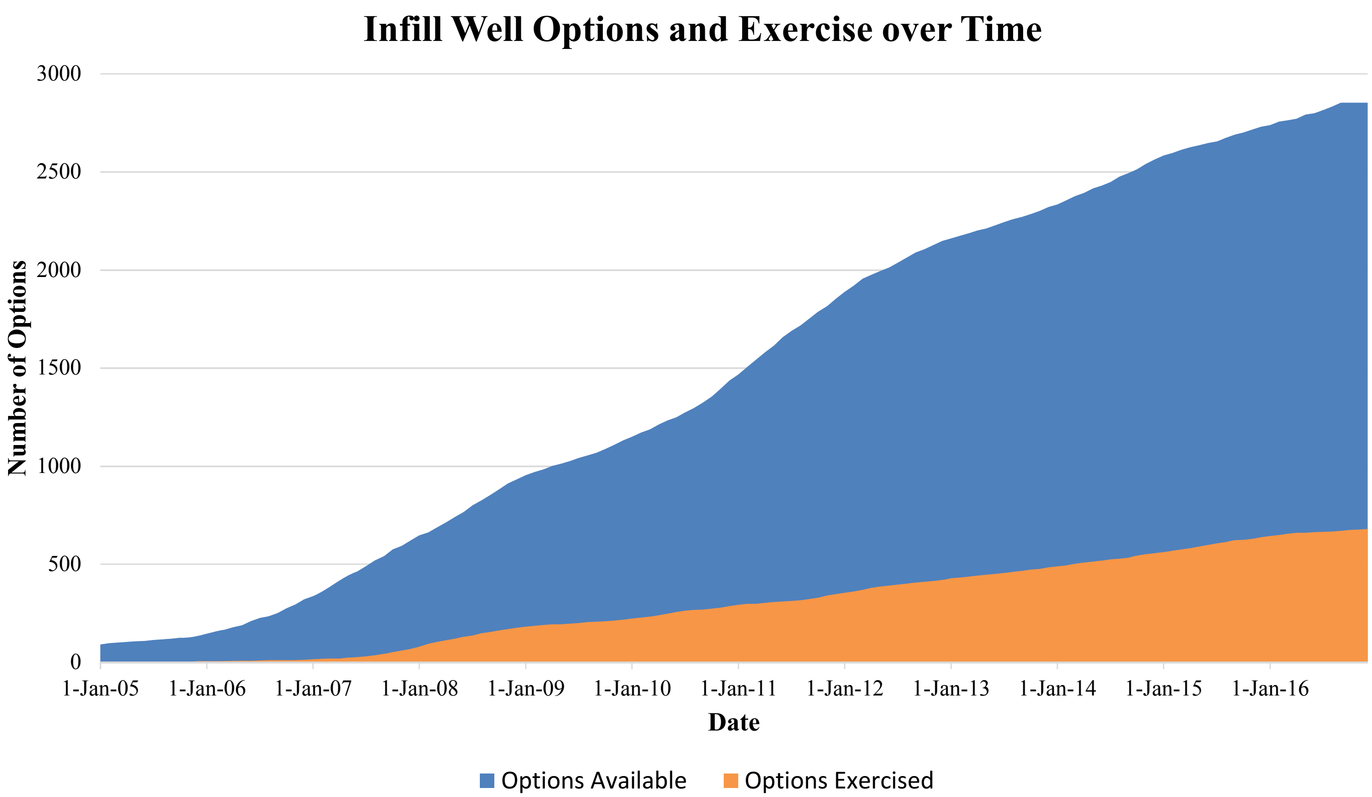


Figure 4C: Infill well options and exercise over time

This figure plots the number of infill drilling options available and the number of options that have been exercised, measured by the number of infill wells drilled, for the time period 2005 through 2016.

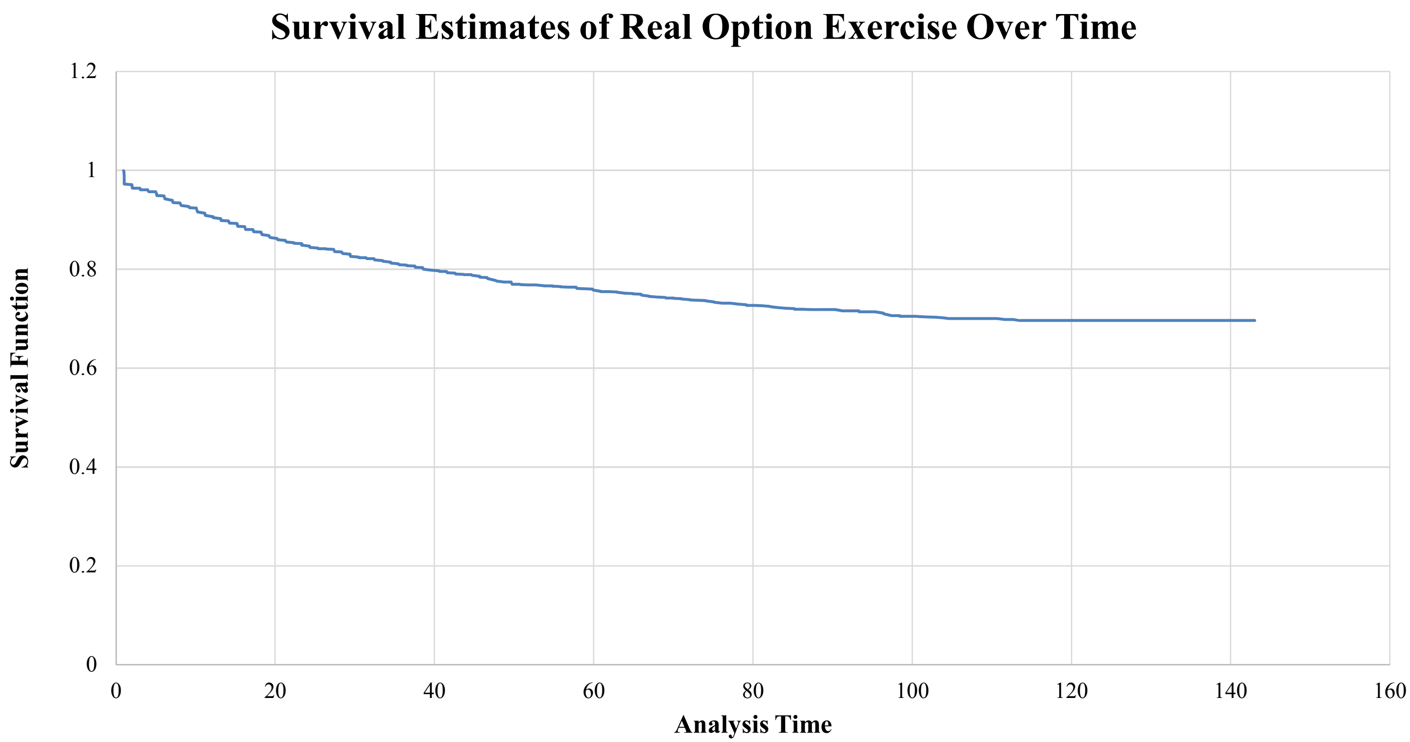


Figure 5: Survival Estimates for the full sample

This figure plots the survival function, measured by the proportion of infill drilling options that remain unexercised (i.e. that have "survived") over our sample period from 2005 through 2016.

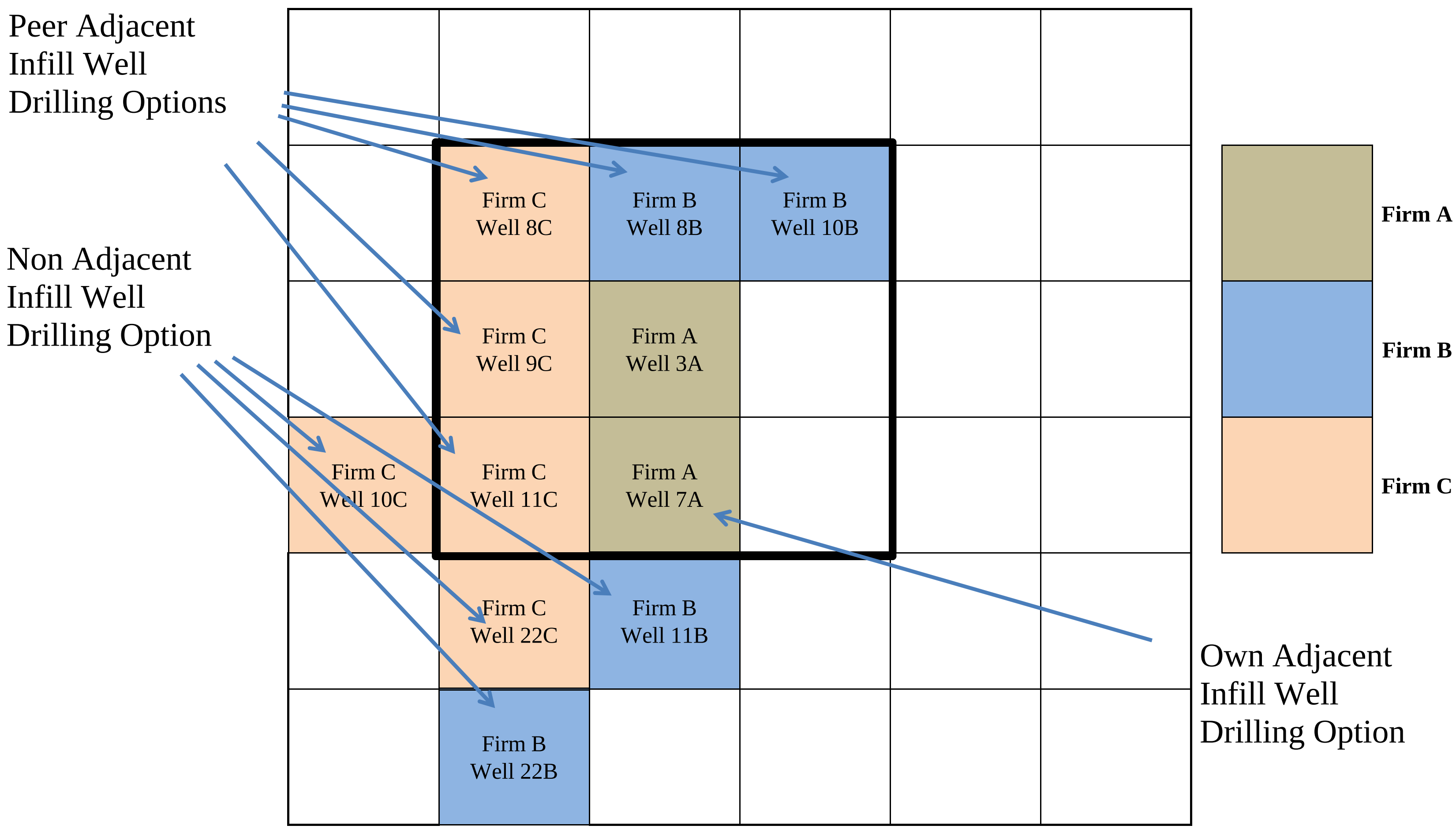
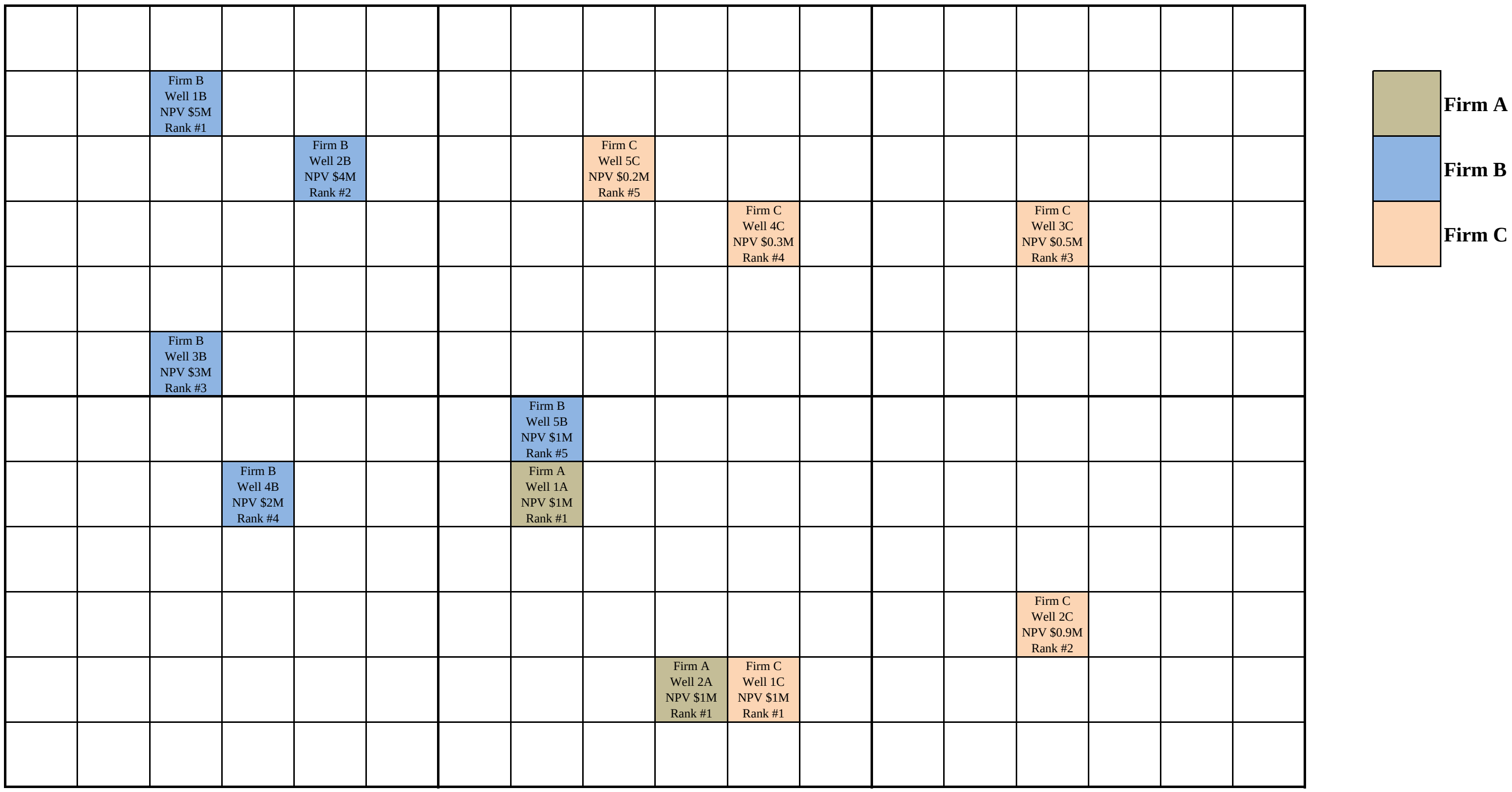


Figure 6: Peer Project Definition

This figure provides an illustrative example of the definition that is used for adjacent peer exercise activity. Specifically, the figure plots a 6 x 6 township which has 36 one mile by one mile drilling units. Due to institutional features of the land survey in our empirical setting, all infill drilling options conform to the above grid layout, and each infill drilling option is linked to a one mile by one mile drilling unit. We compute adjacent activity as the number of adjacent infill options that have been drilled by firms on the 8 adjacent drilling units, we further subdivide this activity by whether peer firms or a firm itself has exercised. For example, for the infill option on well 3A, if Firm C exercised option 8C and 9C and no other options were exercised, the number of adjacent peer options exercised would be 2. If Firm A exercised option 7A, then its own adjacent options exercised would increase to 1, while peer adjacent exercise would remain at 2.



Firm B Infill Real Option Portfolio

Well	Rank	NPV ($M)	Percentile
1B	1	5.0	0.80
2B	2	4.0	0.60
3B	3	3.0	0.40
4B	4	2.0	0.20
5B	5	1.0	0.00

Firm C Infill Real Option Portfolio

Well	Rank	NPV	Percentile
1C	1	1.0	0.80
2C	2	0.9	0.60
3C	3	0.5	0.40
4C	4	0.3	0.20
5C	5	0.2	0.00

Firm A Adjacent Peer Project Percentile Ranks

Well	Adjacent Well	Peer Rank	Peer Rank %	Peer NPV
1A	5B	5	0.00	1.0
2A	1C	1	0.80	1.0

Figure 7: Identification Strategy

This figure provides an illustrative example of the variation we are using for our primary identification strategy. The figure plots different infill well options owned by Firm A, Firm B, and Firm C, along with NPVs and the relative rank of the options in a firm's capital project portfolio. Firm A has two different infill options, one of which is adjacent to Firm B and one of which is adjacent to Firm C. As the example shows, the NPV of Firm A's options and the options adjacent to it are $1 million. However, for infill option 1A the adjacent option owned by firm B (5B) is the fifth ranked option in Firm B's portfolio. Alternatively for infill option 2A, the adjacent option owned by firm C (1C) is the top project in Firm C's portfolio. Our instrument relies on the idea that because option 2A is adjacent to a peer project that is in the 80th percentile of that peer's portfolio (1C) and not the 0th percentile (5B), that 1C is more likely to be exercised, even though the absolute NPVs 1C and 5B are similar.

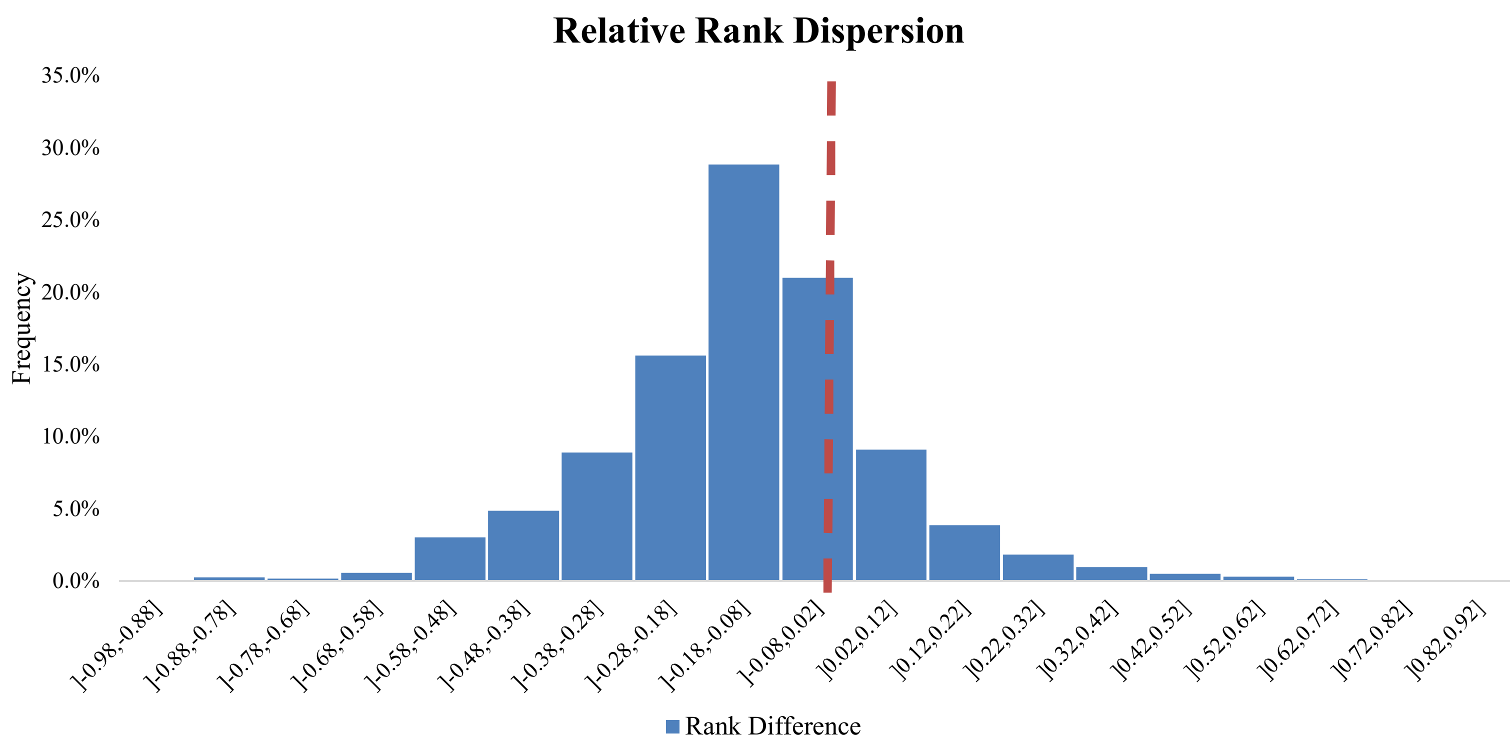


Figure 8: Relative Rank Dispersion

This figure plots the distribution of the difference in the relative rank percentile of the NPV of infill options in a given firm's portfolio versus the relative rank percentile's of the NPV of infill options that are adjacent to the project, but owned by peer firms. Positive numbers mean that the peer relative rank percentile is higher, while negative numbers mean that the peer relative rank percentile is lower. The rank of an infill option's NPV is based on the quality of the initial test well that is drilled to hold the acreage by production (HBP) for the infill option. At every time period all initial wells in a firm's project portfolio are ranked and percentiles are computed. The vertical bar in red marks a rank difference of zero.

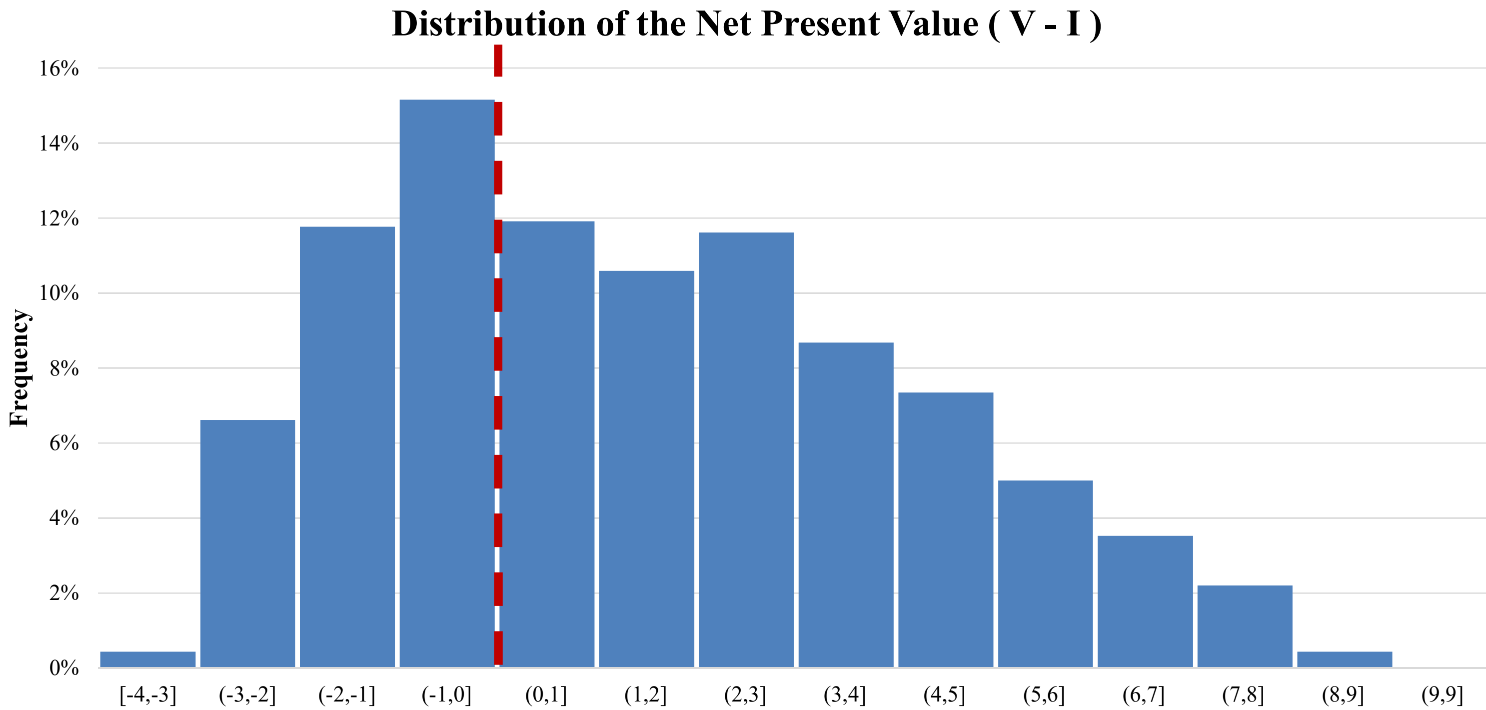


Figure 9: Distribution of the Net Present value (V-I)

This figure plots the distribution of the net present value (V-I) of an infill well at exercise time under the base case scenario, such that we set the depletion rate (ω) at 27%, the accounting depreciation rate (Θ) at 40%, the operational cost (φ) at 20%, the royalty rate (ρ) at 18.75%, the tax rate (τ) at 0% and the discount rate (μ) at 10%. The vertical dotted bar marks an NPV of zero.

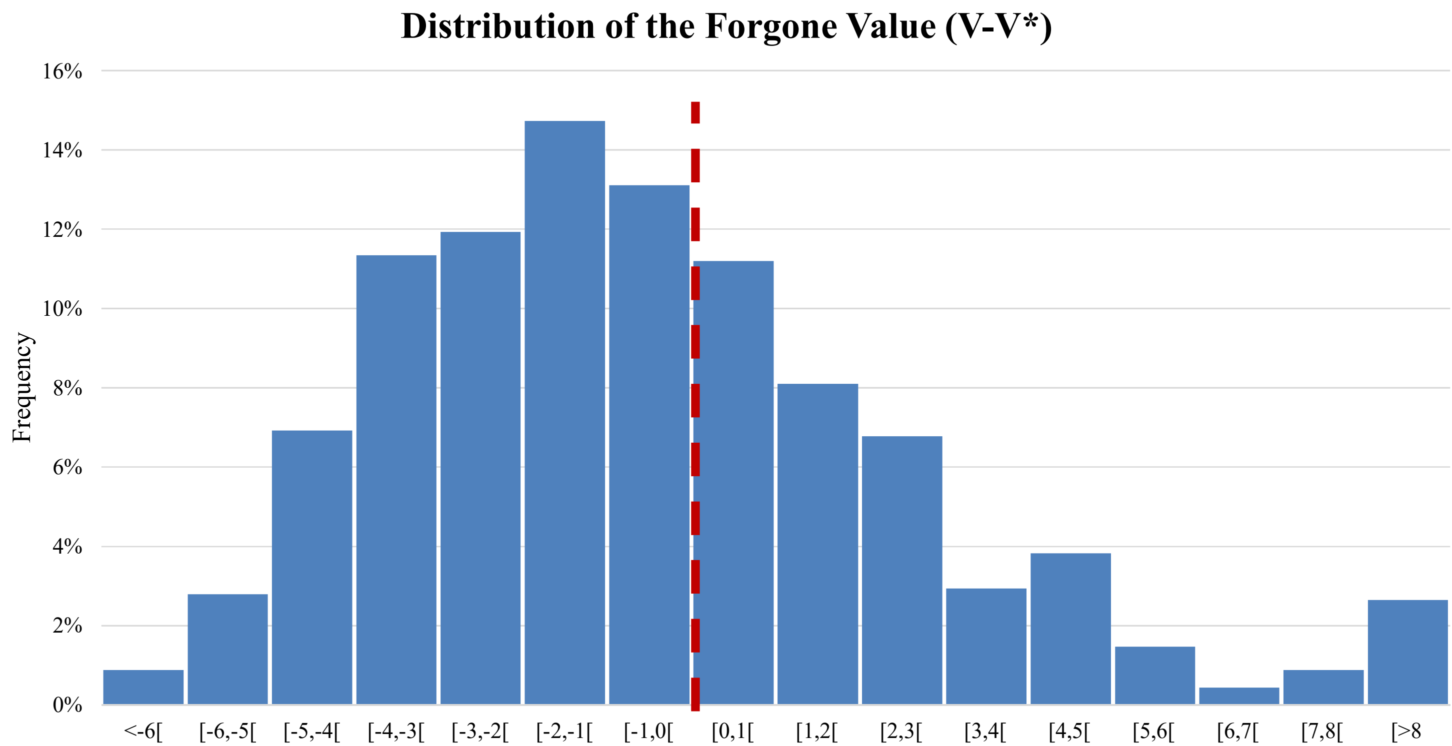


Figure 10: Distribution of the Forgone Value (V-V*)

This figure plots the distribution of the forgone value (V-V*) estimated using the Dixit-Pindyck framework under the base case scenario, such that we set the depletion rate (ω) at 27%, the accounting depreciation rate (Θ) at 40%, the operational cost (φ) at 20%, the royalty rate (ρ) at 18.75%, the tax rate (τ) at 0% and the discount rate (μ) at 10%. The histogram is shown for all exercised real options in the sample. A value to the left of the vertical line at zero corresponds to exercising too early as the trigger rule from optimal stopping time theory prescribes that options are only exercised optimally if exercised the first time the underlying value (V) crosses from below the trigger value (V*).

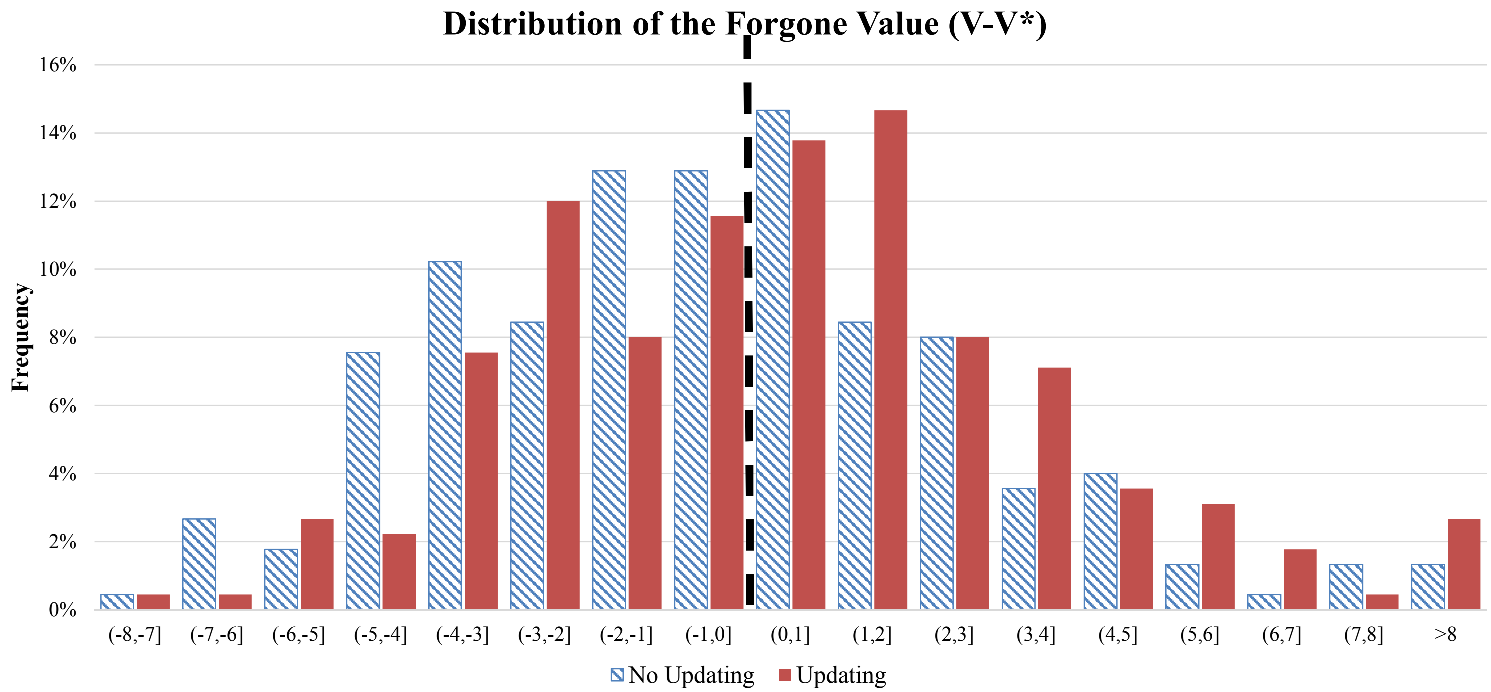


Figure 11A: Distribution of the Forgone Value (V-V*) for Wells with Adjacent Activity

This figure plots the distribution of the forgone value (V-V*) for all the wells with adjacent activity under the base case scenario, such that we set the depletion rate (ω) at 27%, the accounting depreciation rate (Θ) at 40%, the operational cost (ϕ) at 20%, the royalty rate (ρ) at 18.75%, the tax rate (τ) at 0% and the discount rate (μ) at 10%. Two histograms are plotted: the histogram in blue (non-solid) is computed using the wells' expected productivity based on the first well realization, and the histogram in red (solid) augments the information set for expected productivity with the adjacent activity of peers. A value to the left of the vertical line at zero corresponds to exercising too early as the trigger rule from optimal stopping time theory prescribes that options are only exercised optimally if exercised the first time the underlying value (V) crosses from below the trigger value (V*).

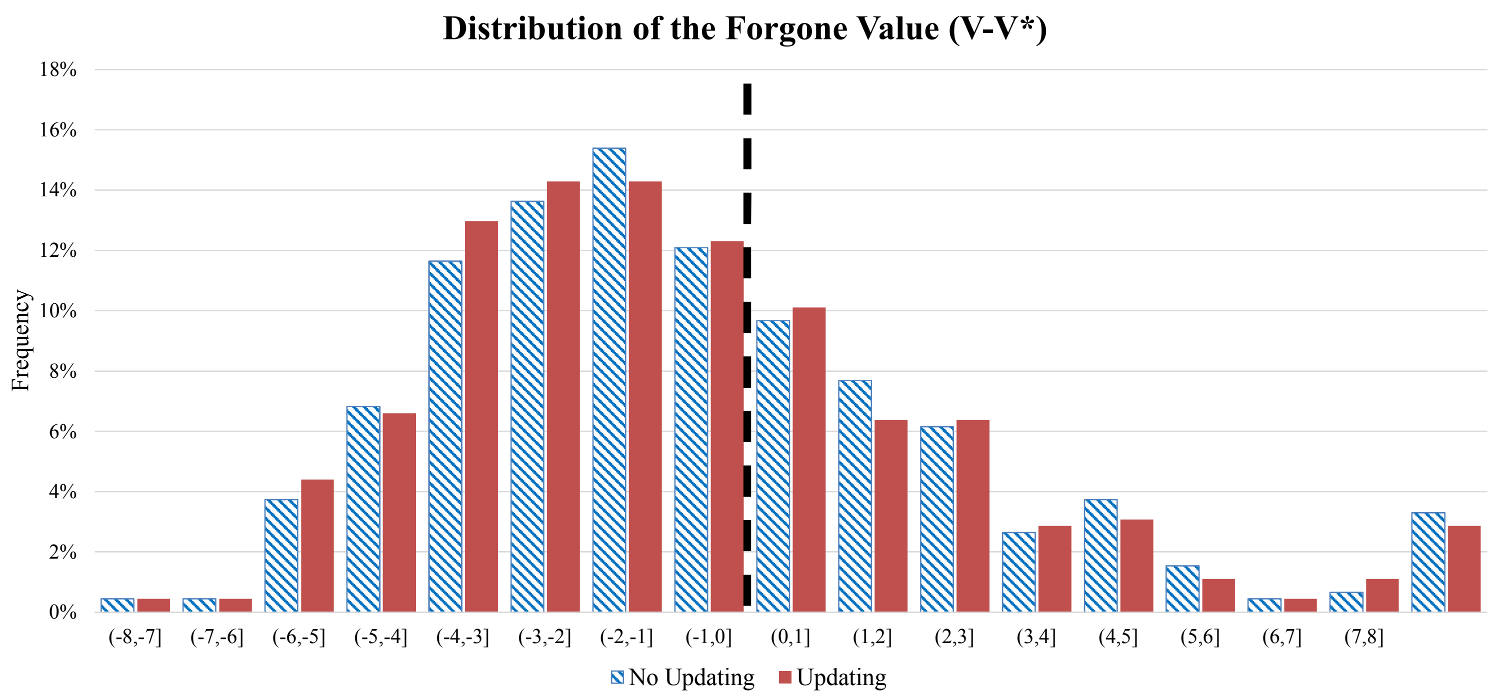


Figure 11B Distribution of the Forgone Value (V-V*) for Wells with no Adjacent Activity

This figure plots the distribution of the forgone value (V-V*) for all the wells with no adjacent activity under the base case scenario, such that we set the depletion rate (ω) at 27%, the accounting depreciation rate (Θ) at 40%, the operational cost (ϕ) at 20%, the royalty rate (ρ) at 18.75%, the tax rate (τ) at 0% and the discount rate (μ) at 10%. Two histograms are plotted: the histogram in blue (non-solid) is computed using the wells' expected productivity based on the first well realization, and the histogram in red (solid) augments the information set for expected productivity with the adjacent activity of peers. A value to the left of the vertical line at zero corresponds to exercising too early as the trigger rule from optimal stopping time theory prescribes that options are only exercised optimally if exercised the first time the underlying value (V) crosses from below the trigger value (V*).

Table 1: Summary Statistics

This table contains summary statistics for the data in our study. Panel A presents an overview of the sample of options on natural gas infill shale drilling opportunities in Oklahoma, including how many real options there are, how many have been exercised, over which sample period, over how many townships and the number of firms (operators) in the sample. Panel B presents summary statistics on the panel data we estimate our hazard models on. The unit of observation in this panel is at the infill option-month level, that is, there is an observation for every infill option available for exercise every month. The baseline variables are all variables used in the hazard model to assess whether exercise is directionally correlated with factors that standard real option theories suggest are important; log(first well production) is a proxy for the underlying reserves in the unit where the infill well can be exercised. The number of adjacent infill options exercised by competitors (peers) corresponds to the variable used to assess whether peer-related real option exercise activity in adjacent drilling units can affect option exercise decisions. We compute a similar measure of adjacent exercise activity for the firm itself (own). The relative rank percentile measures are used to instrument real option exercise activity.

Panel A: Sample Statistics

Time Period	2005-2016
Total Number of Real Options	2853
Number of Exercised Options over Sample Period	680
Number of Townships	442
Number of Firms	159

Panel B: Panel Data Summary Statistics

Baseline Variables	N	Mean	Median	Std Dev
Natural Gas Price	162905	4.56	4.05	1.77
Implied Volatility of Natural Gas	162905	26.58	25.44	4.32
Interest Rates	162905	1.63	1.51	0.85
Log(First Well Production)	162905	12.40	12.68	1.72
Peer Effect Variables	*N*	Mean	Median	Std Dev
Adjacent Competitor Options Exercised	162905	0.34	0.00	0.86
Adjacent Own Firm Options Exercised	162905	0.40	0.00	0.88
Relative rank percentile (own infill option)	103451	0.46	0.45	0.29
Relative rank percentile (adjacent peer infill options)	103451	0.57	0.58	0.29

Table 2: Peer Effects and Real Option Exercise

This table reports coefficient estimates from a Cox hazard model of real option exercise. The time period of the sample is from 2005 to 2016. The unit of observation in the underlying panel is at the "infill drill option" i, month t level. The number of adjacent exercised option (competitor) for an unexercised option i at time t is the number of the adjacent 8 drilling units owned by competitors in which the "infill drill option" has been exercised by time t. The number of "own" adjacent options exercised for an unexercised option i at time t is the number of the adjacent 8 drilling units owned by the firm itself in which the "infill drill option" has been exercised. The implied volatility of natural gas is the implied volatility based on option prices 18 months in the future, and the natural gas price is the price of the natural gas futures contract 18 months out into the future. The five year risk free rate is the 5 year nominal risk free rate on U.S. Treasury bonds. The log of drilling costs is a time-varying estimate of drilling costs for an infill well (analogous to the strike price of the real option). The log first well production variable is fixed for a given option, and is the logarithm of the first year of production of the first well on the drilling unit, which corresponds to production prior to the exercise of the infill option. The following variables have been scaled to have mean 0 and standard deviation of 1, to facilitate economic interpretations: Number of adjacent exercised options (own) and Number of adjacent exercised options (peer). The hazard impact percentage (HI), which is the percentage change in the hazard rate per unit change of the covariate is reported next to the coefficient. z-statistics are reported in brackets below the coefficients. Standard errors are clustered by township. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Hazard model for infill option exercise					
	(1)		(2)		(3)	
	Estimates	HI (%)	Estimates	HI (%)	Estimates	HI (%)
(β_1) Implied volatility of natural gas (percent)$_t$	-0.0328**	-3.23	-0.0337***	-3.31	-0.0282**	-2.79
	[-2.47]		[-2.58]		[-2.22]	
(β_2) Natural gas price ($/mcf)$_t$	0.1378***	14.77	0.1412***	15.17	0.1839***	20.19
	[3.68]		[2.98]		[4.10]	
(β_3) Log drilling cost$_t$			-0.0079	-0.79	0.0667	6.90
			[-0.03]		[0.24]	
(β_4) 5 years risk free interest rate$_t$			0.1382	14.82	0.0749	7.77
			[1.50]		[0.78]	
(β_5) Log first well production$_i$			0.4153***	51.48	0.3302***	39.13
			[5.74]		[3.08]	
(β_6) Number of adjacent exercised options (own)$_{i,t}$	0.546***	72.63	0.5263***	69.26	0.3781***	45.95
	[15.22]		[14.24]		[8.83]	
(β_7) Number of adjacent exercised options (peer)$_{i,t}$	0.3233***	38.17	0.2821***	32.59	0.1038*	10.94
	[8.75]		[7.58]		[1.96]	
Township FE	No		No		Yes	
N	162,905		162,905		162,905	

Table 3: Project Relative Rank Percentile and Option Exercise

This table reports the effect of the relative project rank percentile within the portfolio of a firm's infill drilling options on the decision to exercise the real option to infill drill. The time period of the sample is from 2005 to 2016. The unit of observation in the underlying panel is at the "infill drill option" i, month t level. The relative project rank percentiles are based on the quality of the project, as measured by the production from the first well on a drilling unit within a firm's portfolio. The percentile is computed as the rank of the project divided by the total number of infill options a firm has, higher percentile projects can be viewed as having a higher relative NPV rank within a firm's portfolio. The variable "Relative rank percentile (own project) has been scaled to have mean 0 and standard deviation of 1, to facilitate economic interpretations. The hazard impact percentage (HI), which is the percentage change in the hazard rate per unit change of the covariate is reported next to the coefficient. z-statistics are reported in brackets below the coefficients. Standard errors are clustered by township. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Hazard model for infill option exercise					
	(1)		(2)		(3)	
	Estimate	HI (%)	Estimate	HI (%)	Estimate	HI (%)
(β_1) Implied volatility of natural gas (percent)$_t$	-0.0252*	-2.49	-0.028**	-2.76	-0.0245*	-2.42
	[-1.92]		[-2.07]		[-1.77]	
(β_2) Natural Gas price ($/mcf)$_t$	0.1751***	19.14	0.1692***	18.44	0.1631***	17.71
	[4.21]		[3.28]		[3.38]	
(β_3) Log drilling cost$_t$			0.1772	19.39	-0.0141	-1.40
			[0.62]		[-0.05]	
(β_4) 5 years risk free interest rate$_t$			0.0533	5.47	-0.0093	-0.93
			[0.60]		[-0.10]	
(β_5) Log first well production$_i$			0.1273	13.57	-0.0974	-9.28
			[1.29]		[-1.44]	
(β_6) Relative rank percentile (own project)$_{i,t}$	0.6147***	84.91	0.5059***	65.84	0.6014***	82.47
	[10.24]		[4.70]		[5.69]	
Township FE	No		No		Yes	
N	162,905		162,905		162,905	

Table 4: Real Option Exercise and Exogenous Peer Effects

This table reports the effects of peer real option exercise decisions based on an exogenous measure of peer exercise activity from an instrument. The time period of the sample is from 2005 to 2016. The unit of observation in the underlying panel is at the "infill drill option" i, month t level. The results in Panel A report coefficient estimates of a hazard model which uses the average adjacent peer project relative rank percentile to instrument (Relative rank percentile (adjacent peer projects) in the table) for the number of adjacent peer infill options which have been exercised. The bottom of Panel A reports the first stage regression of the two-stage estimation approach. Correction for estimation error in the first stage in our Cox two-stage IV model is performed by bootstrapping the standard errors (MacKinnon (2002). The Cox IV two-stage approach follows Tchetgen et al. (2015). Panel B reports the direct effect of our instrument in the Cox hazard model. The following variables have been scaled to have mean 0 and standard deviation of 1, to facilitate economic interpretations: Number of adjacent exercised options (own), Number of adjacent exercised options (peer), Relative rank percentile (own project), and Relative rank percentile (adjacent peer projects). The hazard impact percentage (HI), which is the percentage change in the hazard rate per unit change of the covariate is reported next to the coefficient. z-statistics are reported in brackets below the coefficients. Standard errors are clustered by township. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 4: Panel A

	Hazard model					
	Instrumented - Number of adjacent exercised options (peer)					
	(1)		(2)		(3)	
	Estimate	HI (%)	Estimate	HI (%)	Estimate	HI (%)
(β_1) Implied volatility of natural gas (percent)$_t$	-0.0245	-2.42	-0.0281	-2.77	-0.0166	-1.64
	[-1.45]		[-1.62]		[-0.98]	
(β_2) Natural gas price ($/mcf)$_t$	0.2062***	22.90	0.1546***	16.72	0.2801***	32.33
	[2.86]		[2.62]		[2.70]	
(β_3) Log drilling cost$_t$	0.0494	5.06	-0.0319	-3.14	0.4462	56.24
	[0.17]		[-0.10]		[1.07]	
(β_4) 5 year risk free interest rate$_t$	0.1325	14.17	0.1564	16.93	0.2168	24.21
	[1.07]		[1.50]		[1.29]	
(β_5) Log first well production$_i$	0.2432***	27.54	-0.0064	-0.64	-0.0565	-5.50
	[2.61]		[-0.07]		[-0.68]	
(β_6) Instrumented - Number of adjacent exercised options (peer)$_{i,t}$	0.595***	81.31	0.5825***	79.06	0.6623**	93.93
	[2.72]		[2.82]		[2.16]	
(β_7) Average log first well production adjacent options (peer)$_{i,t}$	-0.0671	-6.49	-0.0746**	-7.18	-0.0236	-2.33
	[-1.63]		[-2.15]		[-0.75]	
(β_8) Number of adjacent exercised options (own)$_{i,t}$			0.3731***	45.22	0.7649***	114.88
			[3.35]		[4.25]	
(β_9) Relative rank percentile (own project)$_{i,t}$			0.311**	36.48	0.2563	29.21
			[2.05]		[1.63]	
Township FE	No		No		Yes	
N	103,451		103,451		103,451	

First Stage Coefficients	First Stage Regression Dependent Variable = Number of adjacent exercised options (peer)$_{i,t}$		
(β_1) Relative rank percentile (adjacent peer projects)$_{i,t}$	0.1306***	0.121***	0.1117***
	[3.28]	[3.14]	[2.93]
Township FE	No	No	Yes
Included Instruments/Controls	Yes	Yes	Yes

Table 4: Panel B

	Hazard model					
	Reduced form - Relative rank percentile (adjacent peer projects)					
	(1)		(2)		(3)	
	Estimate	HI (%)	Estimate	HI (%)	Estimate	HI (%)
(β_1) Implied volatility of natural gas (percent)$_t$	-0.0242	-2.39	-0.028*	-2.76	-0.0211	-2.09
	[-1.55]		[-1.84]		[-1.43]	
(β_2) Natural gas price ($/mcf)$_t$	0.1838***	20.18	0.1301***	13.90	0.1588***	17.21
	[3.64]		[2.71]		[3.12]	
(β_3) Log drilling cost$_t$	0.1768	19.34	0.0805	8.39	0.2025	22.45
	[0.55]		[0.26]		[0.55]	
(β_4) 5 year risk free interest rate$_t$	0.0704	7.30	0.1022	10.76	0.0441	4.51
	[0.75]		[1.11]		[0.45]	
(β_5) Log first well production$_i$	0.295***	34.31	0.0045	0.45	-0.0446	-4.36
	[3.40]		[0.06]		[-0.60]	
(β_6) Relative rank percentile (adjacent peer projects)$_{i,t}$	0.3043***	35.57	0.2676***	30.69	0.2417***	27.34
	[3.31]		[3.65]		[2.90]	
(β_7) Average log first well production adjacent options (peer)$_{i,t}$	0.0567***	5.84	0.0365**	3.72	0.0566***	5.82
	[3.46]		[2.31]		[3.06]	
(β_8) Number of adjacent exercised options (own)$_{i,t}$			0.5002***	64.90	0.3985***	48.96
			[11.01]		[7.67]	
(β_9) Relative rank percentile (own project)$_{i,t}$			0.4146***	51.37	0.4256***	53.05
			[3.86]		[3.67]	
Township FE	No		No		Yes	
N	103,451		103,451		103,451	

Table 5: Internal Validity - Correlation of Project Relative Rank Percentiles

This table reports the coefficient estimates of an ordinary least squares (OLS) regression of the relative rank percentile of a firm's own project on the relative rank percentiles of adjacent infill options owned by peer firms. The unit of observation in the underlying panel is at the "infill drill option" *i* , month *t* level. *t* -statistics are reported in brackets below the coefficients. Standard errors are clustered by township. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Dependent variable = Relative rank percentile (own project)
	(1)
(β_1) Relative rank percentile (adjacent peer projects)$_{i,t}$	-0.0359
	[-1.11]
Township FE	Yes
N	103,451

Table 6: Internal Validity - Subsample Analysis

This table reports coefficient estimates from a Cox hazard model of real option exercise on a specific subsample to test instrument validity. The time period of the samples are from 2005 to 2016. The unit of observation in the underlying panel is at the "infill drill option" i, month t level. Specifications (1) and (2) report exercise behavior for the subsample of real options where a project's relative rank percentile within a given firm's portfolio is below the median for that firm, but adjacent projects owned by peers have relative rank percentiles in peer project portfolios that are above median. The following variables have been scaled to have mean 0 and standard deviation of 1, to facilitate economic interpretations: Number of adjacent exercised options (own), Number of adjacent exercised options (peer), Relative rank percentile (own project), and Relative rank percentile (adjacent peer projects). The hazard impact percentage (HI), which is the percentage change in the hazard rate per unit change of the covariate is reported next to the coefficient. z-statistics are reported in brackets below the coefficients. Standard errors are clustered by township. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Hazard model			
	Low rank own project vs. High rank adjacent peer project			
	Reduced Form Peer Effects		Instrumented Peer Effects	
	(1)		(2)	
	Estimates	HI (%)	Estimates	HI (%)
(β_1) Implied volatility of natural gas (percent)$_t$	-0.0107	-1.06	-0.0157	-1.56
	[-0.38]		[-0.56]	
(β_2) Natural gas price ($/mcf)$_t$	0.0584	6.02	0.1579**	17.11
	[0.66]		[2.03]	
(β_3) Log drilling cost$_t$	-0.0701	-6.77	-0.0328	-3.23
	[-0.15]		[-0.06]	
(β_4) 5 year risk free interest rate$_t$	0.1505	16.25	0.1107	11.71
	[1.01]		[0.78]	
(β_5) Log first well production$_i$	0.0241	2.44	0.0308	3.12
	[0.27]		[0.29]	
(β_6) Relative rank percentile (adjacent peer projects)$_{i,t}$	0.4345***	54.42		
	[3.46]			
(β_7) Average log first well production adjacent options (peer)$_{i,t}$	0.0793***	8.26	-0.0936	-8.93
	[3.02]		[-1.17]	
(β_8) Number of adjacent exercised options (own)$_{i,t}$	0.399***	49.03	0.0166	1.67
	[5.27]		[0.04]	
(β_9) Relative rank percentile (own project)$_{i,t}$	0.1793	19.64	-0.2562	-22.60
	[0.83]		[-0.85]	
(β_{10}) Instrumented - Number of adjacent exercised options (peer)$_{i,t}$			1.1038**	201.57
			[2.21]	
Township FE	No		No	
N	43,686		43,686	

Table 7: Actions vs. Characteristics

This table reports coefficient estimates from a Cox hazard model of real option exercise. The time period of the sample is from 2005 to 2016. The unit of observation in the underlying panel is at the "infill drill option" i, month t level. The spell in the hazard model is defined as the time period from which an infill option becomes available (first well drilled) to when the infill option (two or more wells) are drilled. The variable "Play activity" is computed in two steps. First, we compute the drilling activity intensity across Oklahoma for each adjacent competitor at a given point in time t, while excluding the wells they drilled in the township of infill option i. Then, for each infill drill option in our sample, we define the variable "Play activity" as the sum of the adjacent competitor activity outside the township of the infill drill option at each point in time. The following variables have been scaled to have mean 0 and standard deviation of 1, to facilitate economic interpretations: Number of adjacent exercised options (own), Number of adjacent exercised options (peer), Relative rank percentile (own project), Play activity (peer), and Relative rank percentile (adjacent peer projects). The hazard impact percentage (HI), which is the percentage change in the hazard rate per unit change of the covariate is reported next to the coefficient. z-statistics are reported in brackets below the coefficients. Standard errors are clustered by township. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Hazard model							
	Reduced Form Peer Effects				Instrumented Peer Effects			
	(1)		(2)		(3)		(4)	
	Estimates	HI (%)	Estimates	HI (%)	Estimates	HI (%)	Estimates	HI (%)
(β_1) Implied volatility of natural gas (percent)$_t$	-0.0272*	-2.69	-0.0203	-2.01	-0.0289*	-2.85	-0.0168	-1.67
	[-1.80]		[-1.38]		[-1.67]		[-0.93]	
(β_2) Natural gas price ($/mcf)$_t$	0.1205**	12.81	0.1548***	16.75	0.1665***	18.12	0.2767***	31.88
	[2.54]		[3.08]		[2.81]		[2.96]	
(β_3) Log drilling cost$_t$	0.0536	5.51	0.2035	22.57	0.0145	1.46	0.4603	58.45
	[0.17]		[0.55]		[0.04]		[1.30]	
(β_4) 5 year risk free interest rate$_t$	0.1052	11.09	0.0501	5.13	0.1471	15.85	0.2193	24.52
	[1.13]		[0.51]		[1.39]		[1.08]	
(β_5) Log first well production$_i$	0.0048	0.48	-0.0413	-4.05	-0.0033	-0.33	-0.0564	-5.49
	[0.06]		[-0.55]		[-0.04]		[-0.56]	
(β_6) Relative rank percentile (adjacent peer projects)$_{i,t}$	0.2689***	30.85	0.2448***	27.74				
	[3.65]		[2.94]					
(β_7) Average log first well production adjacent options (peer)$_{i,t}$	0.0307**	3.12	0.0535***	5.50	-0.0665**	-6.43	-0.0200	-1.98
	[1.96]		[2.91]		[-2.17]		[-0.57]	
(β_8) Number of adjacent exercised options (own)$_{i,t}$	0.5106***	66.62	0.4066***	50.18	0.355***	42.62	0.7553***	112.83
	[11.11]		[7.93]		[3.01]		[4.07]	
(β_9) Relative rank percentile (own project)$_{i,t}$	0.4107***	50.79	0.4207***	52.30	0.3062**	35.82	0.2527	28.75
	[3.84]		[3.63]		[2.02]		[1.26]	
(β_{10}) Play activity (peer)$_{i,t}$	0.0916***	9.60	0.0575**	5.92	-0.2544**	-22.46	-0.1165	-11.00
	[4.24]		[2.26]		[-1.96]		[-1.00]	
(β_{11}) Instrumented - Number of adjacent exercised options (peer)$_{i,t}$					0.7351***	108.58	0.6926*	99.89
					[2.79]		[1.76]	
Township FE	No		Yes		No		Yes	
N	103,451		103,451		103,451		103,451	

Table 8: Real Option Exercise and Experienced Peers

This table reports coefficient estimates from a Cox hazard model of real option exercise. The time period of the sample is from 2005 to 2016. The unit of observation in the underlying panel is at the "infill drill option" i, month t level. The signal quality variables (Adjacent Experienced / Adjacent Inexperienced) are constructed in two steps. First, we identify if the adjacent firms exercising their drilling option are more (less) experienced than the median firm in the sample based on the number of wells drilled and, accordingly, we define them as Experienced (Inexperienced). In the second step, we aggregate the wells that are drilled by experienced firms into the variable number of adjacent experienced options (experienced peer) and those drilled by unexperienced firms into the number adjacent exercised options (inexperienced peer). The following variables have been scaled to have mean 0 and standard deviation of 1, to facilitate economic interpretations: Number of adjacent exercised options (own), Number of adjacent exercised options (experienced peer), Number of adjacent exercised options (inexperienced peer). The hazard impact percentage (HI), which is the percentage change in the hazard rate per unit change of the covariate is reported next to the coefficient. z-statistics are reported in brackets below the coefficients. Standard errors are clustered by township. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Hazard Model			
	(1)		(2)	
	Estimates	HI (%)	Estimates	HI (%)
(β_1) Implied volatility of natural gas (percent)$_t$	-0.0339***	-3.33	-0.0282**	-2.78
	[-2.59]		[-2.22]	
(β_2) Natural gas price ($/mcf)$_t$	0.1407***	15.10	0.1842***	20.23
	[2.96]		[4.09]	
(β_3) Log drilling cost$_t$	0.0009	0.09	0.0669	6.91
	[0.00]		[0.24]	
(β_4) 5 year risk free interest rate$_t$	0.1398	15.00	0.0746	7.74
	[1.51]		[0.77]	
(β_5) Log first well production$_i$	0.4144***	51.35	0.3309***	39.22
	[5.72]		[3.08]	
(β_6) Number of adjacent exercised options (own)$_{i,t}$	0.525***	69.04	0.3786***	46.03
	[14.32]		[8.89]	
(β_7) Number of adjacent exercised options (experienced peer)$_{i,t}$	0.2658***	30.45	0.1022**	10.76
	[7.54]		[2.14]	
(β_8) Number of adjacent exercised options (inexperienced peer)$_{i,t}$	0.0871***	9.10	0.0152	1.53
	[3.03]		[0.30]	
Township FE	No		Yes	
N	162,905		162,905	

Table 9: Real Option Value Estimates and Sensitivity Analysis

This table first reports in Panel A summary statistics on well costs (I), present value of cash flows (V), the optimal trigger value (V*), and the net present value (NPV=V-I) at the time of exercise, as generated by a baseline real options model (see Paddock et al. (1988) and Dixit Pindyck (1994)). Panel B reports a sensitivity analysis for the Net Present Value (NPV=V-I) and forgone value (V*-V) at time of exercise. The sensitivity analysis is performed on the different assumptions for several model parameters. Depletion rate is the rate at which a well depletes its reserves, a rate of 27 can be interpreted as a wells production declining at a rate of 27% a year. Operational Cost is the percentage of cash flows going towards lease operating expenses, a rate of 20 can be interpreted as 20% of cash flows going to pay for ongoing operating costs. The discount rate is the firm's cost of capital, and the tax rate is the rate used to compute after tax cash flow.

Panel A: Summary statistics

Well-Level Statistics at Time of Exercise	N	Mean	Median	Std Dev
Well Costs (I)	680	$ 4,740,347	$ 4,798,365	$ 651,954
Present Value of Well Cash Flow (V)	680	$ 6,656,654	$ 6,146,015	$ 3,390,738
Optimal Trigger Value (V*)	680	$ 7,079,647	$ 7,307,629	$ 1,407,007
Net Present Value (V-I)	680	$ 1,916,307	$ 1,374,095	$ 3,414,682

Panel B: Sensitivity Analysis

Depletion Rate Sensitivity	Mean	Pr(Mean = 0)	Median	Pr(Median = 0)
Net Present Value (V-I at Exercise)				
Depletion Rate (ω = 25)	$2,101,214	0.00	$1,540,539	0.00
Depletion Rate (ω = 27)	$1,916,307	0.00	$1,374,095	0.00
Depletion Rate (ω = 29)	$1,574,940	0.00	$1,078,187	0.00
Forgone Value (V-V at Exercise)*				
Depletion Rate (ω = 25)	$238,086	0.07	$702,787	0.00
Depletion Rate (ω = 27)	$422,993	0.00	$866,943	0.00
Depletion Rate (ω = 29)	$764,360	0.00	$1,144,138	0.00
Operational Cost Sensitivity	Mean	Pr(Mean = 0)	Median	Pr(Median = 0)
Net Present Value (V-I at Exercise)				
Operational Cost (ϕ = 15)	$2,459,708	0.00	$1,864,696	0.00
Operational Cost (ϕ = 20)	$1,916,307	0.00	$1,374,095	0.00
Operational Cost (ϕ = 25)	$1,372,907	0.00	$903,057	0.00
Forgone Value (V-V at Exercise)*				
Operational Cost (ϕ = 15)	$259,973	0.05	$721,624	0.00
Operational Cost (ϕ = 20)	$422,993	0.00	$866,943	0.00
Operational Cost (ϕ = 25)	$694,693	0.00	$1,078,926	0.00
Discount Rate Sensitivity	Mean	Pr(Mean = 0)	Median	Pr(Median = 0)
Net Present Value (V-I at Exercise)				
Discount Rate (μ = 7.5%)	$2,398,674	0.00	$1,811,878	0.00
Discount Rate (μ = 10%)	$1,916,307	0.00	$1,374,095	0.00
Discount Rate (μ = 12.5%)	$1,495,000	0.00	$1,008,892	0.00
Forgone Value (V-V at Exercise)*				
Discount Rate (μ = 7.5%)	$857,331	0.00	$1,346,908	0.00
Discount Rate (μ = 10%)	$422,993	0.00	$866,943	0.00
Discount Rate (μ = 12.5%)	$318,536	0.01	$760,092	0.00
Tax Rate Sensitivity	Mean	Pr(Mean = 0)	Median	Pr(Median = 0)
Net Present Value (V-I at Exercise)				
Tax Rate (τ = 0 %)	$1,916,307	0.00	$1,374,095	0.00
Tax Rate (τ = 15 %)	$1,569,890	0.00	$1,073,808	0.00
Tax Rate (τ = 30 %)	$1,223,472	0.00	$772,849	0.00
Forgone Value (V-V at Exercise)*				
Tax Rate (τ = 0 %)	$422,993	0.00	$866,943	0.00
Tax Rate (τ = 15 %)	$769,411	0.00	$1,148,941	0.00
Tax Rate (τ = 30 %)	$1,115,829	0.00	$1,470,295	0.00
Royalty Rate Sensitivity	Mean	Pr(Mean = 0)	Median	Pr(Median = 0)
Net Present Value (V-I at Exercise)				
Royalty Rate (ρ = 13.75%)	$2,459,708	0.00	$1,864,696	0.00
Royalty Rate (ρ = 18.75%)	$1,916,307	0.00	$1,374,095	0.00
Royalty Rate (ρ = 23.75%)	$1,372,907	0.00	$903,057	0.00
Forgone Value (V-V at Exercise)*				
Royalty Rate (ρ = 13.75%)	$259,973	0.05	$721,624	0.00
Royalty Rate (ρ = 18.75%)	$422,993	0.00	$866,943	0.00
Royalty Rate (ρ = 23.75%)	$694,693	0.00	$1,078,926	0.00