Information Asymmetry, Trade, and Drilling: Evidence from an Oil Lease Lottery

Paul A. Brehm* and Eric Lewis*

ONLINE APPENDIX

A Online Appendix: Theory

This section discusses some alternative models:

- A.1: Modifications of the informed buyer model in Section 3. These include multiple buyers and situations where the seller has more bargaining power. Solutions indicate that the presence of at least some buyer bargaining power is important for rationalizing our findings.

- A.2: We discuss a range of mechanisms that could cause firms to delay trade relative to individuals, including multiple offers that do not arrive simultaneously, risk aversion, and price volatility.

- A.3: A more detailed exploration of one mechanism – liquidity constraints – that could affect the timing of trade. In the absence of information asymmetries, liquidity constraints have no effect on drilling outcomes. When coupled with information asymmetry caused by an informed buyer, leases won by firms are less likely to be traded, have an overall later trading date, and have lower drilling rates than those won by individuals.

*pbrehm@oberlin.edu (Oberlin College) and ericlewis@tamu.edu (Texas A&M University).

A.1 Further discussion of informed buyer model

We discuss two adjustments to the informed buyer in Section 3. First, we consider the case where there are multiple informed buyers. We then examine the model as a seller-offer game.

Multiple buyers

We explore what happens if there are multiple informed buyers who compete with each other for a neighboring lease: Suppose that there are two potential buyers, each with identical perfect knowledge of the true value of θ as well as the same cost C_B, where because of the buyers' previous drilling, $C_B < C_j$. As in the main article text, suppose that the buyers are the ones that make the offers. Here the two potential buyers are competing against each other. Assuming Bertrand competition, the two potential buyers will compete their profits down to zero, each making the offer $O_j^* = \theta - C_B$ as long as $\theta - C_B \geq 0$.[1] Because this offer gives all of the gains from trade to the seller, and because the gains from trade are positive ($C_B < C_j$), then conditional on receiving an offer, the seller always prefers to sell rather than reject the offer. If $\theta - C_B < 0$, then the buyers will not make offers and the seller will infer that $E(\theta | \text{no offer}) < C_B$. Because $C_B < C_j$, the fact that the buyer was unwilling to drill leads the winner to infer that it would be even more unprofitable if they were to drill. Therefore, no drilling occurs.

In this model there is no difference in the probability of trade nor the probability of drilling between leases won by individuals and leases won by firms. Leases are always re-allocated to the buyer if the buyer's profits from drilling are positive. Drilling only happens if a buyer buys. The competition between the two potential buyers eliminates the inefficiencies caused by their private information. This model

[1] The results are similar if the potential buyers have different costs. With different buyer costs, the two buyers will compete profits down to the higher cost buyer's break-even point.

therefore suggests that in order to rationalize our findings that leases won by individuals have a higher probability of drilling than those won by firms, the buyer must have at least some market power.

Seller bargaining power

We now examine the same informed buyer model as in Section 3, only we assume instead that it is the seller that makes a TIOLI offer to the buyer. As in the buyer-offer model, this seller-offer model allows the seller to update beliefs depending on the buyer: In particular, if the buyer refuses the seller's offered asking price, the seller updates beliefs about the value of the lease and may still drill if the updated value of the lease exceeds drilling costs.

In the following, we discuss first the general form of this model. After that we discuss a simple example using the uniform distribution. Under this distributional assumption, we show that leases won by individuals will have a weakly higher probability of trade and a weakly lower probability of drilling. This therefore implies that the buyer likely needs to have some bargaining power in order to rationalize our empirical findings that for leases close to existing production, those won by individuals are more likely to be drilled than leases won by firms.

General form of seller-offer model: As before, the value of θ is taken from a distribution with cdf $F(\theta)$. The seller makes an offer (an "ask") – the amount that it asks from the seller for the lease, which we denote as a. The buyer only accepts the offer if $\theta \geq a + C_B$, which happens with probability $1 - F(a + C_B)$. If the buyer does not accept the offer, then the seller will drill on their own if $E(\theta|\theta < a+C_B) - C_j \geq 0$. The seller can also choose to drill or abandon the lease without making an offer.

Therefore, the offer a that the seller will make will maximize:

$$a[1 - F(a + C_B)] + F(a + C_B) \max\{E(\theta|\theta < a + C_B) - C_j, 0\} \tag{14}$$

subject to:

$$a \geq \max\{E(\theta) - C_j, 0\} \tag{15}$$

In other words, the offer a maximizes the offer multiplied by the probability that the offer is accepted plus the probability that the offer is not accepted multiplied by the updated beliefs about profits from the well if the offer is not accepted. The constraint shows that the offer must be at least as good as the expected value of the lease to a seller if it chose not to make an offer.

This type of model is significantly more complicated to solve than the buyer-offer model because there are multiple cases that must be checked. These include the max function in the optimization function, the max function in the constraint, as well as whether the constraint binds. Because of this complexity, we examine the special case of the uniform distribution. There, we show that the max function in the objective function is analytically important and that the constraint does not bind.

Seller offer model with a uniform distribution: Here we focus on the case where $\theta \sim U(0, 1)$. For demonstration and ease of calculation, we assume that the buyer's cost of drilling is $C_B = 0$, but results are qualitatively similar for other values of C_B. We also assume that $C_j \in (0, 1)$, meaning that there is some probability that drilling the lease on its own would lead to positive profits. Under this framework, we can write the seller's problem from Equations 14 and 15 as:

$$a(1 - a) + a \max\{a/2 - C_j, 0\} \tag{16}$$

subject to:

$$a \geq \max\{1/2 - C_j, 0\} \tag{17}$$

Passing the other terms in Equation 16 into the max function, we can rewrite the problem as choosing the offer a that maximizes:

$$\max\{V_1(a; C_j), V_2(a)\} \tag{18}$$

subject to the same constraint in Equation 17, where V_1 and V_2 are defined as:

$$V_1(a; C_j) = a(1-a) + a(a/2 - C_j) \tag{19}$$

$$V_2(a) = a(1-a) \tag{20}$$

Both V_1 and V_2 are parabolas with global maximum. Denoting the arg max of V_1 and arg max of V_2 respectively as $a_1^*(C_j)$ and a_2^*, we can solve for these by taking first order conditions:

$$a_1^*(C_j) = 1 - C_j \tag{21}$$

$$a_2^* = 1/2 \tag{22}$$

In other words, if the seller plans to drill the lease in the event of the offer not being accepted, it will make the offer $1 - C_j$. In this case, its optimal action and payoffs depend on its own costs because drilling the well in the event of no sale requires paying drilling costs. In contrast, if the seller does not plan to drill the lease in the event of the offer not being accepted, it will make the offer of $1/2$, and the offer will not depend on the seller's drilling costs.

This is illustrated graphically in Figure A.2 panel (a). There we plot V_1 and V_2 for various values of a. There we show how the shape and position of V_1 changes depending on the seller's drilling cost C_j, where a lower seller drilling cost C_j leads to both a higher overall curve V_1 as well as a higher a_1^*.

The optimal offer $a^*(C_j)$ will depend on whether $V_1(a_1^*(C_j))$ or $V_2(a_2^*)$ is higher. In other words, the correspondence $a^*(C_j)$ is defined below in Equation 23 and is graphed in Figure A.2 panel (b):

$$a^*(C_j) = \begin{cases} 1 - C_j & \text{if } V_1(1 - C_j; C_j) \geq V_2(1/2) \quad \text{(23a)} \\ 1/2 & \text{if } V_1(1 - C_j; C_j) \leq V_2(1/2) \quad \text{(23b)} \end{cases}$$

The solution for a^* given in Equation 23 is solved for ignoring the constraint in Equation 17. This constraint will not be binding: Regardless of the value of C_j, we have that $a^*(C_j) > \max\{1/2 - C_j, 0\}$ for all values of $C_j \in (0, 1)$.

A little algebra shows that the value of C_j for which $V_1(1 - C_j; C_j) = V_2(1/2)$ is $C_j = (2 - \sqrt{2})/2$, or about 0.293. At this cost C_j, the ask correspondence $a^*(C_j)$ exhibits a discontinuity, where the seller is indifferent between asking for $1 - C_j$ and asking for 1/2. Both give identical payoffs. Asking for $1 - C_j$ gives a low probability of trade, but this is balanced out by large payoffs conditional on trade as well as the promise of expected positive payoffs to drilling if trade does not happen. In contrast, asking for $a = 1/2$ yields much lower payoffs conditional on trade but a much higher probability of trade. Both strategies give expected profits of 1/4.

In this model, the correspondence $a^*(C_j)$ is weakly decreasing in C_j. This implies that individuals will have weakly higher probability of trade than firms. The figure also shows that probability of drilling will be weakly decreasing in C_j: For $C_j < (2 - \sqrt{2})/2$, drilling will happen regardless of trade. For $C_j > (2 - \sqrt{2})/2$, drilling only happens if there is trade, which happens with probability 1/2.

A.2 Trade Delay models

We next explore a variety of mechanisms whereby initial assignment affects the timing of trade. This section shows that a range of mechanisms can lead to our findings that

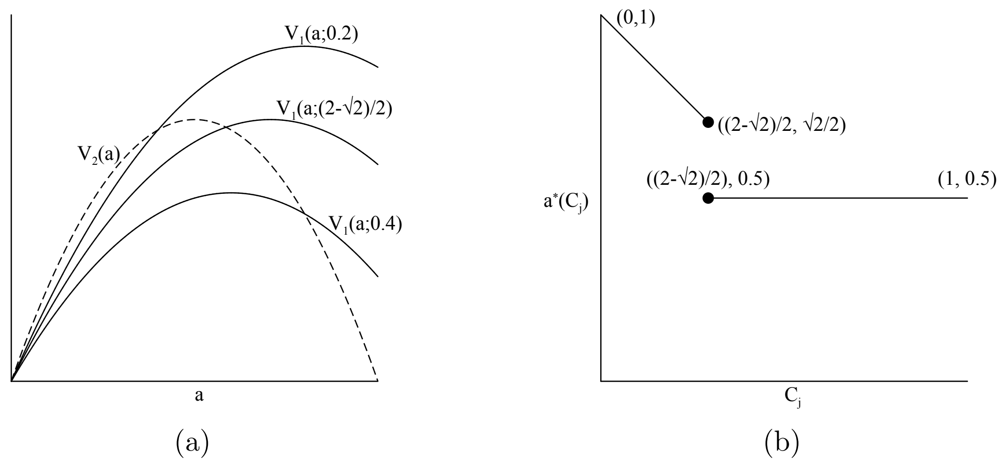


Figure A.2: Figures for the seller-offer game with $\theta \sim U(0,1)$ and buyer driller cost $C_B = 0$. Panel (a) graphs the functions $V_1(a; C_j)$ and $V_2(a)$ as given in Equations 19 and 20. The solid line parabolas are graphs $V_1(a; C_j)$ for three seller driller cost values: $C_j = 0.2$, $C_j = (2-\sqrt{2})/2$, and $C_j = 0.4$. The dashed line parabola is a graph of $V_2(a)$. Panel (b) shows the correspondence between seller drilling cost C_j and optimal ask $a^*(C_j)$, as given in Equation 23. The coordinates in panel (b) give the (x,y) coordinates of various end points in the correspondence $a^*(C_j)$.

leases won by individuals tend to be traded earlier than leases won by firms. These mechanisms include liquidity constraints, rental payments, and risk aversion.

Firms trading later than individuals can hypothetically lead to a lower probability of drilling because leases have a finite length. If individuals tend to sell their leases to buyers at an earlier date than firms, this will mean that a buyer who buys from a firm will typically have less time remaining to drill the lease. Less time remaining in a lease means less time to drill, which could hypothetically lead to a lower drilling probability. However, as we discuss in Section 6, it is unlikely that delay mechanisms are driving our findings about drilling for leases close to existing production: Trade patterns for firms and individuals are similar when comparing leases close to and far from existing production, but only leases close to existing production show large drilling differentials between those won by firms and those won by individuals.

Risk aversion: Individuals may be more risk averse than firms. The processes of oil drilling and finding a buyer are each riddled with uncertainty, especially for a winner with little experience in the industry. The individual may therefore prefer to sell the lease early in order to avoid future uncertainty.

For example, denote the discount factor as $\delta \in (0,1)$. if the individual can either get an offer X today or wait until next year and get an offer Y with probability p, where $X < \delta p Y$, then a risk neutral winner would wait until tomorrow. However, for a risk averse winner with a utility function $U(X) = X^{\sigma}$, where $\sigma < 1$, the winner will prefer to sell today if $X^{\sigma} \geq \delta p Y^{\sigma}$ – or, with a little algebra – if $\sigma \leq -(\log p + \log \delta)/(\log Y - \log X)$. Therefore, as long as individuals are sufficiently risk averse (σ is small enough), individuals will trade earlier than firms.[2]

Multiple offers that do not arrive simultaneously: There may be multiple potential buyers, not all of whom make offers immediately. This can lead to on average longer time until trade when the initial winner is a firm. To see this, suppose that the seller has the option to either accept an offer b, drill and get profits $1 - C$, abandon and get payoff 0, or wait until the next period to potentially receive some other offer. In the current period, the value of waiting until tomorrow can be denoted as $\beta V(C)$. Here β is a discount factor and $V(C)$ takes the expectation over all possible future offers, assuming optimal future behavior (i.e., $V(C) = E_b V(b, C)$). $V(C)$ will be at least weakly decreasing in C because in future periods if one chooses to drill, the value of drilling is lower due to the higher cost.

To examine the comparative statics of how seller cost affects incentives to wait to trade, we give the example where choice's payoff is accompanied by a type 1 extreme value shock. Therefore, the probability that the seller accepts the current period offer b:

$$\frac{\exp(b)}{\exp(b) + \exp(1 - C) + \exp(0) + \exp(\beta V(C))} \tag{24}$$

[2]If $X > \delta p Y$, then both risk averse and risk neutral lessees will prefer to trade in period 1.

Taking derivatives with respect to C yields:

$$\frac{\exp(b)[\exp(1-C)-\exp(\beta V(C))\beta(\partial V/\partial C)]}{[\exp(b)+\exp(1-C)+\exp(0)+\exp(\beta V(C))]^2} \tag{25}$$

Because $\partial V/\partial C \leq 0$, the above expression is positive: Higher drilling costs will increase the probability of accepting the offer b today. Therefore, individuals with higher drilling costs will be more likely than firms to trade in the first period.

An increase in costs will have an ambiguous effect on the probability of waiting until the next period. This is because an increase in drilling cost will decrease the payoff of drilling in the current period as well as a smaller decrease in the payoff of waiting until the next period. If the lessee's top two options are to either drill or to wait, then an increase in cost will lead to an increased probability of waiting. On the other hand, if the lessee's top two options are to accept an offer or to wait, then an increase in cost will lead to a decreased probability of waiting.

As a result of this, it is theoretically ambiguous whether an increase in cost will increase or decrease the probability of drilling in later periods. This also implies that an increase in cost will have an ambiguous effect on the expected date of trade conditional on trade happening.

Price volatility and real options with rental payments: In the presence of price volatility, there can be option value to waiting to trade and/or drill (Dixit and Pindyck, 1994). Here, assignment to a firm versus an individual can affect the probability and timing of trade. We discuss a simple model of price volatility that shows that trade will be more likely to happen and tend to happen earlier when an individual initially wins the lease. The intuition is that because of the rental payment required to continue the lease into the next period, individuals with their higher costs are more likely to abandon the lease initially, and so the buyer must make an early offer in order to secure the lease before it is abandoned.

Suppose there are only two periods. Price is normalized to 1 in period 1; price is drawn from a distribution $f(p)$ in period 2. The quantity of oil under the ground is known and is equal to 1. So, if the cost of drilling is C, the profits from drilling when price is p is $p - C$. We focus on the interesting cases where buyer cost C_B is less than seller cost C_S – e.g., there are gains to trade. We examine the comparative statics of C_S on trade timing and probability.

In each time period $t \in \{1, 2\}$, there are a series of actions. First, trade may happen, where buyer buys from the seller. Second, whoever owns the lease may choose to drill. Payoffs in period 2 are discounted with the discount factor δ. Continuing into the second period requires paying a rental fee $r \geq 0$.

Absent any possibility of trade, the value of a lease in $t = 1$ to a lessee with drilling cost C is:

$$V(C) = \max\left\{1 - C, 0, -r + \delta \int_{C}^{\infty} (p - C) f(p) dp\right\} \tag{26}$$

Here the first term in the max function is the profits of drilling in $t = 1$; the second is the profits of abandoning the lease in period 1, and the third is the profits of waiting until $t = 2$ to potentially drill, where drilling only happens if $p \geq C$.

From this it is clear that the profits of drilling today and the profits of waiting until $t = 2$ are both strictly decreasing in C. Therefore, $V(C)$ is weakly decreasing in C. As long as $C_B < C_S$, gains to trade are at least zero. If in addition $V(C_B) > 0$, then gains to trade are strictly positive.

To model trade, we assume that the buyer has all of the bargaining power. As a result, if the buyer trades in period 1 it will set an offer equal to the seller's outside option $V(C_S)$. If it trades in period 2, after the period 2 price has been realized, it will set an offer equal to $\max\{p - C_S, 0\}$ so long as $p \geq C_B$.

Because the buyer sets an offer equal to the seller's outside option, the buyer's payoff is the gains to trade. Therefore, the buyer will choose to trade in period 1 if the gains to trade in period 1 are higher than the expected gains to trade in period 2. The gains to trade in period 1 are $V(C_B) - V(C_S)$. The expected gains to trade in period 2 are:

$$\delta\left[\int_{C_S}^{\infty}(C_S - C_B)f(p)dp + \int_{C_B}^{C_S}(p - C_B)f(p)dp\right] \tag{27}$$

In other words, if in the future $p \geq C_S$, then the seller's best option is to drill, which means that the gains to trade are the differences in costs $C_S - C_B$. If instead in the future $p \in (C_B, C_S)$, then the seller's best option would be to abandon and get a payoff of zero, which means the gains to trade are $p - C_B$.

From inspection, it is easy to see that the expected gains to trade in period 2 given by Equation 27 is strictly less than $C_S - C_B$. This is due to the fact that $\delta < 1$, that the second integrand is less than $C_S - C_B$ for the range of integration, and that with some probability it will be the case that $p < C_B$ which would imply no positive gains to trade.

When will trade happen, if at all? We examine multiple cases corresponding to the various potential values of $V(C_B)$ and $V(C_S)$:[3]

1. $V(C_B) > 0$ and $V(C_S) = 0$:

   Here, the seller will abandon the lease in period 2. Because the buyer values the lease, they prefer trading to not trading, and therefore must trade in period 1 to secure the lease before the seller abandons it.

2. $V(C_B) = 0$ and $V(C_S) = 0$:

[3] There are only five cases below because other cases are not possible when $C_B < C_S$: Because $V(C)$ is weakly decreasing in C, we will never have the case that $V(C_S) > 0$ and $V(C_B) = 0$. Similarly, we will never have the case that $V(C_S) = 1 - C_S$ but $V(C_B) = -r + \delta\int_{C_B}^{\infty} f(p)dp$: Because $\partial[-r + \delta\int_{C}^{\infty} f(p)dp]/\partial C \in (-1, 0)$ and $\partial(1 - C)/\partial C = -1$, then if $V(C_B) = -r + \delta\int_{C_B}^{\infty} f(p)dp > 1 - C_B$, then the same inequality will hold for higher cost values $C_S > C_B$.

Neither the seller nor the buyer values the lease. The buyer is indifferent between making an offer of zero in period 1 and not making an offer. In the event of not receiving an offer in period 1, the seller will abandon the lease.

3. $V(C_B) = 1 - C_B$ and $V(C_S) = 1 - C_S$:

   The first-period gains to trade are $C_S - C_B$. The expected gains to second-period trade given in Equation 27 are strictly less than $C_S - C_B$, meaning that the buyer will trade in period 1 rather than period 2.

4. $V(C_B) = 1 - C_B$ and $V(C_S) = -r + \delta \int_{C_S}^{\infty} (p - C_S) f(p) dp$:

   The gains to trade in period 1 are:

$$1 - C_B - \left[-r + \delta \int_{C_S}^{\infty} (p - C_S) f(p) dp \right] \tag{28}$$

   Using a little algebra, the difference between the gains to trade in period 1 in Equation 28 and the expected gains to trade in period 2 from Equation 27 can be written as:

$$1 - C_B - \left[-r + \delta \int_{C_B}^{\infty} (p - C_B) f(p) dp \right] \tag{29}$$

   Because this is the case where $V(C_B) = 1 - C_B$, the above equation is positive, meaning that the buyer will prefer to buy in period 1 than period 2.

5. $V(C_B) = -r + \delta \int_{C_B}^{\infty} (p - C_B) f(p) dp$, $V(C_S) = -r + \delta \int_{C_S}^{\infty} (p - C_S) f(p) dp$:

   In this case, the gains to trade from trading in period 1 is exactly equal to the expected gains to trade from trading in period 2 in Equation 27. Therefore, in this case, the buyer will be indifferent between making an offer in period 1 and waiting until period 2.

Given the above, how do seller costs affect the timing of trade? We examine three cases for the three possible values of $V(C_B)$:

1. $V(C_B) = 1 - C_B$. In this case the buyer will trade in period 1 regardless of seller costs, so seller costs have no effect on the timing of trade.

2. $V(C_B) = 0$. In this case, both buyer and seller valuation are zero, and so seller drilling cost will have no effect on the timing of trade.

3. $V(C_B) = -r + \delta \int_{C_B}^{\infty} (p - C_B) f(p) dp$. In this case, a sufficiently large increase in seller drilling cost will shift the seller's lease value from being $V(C_S) = -r + \delta \int_{C_B}^{\infty} (p - C_B) f(p) dp$ to being $V(C_S) = 0$. This shift in seller drilling cost will shift the buyer from being indifferent in whether it trades in period 1 or 2 to strictly preferring to trade in period 2. The intuition is that when $V(C_S) = 0$, the buyer must buy immediately in order to prevent the lease being abandoned by the seller and therefore no longer available on the market.

Therefore, a higher seller drilling cost C_S will cause weakly earlier trade. It will also cause a weakly higher probability of trade, because trade does not happen in the second period when the price draw is low. Therefore, trade will be (weakly) more likely and (weakly) earlier when the winner is an individual rather than a firm.

Although this model predicts that trade outcomes will depend on initial assignment, drilling outcomes will not depend on initial assignment. Within this model, there are no frictions nor inefficiencies, implying that the buyer will drill the well (if it is profitable to do so) regardless of the drilling costs of the initial winner.

Liquidity constraints / rental payments: A lease requires paying regular rental payments of (at the time) \$1/acre/year. Individuals may have found it difficult to borrow funds to pay rental payments as they searched for the buyer who could offer them the highest price. In the following section we discuss a model of liquidity constraints, first for both the case with no informed buyer, and then for the case with an informed buyer that knows the true value of θ.

A.3 Liquidity Constraints with and without an Informed Buyer

In this section we write a model that incorporates liquidity constraints. The model illustrates how the presence of liquidity constraints affects the timing of trade. We show how when individuals are more likely to be liquidity constrained than firms, leases won by individuals will tend to be traded earlier than leases won by firms.

We first explore a model of liquidity constraints with no information asymmetry. Then we explore a model of liquidity constraints augmented with information asymmetry and an informed buyer, similar to that in Section 3 and Appendix A.

For both the case with no information asymmetry and the case with information asymmetry, the game proceeds in three stages (outlined in Figure A.3). The first two stages are reassignment stages, where in each period the initial winner j and the buyer may bargain over the lease. In between the first and second reassignment stage, a rental fee, r, must be paid by whoever owns the lease at the end of period 1 – either the initial winner if trade does not happen in period 1, or the buyer if trade does happen in period 1. We assume that bargaining always takes place by the buyer making take-it-or-leave-it offers O_1 and O_2 in periods $t = 1$ and $t = 2$ respectively.

The third stage is the drilling stage, where the final lease owner decides whether to drill. Payoffs from drilling equal the realized value of θ less the cost of drilling (C_j). For simplicity, we assume that there is no discounting of the payoffs in stages two and three.

Figure A.3: Timeline with Multiple Trading Periods and Liquidity Constraints

To model liquidity constraints, we assume that each type of winner $j \in \{I, F\}$ has some probability of being liquidity constrained, where the probability of a liquidity constraint is given by p_j. We focus on the case where individuals are more likely to be liquidity constrained than firms: $p_I > p_F$. We model liquidity constraints as increasing the effective rental payment when moving from stage one to stage two: If the initial winner wants to wait until stage two before selling to the buyer, a non-liquidity constrained winner will need to pay r to continue but a liquidity constrained winner will need to pay $r(1+\delta)$, with $\delta > 0$. This δ reflects the idea that liquidity constrained individuals will find it especially difficult to find financing. We assume that the buyer can observe the winner's type $j \in \{I, F\}$, but does not observe whether the winner is liquidity constrained. Therefore, the buyer will use two different offers in stages one and two to separate between the liquidity constrained and the non-liquidity constrained.

Throughout, we focus on the equilibrium where the buyer makes take-it-or-leave-it offers to the seller. This assumption means that the buyer will set the offer equal to the seller's outside option. The models are solved with backwards induction, first solving for the stage two offer and then solving for the stage one offer. We solve for the optimal buyer offers assuming that the buyer targets the liquidity constrained types in stage one and the non-liquidity constrained types in stage two. Then after solving for these offers, we discuss why it is optimal for the buyer to differentiate between the two types by using two offers, the first to the liquidity constrained and the second to the non-liquidity constrained.

Buyer does not know θ

In a setting with no informed buyer, we assume that all agents have the same beliefs about the distribution of θ, including the same belief about $E(\theta)$.

The initial winner will either be an individual ($j = I$) or a firm ($j = F$). We denote the potential buyer as LCF, meaning a lowest cost firm.[4] Costs are publicly

[4] As this model assumes that all agents have the same information about all costs and the distribution of θ, the only reasonable buyer of the lease is a buyer with the lowest drilling costs. Therefore,

known, with $C_I > C_F \geq C_{LCF}$. We also allow for there to be some probability that the firm that wins the lease is a lowest cost firm ($C_F = C_{LCF}$), which denote as $p(C_F = C_{LCF})$. This may stem from the case where there is no buyer with lower costs than the lease winner, or from there being multiple firms – both winner and buyer – who are tied for lowest cost, such that the gains to trade are zero. If $C_F = C_{LCF}$ and the initial winner is not liquidity constrained, there is no reason to trade.

In this game, the buyer can make an offer O_1 in stage one and an offer O_2 in stage two if the first one is not accepted. If the lease is traded in stage one, then the buyer (the LCF) acquires the lease, pays the rental payment r, and then chooses whether to drill or not in stage three. If the seller waits until stage two to sell, it must pay the rental payment itself before considering stage two's offer. If it is not liquidity constrained, it will pay r to continue; if it is liquidity constrained, it pays $r(1+\delta)$ to continue.

We solve this by backward induction. In the last period, the expected profits of owning the lease are $\max\{E(\theta) - C_j, 0\}$, where C_j is the drilling cost for the lease holder in stage two. Therefore, the offer O_2 is set to equal the winner's outside option conditional on being in stage two. Because the rental payment has already been paid, it is sunk, and does not affect the second stage payoff. Therefore, the second stage offer is:

$$O_2 = \max\{E(\theta) - C_j, 0\} \tag{30}$$

In stage one, the offer O_1 is set to target the liquidity constrained winner, and is set so that the liquidity constrained winner is indifferent between accepting the offer O_1 and its outside option. The outside option is equal to the max payoff over the possibilities of waiting until the stage two offer, drilling on its own in stage three, or choosing not to drill the lease in stage three, all of which require the seller to pay

we assume that the buyer is the lowest cost firm.

a continuation payment $r(1+\delta)$. Therefore:

$$O_1 = \max\{O_2 - r(1+\delta), E(\theta) - C_j - r(1+\delta), -r(1+\delta)\} \tag{31}$$

$$\Rightarrow O_1 = \max\{E(\theta) - C_j, 0\} - r(1+\delta) \tag{32}$$

Above we have posited that an equilibrium consists of an offer O_1 that targets the liquidity constrained and then a second offer that targets the non-liquidity constrained. Below we discuss why this is an equilibrium because both types of sellers are incentivized to accept the offer targeted to them and because the buyer maximizes payoffs by making the offers.

For each seller type, it is optimal for it to accept the offer targeted to it: The liquidity constrained winner is indifferent between trading immediately in stage one and getting O_1 and waiting, because O_1 is set to be equal to the outside option of waiting, and so an equilibrium where it takes the offer O_1 maximizes its payoffs. The non-liquidity constrained winner strictly prefers to wait rather than trade in stage one: If it waits, it receives a total payoff of $\max\{E(\theta) - C_j, 0\} - r$; if it accepts O_1, it receives a total payoff of $\max\{E(\theta) - C_j, 0\} - r(1+\delta)$, which is strictly worse. And as long as $C_j > C_{LCF}$, it prefers to accept O_2 rather than drill itself.

This separating strategy is also optimal for the buyer: By making a lower offer O_1 in stage one rather than offer O_2 in both periods, the buyer makes an additional expected profit of $r\delta \cdot p_j$ – the intertemporal gains to trade multiplied by the probability that the winner is liquidity constrained.

What are the predictions of this model for trade and drilling? In this model, the probability of trade will be higher when the winner is an individual than when the winner is a firm. Trade happens as long as $E(\theta) \geq C_{LCF}$ and $C_{LCF} < C_j$. With the assumption that a winning firm may have $C_F = C_{LCF}$ but that for individuals it is always the case that $C_I > C_{LCF}$, the probability of trade will be higher for individuals than firms.

The expected time of trade will be earlier when the winner is an individual relative to when the winner is a firm. Conditional on trade happening (e.g., conditional on $C_F > C_{LCF}$), the expected date of trade for a winner of type j is $t = 1 + (1 - p_j)$. As long as $p_I > p_F$, leases won by individuals will be traded earlier than leases won by firms.

In this model, the probability of drilling will be identical for leases won by firms and leases won by individuals. Drilling happens so long as $E(\theta) \geq C_{LCF}$, which does not depend on the identity of the initial assignment.

Buyer knows θ

Next, we examine how the model changes if we introduce information asymmetry via an informed buyer. Our model of information asymmetry is very similar to that in Section 3; here we enrich it by allowing for the possibility of liquidity constrained winners, showing how the information asymmetry interacts with the liquidity constraints.

We now assume that the buyer is the nearby producing firm (NPF). As discussed in Section 2, the nearby producing firm will typically have a cost advantage and will have a more precise signal of the true underlying amount of oil that can be produced from a lease. Therefore here, as in Section 3, we assume that the NPF is the potential buyer and that the NPF knows the true value of θ. We also assume that $C_j > C_{NPF}$ – the nearby producing firm has a strict cost advantage over the initial winner, e.g., due to its economies of scale. And given that the nearby producing firms very rarely win in the data, we assume that the NPF is never the initial winner. Because of the complexity of the model, we focus on the comparative statics of liquidity constraints (p_j), which are easier to solve for, rather than on the comparative statics of drilling cost C_j, which are significantly more challenging to solve for. Below we show that even if firm and individual drilling costs are equal, a higher probability that the individual is liquidity constrained will lead to leases won by individuals having a higher probability of trade, earlier timing of trade, and potentially a higher probability of trade.

As in the previous model, we solve for two offers: a first-stage offer O_1^* that targets the liquidity constrained winner and a second-stage offer O_2^* that targets the non-liquidity constrained. As in Section 3, we solve for strategies that follow a threshold pooling rule: The buyer will make an offer of O_1^* in stage one if $\theta \geq \theta_1^* \equiv O_1^* + C_{NPF} + r$. The buyer will make an offer of O_2^* in stage two if $\theta \geq \theta_2^* \equiv O_2^* + C_{NPF}$. The first stage offer O_1^* will be made only if $\theta \geq \theta_1^* \equiv O_1^* + C_{NPF} + r$ because if the buyer buys in stage one, it will have to pay the rental fee to continue to stage two. In contrast, the second stage offer O_2^* will be made only if $\theta \geq \theta_2^* \equiv O_2^* + C_{NPF}$, reflecting that if the buyer buys in stage two, the initial winner already paid the rental fee to move from stage one to stage two, and therefore that fee is sunk. Later we show that it is optimal for the buyer to make these offers and show that it is optimal for each of the seller types to accept the offers targeted to them.

In the derivation below, we assume that the distribution $F(\theta)$ satisfies $\partial E(\theta|\theta \geq Y)/\partial Y \in (0,1)$. This distribution assumption is satisfied by a number of distributions including the normal and the uniform. This assumption helps avoid the possibility that there are multiple offers that satisfy the equilibrium condition on the offer amount.[5]

Second stage: We again solve the model using backwards induction. During the second stage, the NPF either makes a take-it-or-leave-it offer $O_2 \geq 0$ to the initial lessee (i) or makes no offer. Because the initial lessee knows that the NPF knows the true value of θ, the initial lessee updates its beliefs about θ based upon whether it re-

[5] A more general assumption is used in Appendix A, where there we only assume that $E(\theta|\theta \geq Y)$ is continuous in Y. One distribution that does not satisfy the assumption that $\partial E(\theta|\theta \geq Y)/\partial Y \in (0,1)$ is the exponential distribution, as discussed in Appendix A.

ceived an offer and the offer size. If it did receive an offer, the initial lessee updates its expectation of the value of retaining the lease to be $\max\{0, E[\theta|O_2] - C_i\}$. It will only trade if the offer exceeds profits of retaining the lease: $O_2 \geq \max\{0, E[\theta|O_2] - C_i\}$. The initial lessee knows that the NPF only makes an offer when it is rational, and therefore the true θ is at least $O_2 + C_{NPF}$. Therefore, the definition of O_2^* is identical to the definition of O^* in Section 3, and is defined by the following implicit equation:

$$O_2^* = \max\{E(\theta|\theta \geq O_2^* + C_{NPF}) - C_j, 0\} \tag{33}$$

and the NPF will make an offer of O_2^* if $\theta \geq O_2^* + C_{NPF} \equiv \theta_2^*$.

First stage: Here we solve for the first stage offer O_1^* that is targeted to the liquidity constrained seller. The buyer will make an offer O_1^* equal to the liquidity constrained seller's outside option. Therefore, the first stage offer O_1^* is implicitly defined by the following equation that sets the offer O_1^* equal to the seller's outside option conditional on receiving an offer O_1^*:

$$\begin{aligned}
O_1^* = {} & p(\theta \geq O_2^* + C_{NPF}|\theta \geq O_1^* + C_{NPF} + r) \\
& \cdot \max\{O_2^*, E(\theta|\theta \geq O_1^* + C_{NPF} + r, \theta \geq O_2^* + C_{NPF}) - C_I, 0\} \\
& + (1 - p(\theta \geq O_2^* + C_{NPF}|\theta \geq O_1^* + C_{NPF} + r)) \\
& \cdot \max\{E(\theta|\theta \geq O_1^* + C_{NPF} + r, \theta < O_2^* + C_{NPF}) - C_i, 0\} \\
& - r(1 + \delta)
\end{aligned} \tag{34}$$

where this offer O_1^* will only be made if $\theta \geq O_1^* + C_{NPF} + r \equiv \theta_1^*$:

In other words, O_1^* is set equal to the expected outside option value minus the continuation cost, where the expected outside option is a probability-weighted average of the outside option in the case where an offer in stage two is received and the outside option in the case where an offer in stage two is not received. In the following two theorems, we prove that $O_1^* < O_2^*$ and that $\theta_1^* < \theta_2^*$, meaning that an offer is more likely to be made in stage one than in stage two and that the offer made in stage one will be lower than that in stage two:

Theorem 1. *The threshold value of θ for making an offer in stage one, $\theta_1^* \equiv O_1^* + C_{NPF} + r$, will be less than the threshold for making an offer in stage two, $\theta_2^* \equiv O_2^* + C_{NPF}$.*

Proof. Proof by contradiction: Suppose that $O_1^* + C_{NPF} + r \geq O_2^* + C_{NPF}$. In that case, $p(\theta \geq O_2^* + C_{NPF} | \theta \geq O_1^* + C_{NPF} + r) = 1$, and we can write Equation 34 as:

$$O_1^* = \max\{O_2^*, E(\theta|\theta \geq O_1^* + C_{NPF} + r) - C_j, 0\} - r(1+\delta) \tag{35}$$

The definition of O_2^* in Equation 33, the assumption that $O_2^* + C_{NPF} \leq O_1^* + C_{NPF} + r$, and the fact that $E(\theta|\theta \geq Y)$ is increasing in Y jointly implies:

$$\theta_2^* = \max\{E(\theta|\theta \geq O_2^* + C_{NPF}) - C_j, 0\} \tag{36}$$

$$\leq \max\{E(\theta|\theta \geq O_1^* + C_{NPF} + r) - C_j, 0\} \tag{37}$$

Therefore, we can rewrite Equation 35 as:

$$O_1^* = \max\{E(\theta|\theta \geq O_1^* + C_{NPF} + r) - C_j, 0\} - r(1+\delta) \tag{38}$$

Notice that if $r = 0$, then $O_1^* = O_2^*$. Similarly, if $r = 0$, then $\theta_1^* \equiv O_1^* + C_{NPF} + r = \theta_2^* \equiv O_2^* + C_{NPF}$. Therefore, we first calculate $\partial O_1^*/\partial r$ and then $\partial \theta_1^*/\partial r$ to show how θ_1^* compares to θ_2^*:

To solve for $\partial O_1^*/\partial r$, we take derivatives with respect to r in Equation 38.[6]

[6] Here we take derivatives under the assumption that the value of the max function above is greater than zero because above we have assumed that the initial offer O_1^* is strictly positive. If the value of the max function is equal to zero, then $\partial O_1^*/\partial r = -(1+\delta)$ and $\partial \theta_1^*/\partial r = -\delta$, which also implies a contradiction of the initial assumption in this proof by contradiction.

Rearranging yields:

$$\frac{\partial O_1^*}{\partial r} = \frac{\left[\frac{\partial E}{\partial Y} - (1+\delta)\right]}{1 - \frac{\partial E}{\partial Y}} \tag{39}$$

where

$$\frac{\partial E}{\partial Y} \equiv \frac{\partial E(\theta|\theta \geq \theta_1^* + C_{NPF} + r)}{\partial(\theta_1^* + C_{NPF} + r)} \tag{40}$$

To solve for $\partial\theta_1^*/\partial r$, we use the fact that if $\theta_1^* = O_1^* + C_{NPF} + r$, then $\partial\theta_1^*/\partial r = \partial O_1^*/\partial r + 1$. This combined with the above equation yields:

$$\frac{\partial \theta_1^*}{\partial r} = \frac{\left[\frac{\partial E}{\partial Y} - (1+\delta) + \left(1 - \frac{\partial E}{\partial Y}\right)\right]}{1 - \frac{\partial E}{\partial Y}} = \frac{-\delta}{1 - \frac{\partial E}{\partial Y}} < 0 \tag{41}$$

Therefore, with a positive rental payment of r, the value of θ_1^* will be less than θ_2^*, which is contrary to the initial assumption that $\theta_1^* \geq \theta_2^*$. □

Theorem 2. $O_1^* \leq O_2^* - r(1+\delta)$. *In other words, accepting the offer of O_1^* in stage one gives a lower payoff than waiting until stage two, paying the rental payment, and getting an offer of O_2^* with surety.*

Proof. Given that $E(\theta|\theta \geq Y)$ is increasing in Y, that the previous theorem shows that $O_2^* + C_{NPF} \geq O_1^* + C_{NPF} + r$, and the definition of O_2^* in Equation 33, the value of the first max function in Equation 34 is equal to O_2^*. Therefore, we can rewrite Equation 34 as:

$$\begin{aligned} O_1^* = {} & p(\theta \geq O_2^* + C_{NPF}|\theta \geq O_1^* + C_{NPF} + r) \cdot O_2^* \\ & + (1 - p(\theta \geq O_2^* + C_{NPF}|\theta \geq O_1^* + C_{NPF} + r)) \\ & \cdot \max\{E(\theta|\theta \geq O_1^* + C_{NPF} + r, \theta < O_2^* + C_{NPF}) - C_i, 0\} \\ & - r(1+\delta) \end{aligned} \tag{42}$$

In addition, the fact that the interval $[O_1^* + C_{NPF} + r, O_2^* + C_{NPF})$ is to the left of the interval $[O_2^* + C_{NPF}, \infty)$ and the fact that $E(\theta|\theta \geq Y)$ is increasing in Y means that the value of the max function in Equation 42 is weakly less than O_2^*. This implies that $O_1^* \leq O_2^* - r(1+\delta)$.

$\square$

Discussion: Do the above strategies constitute an equilibrium? As can be seen in the analysis above, the liquidity constrained seller will be indifferent between accepting an offer O_1^* and waiting. Therefore, for the liquidity constrained seller, accepting the offer O_1^* is optimal.

Would the non-liquidity constrained seller prefer to accept an offer O_1^* or wait until the second stage? Because O_1^* is set to the liquidity constrained seller's outside option which includes paying a continuation cost $r(1+\delta)$, a non-liquidity constrained seller would strictly prefer to reject the offer and continue in the lease because by so doing it only pays a continuation cost of r. Therefore, the buyer's strategy of using two offers in the two stages induces the liquidity constrained and non-liquidity constrained types to reveal their types.

Would the buyer prefer to use the two offers O_1^* and O_2^* targeted to the liquidity constrained and non-liquidity constrained types respectively, or some other strategy? From above, the strategy of offering O_2^* as long as $\theta \geq O_2^*$ is optimal in the second stage because all rental payments are sunk. Is it optimal to offer O_1^* or some other offer in stage one? Making a lower offer will ensure that neither type of seller will want to buy – both find it better to wait for the possibility of O_2^* in stage two. Making an offer higher than O_1^* will ensure that the liquidity constrained type sells, but it will give the buyer a lower payoff than if it offered O_1^*. Therefore, an offer of O_1^* is optimal from the buyer's perspective.

What are the implications of the model for trade and drilling outcomes? Here, for simplicity, we focus on the comparative statics of liquidity constraints of p_j, with

$p_I > p_F$. To isolate the effect of liquidity constraints, we assume seller drilling cost is fixed, with $C_I = C_F$. We take this approach to illustrate the intuition of the effect of liquidity constraints in the presence of an informed buyer.

The buyer's strategy will remain the same regardless of whether the winner is an individual or firm: Offer O_1^* in stage one if $\theta \geq O_1^* + C_{NPF} + r$ and offer O_2^* in stage two if $\theta \geq O_2^* + C_{NPF}$. This affects the probability of trade: Because it is more likely that the buyer is willing to offer O_1^* in stage one than it is to offer O_2^* in stage two, individuals will have a higher probability of trade than firms because individuals are more likely to be liquidity constrained and therefore to accept O_1^* when it is offered.[7]

This model also implies that conditional on trade happening, trade will happen earlier on average when individuals win, because individuals are more likely than firms to trade when O_1^* is offered.[8] Therefore, this model shows how the combination of delay mechanisms (liquidity constraints) with an informed buyer can lead to both the trade probability and trade timing patterns we observe in our data.

Similar intuition holds in examining drilling probabilities. Holding seller drilling cost fixed but allowing the probability that the seller is liquidity constrained to vary depending on whether the winner is an individual or a firm (with $p_I > p_F$), one possible outcome of the model is that no agent drills in the absence of trade. Because individuals are more likely to trade overall, leases won by individuals may be also more likely to be drilled than leases won by firms.

[7] More formally, the probability of trade is $p_j(1 - F(\theta_1^*)) + (1 - p_j)(1 - F(\theta_2^*))$ which is increasing in p_j because $(1 - F(\theta_1^*)) \geq (1 - F(\theta_2^*))$.

[8] More formally, the expected time of trade, conditional on trade, is $[p_j(1 - F(\theta_1^*)) + 2(1 - p_j)(1 - F(\theta_2^*))]/[p_j(1 - F(\theta_1^*)) + (1 - p_j)(1 - F(\theta_2^*))]$. Taking derivatives with respect to p_j yields gives $-(1 - F(\theta_1^*))(1 - F(\theta_2^*))/[p_j(1 - F(\theta_1^*)) + (1 - p_j)(1 - F(\theta_2^*))]^2 \leq 0$, meaning that an increase in the probability the seller is liquidity constrained will decrease the expected time to trade conditional on trade.

B Online Appendix: Empirics

This section presents additional details about the data, evidence of exogeneity, results, and analysis about different entrant types. We close by analyzing the effect of oil prices on our analysis.

B.1 Further data discussion

We check whether it appears that entrants followed the legal requirement that they do not enter a lottery multiple times. Using the names of the first-, second-, and third-place winners, we find only four cases of specific leases where there were two winners that had both the same first and last name. Three of the four cases were those where one of the winners had the suffix "Jr.". There was only one case which either involved an entrant entering multiple times (a violation of the rules) or some coincidence (two different individuals with identical names both entering submissions for the same lease).

Our analysis relies on measures of whether the lease buyer was a nearby producing firm (e.g., Table 4). To construct this measure, we combine information on the buyer identity from the LR2000 with WOGCC data on operator identities for nearby parcels.

It is not possible to perfectly identify whether a given lease was sold to a particular firm. There are a number of reasons for this: First, operators are not always the same as lessees. Operator information will be unable to identify cases where the buyer was a lessee but not an operator. Second, operator data was only available for 1978 and later. Third, both operator data and LR2000 data are missing corporate relationships such as subsidiaries, which is problematic when the lease buyer and the nearby firm are under the same corporate umbrella but listed under different names. Fourth, identifying whether a buyer was a nearby producing firm requires matching text strings, which is rarely perfect. As a result, our measure of whether a lease was sold to a nearby producing firm is a binary variable that exhibits measurement error. In Table 4, where this variable is the dependent variable, measurement error is non-classical and results in attenuation bias.

SIXTH PRINCIPAL MERIDIAN

WYOMING

#1096 W 0316078
T 17 N, R 60 W, Laramie
Sec 6: Lots 3, 7, S½NE¼, SE¼NW¼, E½SW¼, N½SE¼
8: N½NW¼
T 18 N, R 60 W
Sec 29: NE¼NW¼, W½SW¼
30: Lot 1, NW¼NE¼, NE¼NW¼, NE¼SE¼
31: NE¼, E½SW¼, W½SE¼
32: S½NE¼, W½
1444.21 A

#1097 W 23542
T 26 N, R 60 W, Goshen
Sec 3: SW¼NW¼, NW¼SW¼
4: SE¼NE¼, E½SW¼, NE¼SE¼
8: SE¼NE¼, SE¼NW¼
10: Lot 4
15: Lot 1, SE¼NW¼, NE¼SW¼
22: Lot 3, SE¼NW¼, NE¼SW¼
28: NE¼SE¼
29: E½SW¼
32: SW¼, S½SE¼
33: SW¼NE¼, NW¼SE¼
1099.17 A

#1098 W 5212
T 35 N, R 60 W, Niobrara
Sec 18: NE¼NE¼
40.00 A

#1099 W 43988
T 36 N, R 60 W, Niobrara
Sec 31: Lot 1, NE¼, E½NW¼, NE¼SE¼
32: S½
638.78 A

#1100 W 0309148
T 36 N, R 62 W, Niobrara
Sec 13: W½NE¼, E½NW¼
160.00 A

#1101 W 0316090
T 39 N, R 62 W, Niobrara
Sec 35: NW¼
160.00 A

#1102 W 0310335
T 43 N, R 62 W, Weston
Sec 11: W½SW¼, W½SE¼
14: Lots 1, 4, E½NW¼, W½SE¼
394.53 A

#1106 W 5223
T 38 N, R 63 W, Niobrara
Sec 13: S½NE¼, SE¼
14: NW¼NE¼, S½NE¼, NW¼, SE¼
23: NW¼NE¼
720.00 A

#1107 W 0316100
T 43 N, R 63 W, Weston
Sec 29: SE¼SE¼
40.00 A

#1108 W 0324155
T 37 N, R 64 W, Niobrara
Sec 13: N½SW¼
80.00 A

#1109 W 23556
T 40 N, R 64 W, Niobrara
Sec 7: Lot 3, S½NE¼, NE¼SW¼
166.31 A

#1110 W 0220694-A
T 40 N, R 64 W, Niobrara
Sec 13: SW¼
25: SW¼NW¼, E½SE¼
280.00 A

#1111 W 0220694
T 40 N, R 64 W, Niobrara
Sec 14: W½NW¼
80.00 A

#1112 W 39112
T 41 N, R 64 W, Weston
Sec 4: S½SE¼
14: S½SE¼
160.00 A

#1113 W 0314722
T 46 N, R 64 W, Weston
Sec 17: NW¼NW¼
40.00 A

#1114 W 11816-A
T 48 N, R 64 W, Weston
Sec 4: SE¼SW¼, NE¼SE¼, S½SE¼
160.00 A

#1115 W 0316107
T 35 N, R 65 W, Niobrara
Sec 9: E½SE¼
10: SW¼NW¼, W½SW¼
200.00 A

Figure B.1: Raw data example of parcels offered in the June 1975 lottery.

B.2 Further Evidence of Exogeneity

This section provides additional results relating to empirical exogeneity. We first discuss further evidence that our restricted sample resolves endogeneity concerns. We then discuss summary statistics about the total number of times that an entrant appears, and show that total entrances within a month is not correlated with the probability of a first place win.

Additional Evidence of Exogeneity within the Restricted Sample

Table B.3 presents correlation tables of selected variables for our full sample, while Table B.4 presents the same correlation table albeit for the restricted sample. Using the restricted sample, we find virtually zero correlation between whether the first place winner is a firm and other ex-ante variables, including number of entries, lease area, and whether there was nearby production within 2.6 miles.

Tables B.5 and B.6 reproduce the balance tests in Table 2, but do so separately for leases not close to existing production (Table B.5) and leases close to existing production (Table B.6). Similarly, Tables B.7 and B.8 reproduce Table B.4, but do it separately for leases not close to existing production (Table B.7) and leases close to existing production (Table B.8). Broadly, we find balance within the restricted sample when limited to leases far from existing production and leases close to existing production. The only statistically significant difference is at the 10% level in the variance of the number of entries.

We examine whether the probability that the firm wins is correlated with other ex-ante features. In Table B.9 we use the restricted sample and regress an indicator for whether the winner was a firm on measures of distance. In Column (1) we use the nearby production flag – whether there was nearby production within 2.6 miles. In Column (2) we use the more finely disaggregate distance bins for leases within 0 and 5.2 miles of existing production. Overall, we do not find that distance bins predict the probability that the firm is the first place winner, as the p-value for a test that the coefficients for all bins is equal to zero is 0.235. Although one bin (1.7-2.1 miles) has a higher likelihood of the firm winning than others, the coefficients for nearby bins (1.2-1.7 miles and 2.1-2.6 miles) are close to zero, which suggests no systematic correlation between distance and the probability that the firm wins.

	1	2	3	4	5	6	7	8	9	10	11	12
1: Number of entries	1.00											
2: Area	0.41	1.00										
	(0.00)											
3: Nearby production	0.14	-0.14	1.00									
	(0.00)	(0.00)										
4: Reassign within 12 years	0.21	0.09	0.06	1.00								
	(0.00)	(0.00)	(0.00)									
5: Log time to first trade to firm	-0.20	-0.07	-0.05	-0.15	1.00							
	(0.00)	(0.00)	(0.00)	(0.00)								
6: Log time to first trade	-0.19	-0.09	-0.05	-0.16	0.78	1.00						
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)							
7: Total trades to firms within 12 years	0.22	0.08	0.09	0.36	-0.21	-0.16	1.00					
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
8: Total trades within 12 years	0.20	0.09	0.09	0.41	-0.12	-0.22	0.83	1.00				
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					
9: Drill within 12 years	0.20	0.08	0.18	0.10	-0.04	-0.06	0.20	0.19	1.00			
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)				
10: Produce within 12 years	0.17	0.02	0.16	0.07	-0.02	-0.03	0.16	0.14	0.63	1.00		
	(0.00)	(0.01)	(0.00)	(0.00)	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)			
11: Number of firms that win	-0.10	-0.06	-0.02	-0.10	0.09	0.11	-0.06	-0.08	-0.03	-0.02	1.00	
	(0.00)	(0.00)	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.08)		
12: First-place winner is firm	-0.06	-0.04	0.00	-0.15	0.13	0.14	-0.08	-0.10	-0.02	-0.02	0.58	1.00
	(0.00)	(0.00)	(0.95)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.10)	(0.08)	(0.00)	

Table B.3: Correlations between selected variables for each observation using the unrestricted sample with all 10,762 parcels. Each parentheses reports the p value of a test where the null hypothesis is that the true correlation is zero.

	1	2	3	4	5	6	7	8	9	10	11	12
1: Number of entries	1.00											
2: Area	0.40	1.00										
	(0.00)											
3: Nearby production	0.09	-0.15	1.00									
	(0.00)	(0.00)										
4: Reassign within 12 years	0.20	0.10	0.02	1.00								
	(0.00)	(0.00)	(0.35)									
5: Log time to first trade to firm	-0.16	-0.03	-0.01	-0.19	1.00							
	(0.00)	(0.35)	(0.80)	(0.00)								
6: Log time to first trade	-0.14	-0.07	0.01	-0.19	0.83	1.00						
	(0.00)	(0.02)	(0.67)	(0.00)	(0.00)							
7: Total trades to firms within 12 years	0.22	0.09	0.03	0.45	-0.24	-0.18	1.00					
	(0.00)	(0.00)	(0.30)	(0.00)	(0.00)	(0.00)						
8: Total trades within 12 years	0.18	0.08	0.02	0.50	-0.18	-0.25	0.84	1.00				
	(0.00)	(0.00)	(0.50)	(0.00)	(0.00)	(0.00)	(0.00)					
9: Drill within 12 years	0.20	0.09	0.16	0.11	-0.00	0.01	0.17	0.17	1.00			
	(0.00)	(0.00)	(0.00)	(0.00)	(0.94)	(0.67)	(0.00)	(0.00)				
10: Produce within 12 years	0.17	0.03	0.16	0.06	0.06	0.04	0.11	0.11	0.63	1.00		
	(0.00)	(0.27)	(0.00)	(0.01)	(0.04)	(0.15)	(0.00)	(0.00)	(0.00)			
11: Number of firms that win	.	.	.	.	.	.	.	.	.	.	.	
	.	.	.	.	.	.	.	.	.	.		
12: First-place winner is firm	-0.01	-0.01	0.03	-0.23	0.23	0.25	-0.15	-0.19	-0.00	-0.02	.	1.00
	(0.74)	(0.59)	(0.19)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.51)	.	

Table B.4: Correlations between selected variables for each observation using the restricted sample with exactly one firm winner (1,800 parcels). Each parentheses reports the p value of a test where the null hypothesis is that the true correlation is zero.

	Parcels won by:		Difference
	Individuals	Firms	(p-value)
Number of Entries Mean	401.82	388.76	0.69
Number of Entries Variance	577.25	600.75	0.31
Acreage Mean	677.78	661.48	0.68
Acreage Variance	707.92	695.40	0.66
Nearby Production Mean	0.00	0.00	.
Nearby Production Variance	0.00	0.00	.

Table B.5: We restrict the sample to the 1,424 parcels where exactly one firm appeared amongst the three winners and there is no nearby production. Statistics for parcels won by individuals are reported in Column (1), while Column (2) reports those won by firms. Column (3) reports the p-value from an equality test.

	Parcels won by:		Difference
	Individuals	Firms	(p-value)
Number of Entries Mean	547.76	528.34	0.83
Number of Entries Variance	864.53	756.39	0.08
Acreage Mean	430.31	435.29	0.92
Acreage Variance	484.73	505.48	0.57
Nearby Production Mean	1.00	1.00	.
Nearby Production Variance	0.00	0.00	.

Table B.6: We restrict the sample to the 376 parcels where exactly one firm appeared amongst the three winners and there is nearby production. Statistics for parcels won by individuals are reported in Column (1), while Column (2) reports those won by firms. Column (3) reports the p-value from an equality test.

	1	2	3	4	5	6	7
1: Number of entries	1.00						
2: Area	0.40	1.00					
	(0.00)						
3: Reassign within 12 years	0.20	0.11	1.00				
	(0.00)	(0.00)					
4: Drill within 12 years	0.16	0.09	0.08	1.00			
	(0.00)	(0.00)	(0.00)				
5: Produce within 12 years	0.10	0.01	0.04	0.55	1.00		
	(0.00)	(0.74)	(0.11)	(0.00)			
6: Number of firms that win	.	.	.	.	.	.	
	.	.	.	.	.		
7: First-place winner is firm	-0.01	-0.01	-0.23	0.04	0.01	.	1.00
	(0.69)	(0.68)	(0.00)	(0.12)	(0.79)	.	

Table B.7: Correlations between selected variables for each observation using the restricted sample with exactly one firm winner and there is no nearby production (1,424 parcels). Each parentheses reports the p value of a test where the null hypothesis is that the true correlation is zero.

	1	2	3	4	5	6	7
1: Number of entries	1.00						
2: Area	0.57	1.00					
	(0.00)						
3: Reassign within 12 years	0.22	0.05	1.00				
	(0.00)	(0.38)					
4: Drill within 12 years	0.24	0.23	0.18	1.00			
	(0.00)	(0.00)	(0.00)				
5: Produce within 12 years	0.25	0.18	0.10	0.71	1.00		
	(0.00)	(0.00)	(0.05)	(0.00)			
6: Number of firms that win	.	.	.	.	.	.	
	.	.	.	.	.		
7: First-place winner is firm	-0.01	0.00	-0.26	-0.11	-0.07	.	1.00
	(0.83)	(0.92)	(0.00)	(0.03)	(0.17)	.	

Table B.8: Correlations between selected variables for each observation using the restricted sample with exactly one firm winner and nearby production exists (376 parcels). Each parentheses reports the p value of a test where the null hypothesis is that the true correlation is zero.

	(1) Firm wins	(2) Firm wins
Nearby Production		
Flag	0.036	
	(0.028)	
<1.2 miles		0.043
		(0.045)
1.2-1.7 miles		-0.002
		(0.059)
1.7-2.1 miles		0.136**
		(0.057)
2.1-2.6 miles		0.003
		(0.049)
2.6-3.2 miles		-0.027
		(0.041)
3.2-4.2 miles		0.064
		(0.043)
4.2-5.2 miles		0.026
		(0.042)
Intercept	0.336***	0.330***
	(0.013)	(0.015)
R squared	0.001	0.005
p value: test intercept=1/3	0.809	0.840
p value: intercept + nearby prod=1/3	0.112	
p value: all distance bin indicators=0		0.235
Observations	1800	1800

Table B.9: Regressions testing for corruption: These regressions use the restricted sample where exactly one of the three winners is a firm. The regressions use conventional OLS standard errors.

Total Appearances and Probability of First-Place Wins

In this section, we first discuss summary statistics in Table B.10 of the total appearances and total first-place wins for any entrant in the data. We find that there are 8,527 total firms and individuals that appear in the data: 8,181 are individuals and 346 are firms. We find that over 50% of them only appear in the data once. The average individual appears 3.7 times in the data; the average firm appears about 5.9 times in the data. The number of times that an individual or firm appears in our data has an extremely long right tail, with a maximum of 320 appearances for individuals and 98 appearances for firms.

We next examine how the probability of winning is correlated with how frequently an entrant appears in the data. If the lottery is random, then for any entrant that appears in the data (e.g., appears as a first-, second-, or third-place winner), the probability that the entrant appears as a first-place winner should equal to 1/3 in expectation, and will not be correlated with other observable and unobservable characteristics, such as the total number of times that the entrant appears in the data in a given month.[9]

Table B.11 verifies that the probability a given entrant appears in first place is uncorrelated with other factors. There, we run regressions where the dependent variable is the number of wins in a given month divided by the total number of appearances in that same given month. Right-hand-side variables include whether the entrant is a firm and the log total number of times the entrant appears in the data for that month. An F test examining whether all explanatory variables (other than the intercept) are equal to zero finds large p values, implying that we cannot reject the null hypothesis of exogenous assignment. In Figure B.2, we graph the

[9] Because winning the lottery may lead an entrant to be more or less likely to submit an entry form in future lotteries, the probability that the entrant appears as a first-place winner is correlated with the total number of times that an entrant appears in the data *over all months in the data*. Future research will explore this, examining the treatment effect of winning a lottery on later industry activity.

number of times each entrant appears in the data and the fraction of times that each entrant wins. We find that the distribution of fraction of times won appears to correspond to a binomial distribution divided by total number of appearances, with a convergence to 1/3 as the total number of appearances grows large.

	Appearances			1st Place Wins		
	All	Individuals	Firms	All	Individuals	Firms
Observations	8527	8181	346	8527	8181	346
Mean	3.79	3.7	5.91	1.26	1.23	2.01
Minimum	1	1	1	0	0	0
25th percentile	1	1	1	0	0	0
50th percentile	1	1	2	0	0	1
75th percentile	3	2	5	1	1	2
90th percentile	7	7	13	3	2	5
95th percentile	13	13	24	5	4	9
99th percentile	44	42	61	15	15	23
Maximum	320	320	98	106	106	33

Table B.10: Summary statistics for the number of times entrants appear in data over all months (as either first-, second-, or third-place winners) and the number of times entrants appear in first place.

	(1)	(2)	(3)	(4)
	Fraction of appearances that are first place wins			
Log(total appearances)	0.002	0.000	0.003	0.001
	(0.004)	(0.012)	(0.005)	(0.012)
(Log(total appearances))^2		0.001		0.002
		(0.006)		(0.006)
Firm			0.017	0.017
			(0.015)	(0.015)
Firm x log(total appearances)			-0.017	-0.013
			(0.019)	(0.054)
Firm x (log(total appearances))^2				-0.004
				(0.032)
Constant	0.332***	0.332***	0.331***	0.331***
	(0.003)	(0.003)	(0.003)	(0.004)
R squared	0.000	0.000	0.000	0.000
p-value	0.604	0.701	0.647	0.782
Observations	23141	23141	23141	23141

Table B.11: Regression estimates where dependent variable is the fraction of times within a given month that an entrant is in first place, conditional on appearing in the data for that month (e.g., in first-, second-, and/or third-place). The data are at the entrant by month level – e.g., one observation per entrant by month, conditional on the entrant appearing in the data for the given month. The reported p value is the p value for an F test that all explanatory variables (other than the constant) is equal to zero. Standard errors are heteroskedasticity-robust.

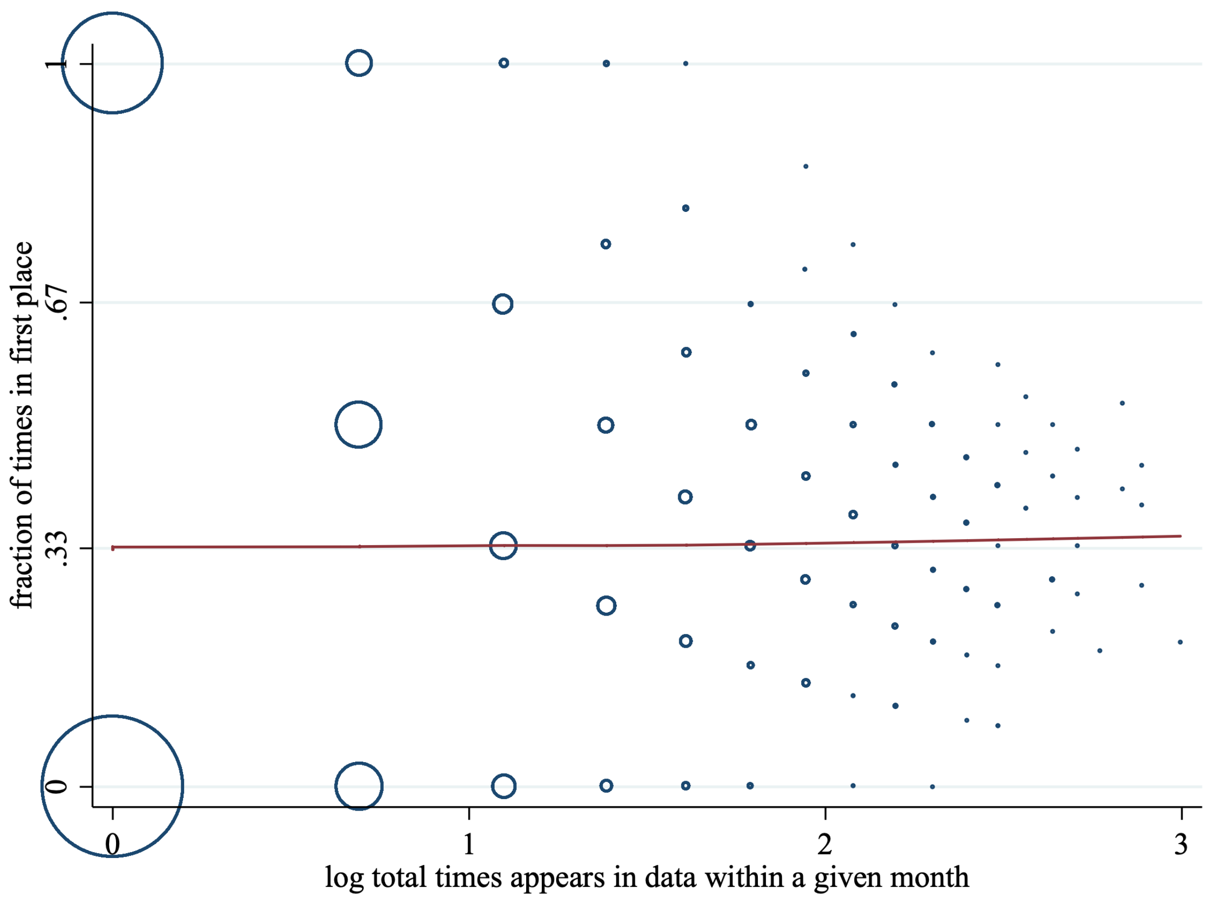


Figure B.2: Fraction of times that an entrant wins first place in a given month relative to the total number of times that an entrant appears in the data within a given month. The figure also plots a local polynomial best fit regression. Size of markers is proportional to the number of observations.

B.3 Additional Trade Analysis

This section presents additional analysis of trade patterns in the data. We first present additional robustness checks to our main trade analysis. Then, we discuss results that show that individuals rarely drilled unless they also traded, indicating that our findings for drilling for leases close to existing production are unlikely to be driven by individuals overoptimistically drilling.

Robustness Results

Table B.12 reports additional regression results examining the timing of trade. These results correspond to the trade timing histograms in Figure 1. Column (1) shows that firm winners are much less likely to have the first transfer be during the lease's first year. Column (2) shows that the likelihood of first transfer within years two through ten is lower when firms initially win the lease. Column (3) shows that firms are slightly more likely than individuals to transfer their leases for the first time after ten years, while Column (4) shows that firms are significantly more likely to never transfer their leases than individuals are.

We next show that trade patterns are similar when we examine the first date a lease was traded *to a firm* specifically, instead of the first date it was traded overall. Table B.13 and Figure B.3 are analogous to Table 3 and Figure 1, respectively, but give summary statistics on whether and when the lease was first transferred to a firm. Columns (1) and (2) of Table B.13 show that lease winners who were individuals were more likely to transfer leases to firms than lease winners that were firms. Columns (3) and (4) of Table B.13 show that conditional on trade happening, lease winners that were individuals tended to transfer their leases to firms faster than lease winners that were firms. Similarly, Figures B.3 and 1 yield similar conclusions. As discussed in Online Appendix Section B.1, the limitations of the LR2000 means that this measure may be understating the extent to which leases were transferred to firms.

	(1) Transfer Within First Year	(2) Transfer During Years 2-10	(3) Transfer After Ten Years	(4) Never Transfer
Firm Winner	-0.204***	-0.029	0.008*	0.221***
	(0.024)	(0.031)	(0.005)	(0.028)
Nearby Production Flag	0.011	0.017	0.003	-0.027
	(0.037)	(0.029)	(0.008)	(0.034)
Firm/Nearby Prod Interaction	-0.020	-0.076	0.019	0.055
	(0.053)	(0.061)	(0.018)	(0.068)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.137	0.041	0.052	0.147
E(Y_i | $F_i = 0, NearbyProd_i$=0)	0.347	0.427	0.004	0.216
p-value: $\beta_1 + \beta_3 = 0$	0.000	0.035	0.105	0.000
Observations	1692	1692	1692	1692

Table B.12: Regressions for restricted sample where there was exactly one firm among the three winners. The first column dependent variable is whether the lease was transferred within 0-1 years of the start. The second column dependent variable is whether it was transferred for the first time within 2-10 years. The third column is whether it was transferred for the first time more than 10 years after start date. And the fourth column is an indicator for whether it was never transferred. For each of these regressions, leases that do not appear in the LR2000 are omitted.

Next, in Table B.14, we provide robustness results related to how initial assignment affected the probability of assignment to a nearby producing firm. In contrast to our main regression results in Table 4, here we have a stricter definition of whether there is trade with a nearby producing firm: Table B.14 only includes trades prior to any drilling on a lease and before twelve years have elapsed. Results are similar to those in Table 4, though less precise. Columns (1) and (3) use the restricted sample, whereas Columns (2) and (4) use the full sample. Columns (3) and (4) further restrict the sample to those observations where trade has happened. For Columns (1) and

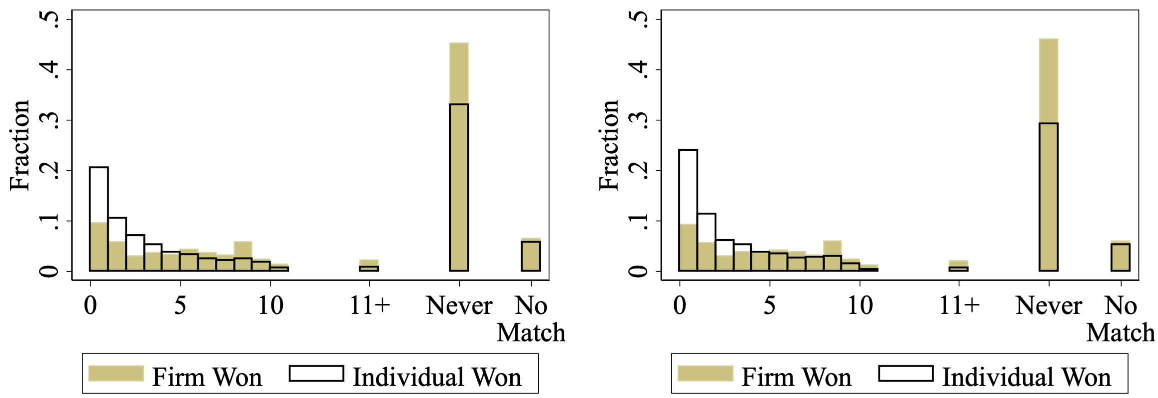
	(1)	(2)	(3)	(4)
	Reassign Probability		Log Time to Reassign	
Firm Winner (β_1)	-0.142***	-0.213***	0.603***	0.565***
	(0.028)	(0.025)	(0.080)	(0.065)
Nearby Production Flag (β_2)	-0.020	0.017	0.043	-0.015
	(0.042)	(0.014)	(0.104)	(0.052)
Firm/Nearby Prod Interaction (β_3)	0.001	-0.054	-0.145	-0.057
	(0.061)	(0.056)	(0.179)	(0.148)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.119	0.131	0.143	0.104
$E(Y_i \mid F_i = 0, NearbyProd_i{=}0)$	0.612	0.762	6.491	6.452
Observations	1800	10762	1022	6920

Table B.13: This table's dependent variables are the probability of reassignment *to a firm* within twelve years of the lease start date (Columns 1 & 2) and the log length of time until reassignment to a firm, given that it happens (Columns 3 & 4). Columns (1) & (3) use our restricted sample, while Columns (2) & (4) use the full sample. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on. Columns (3) and (4) do not correct for selection into reassignment.

(2), we find that for leases close to existing production, individuals are more likely to transfer their leases to a nearby producing firm, with the p-value for the test that $\beta_1 + \beta_3 = 0$ is equal to 0.051 in Column (1) and equal to 0.001 in Column (2). When restricting to cases where trade happened (Columns 3 and 4), our estimates of $\beta_1 + \beta_3$ are no longer statistically significant but are still negative, with p-values of 0.236 and 0.101 respectively.

Reassignment Before Drilling Analysis

In this section we explore the extent to which drilling happened without trade, especially for leases won by individuals, and find that it was rare. This suggests our



(a) Restricted Sample (b) Full Sample

Figure B.3: Histogram of the amount of time, in years, until a lease is first transferred to a firm. Panel (a) is limited to leases where exactly one firm appeared among the first-, second-, and third-place winners. Panel (b) displays the full sample. Individuals are in black outlines and firms are in green. Leases that do not have a recorded transfer are in the "Never" category, while leases that could not be linked to the LR2000 data are in the "No Match" category.

results are not driven by overoptimistic individuals drilling without knowing what they were doing.

In Table B.15, we examine drilling and production that happened prior to trade. Columns (1) and (2) use a dependent variable equal to one if the lease was drilled before there was trade, and zero otherwise. Columns (3) and (4) construct a dependent variable equal to one if the lease produced before there was trade, and zero otherwise. Overall, we find that it was rare for leases to be drilled or to produce prior to their being trade. Firms are more likely to do this than individuals, although the point estimates are not statistically significant.

Table B.16 limits the observations to cases where drilling happened, and examines the fraction of time that trade happened before drilling. Column (1) shows that for the restricted sample, conditional on drilling, 81% of leases won by individuals were traded prior to drilling. That this number rises to 88% if we examine "unsophisticated individuals" – those individuals who appear relatively infrequently

	(1)	(2)	(3)	(4)
	Trade with Nearby Producing Firm			
Firm Winner	-0.024	-0.003	-0.020	0.016
	(0.023)	(0.019)	(0.042)	(0.034)
Nearby Production Flag	0.011	0.022***	0.013	0.025**
	(0.023)	(0.008)	(0.030)	(0.010)
Firm/Nearby Prod Interaction	-0.027	-0.049*	-0.024	-0.059
	(0.035)	(0.025)	(0.056)	(0.044)
Full Sample	No	Yes	No	Yes
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.083	0.031	0.102	0.032
E($Y_i \mid F_i = 0, NearbyProd_i$=0)	0.070	0.059	0.092	0.073
p-value: $\beta_1 + \beta_3 = 0$	0.051	0.001	0.236	0.101
Observations	819	5043	550	3965

Table B.14: We examine whether leases are reassigned to firms that have nearby production, where we use a more restrictive definition of the dependent variable, only allowing the dependent variable to equal 1 if there was trade before any well was drilled and trade before twelve years had elapsed. Columns (1) and (3) uses our restricted sample and Columns (2) and (4) use the full sample. Columns (3) and (4) further restrict the sample to those where trade has occurred.

in the data.[10] In contrast, only 58% of leases won by firms that were drilled were traded prior to drilling.

For leases close to existing production where we find that leases won by individuals had a higher probability of drilling than leases won by firms, an alternative explanation of this is "overoptimism" on the part of individuals, where some individuals may have drilled because of overoptimism about lease productivity. Therefore, we next conduct a modified version of our drilling regressions where we adjust the dependent variable depending on whether there was trade. In particular, for leases won by individuals, we only count them as being drilled if they were also traded.

[10] We discuss our definition of sophisticated and unsophisticated individuals in Section B.5.

	Drilling	Drilling	Production	Production
Firm Winner	0.008	0.011	0.007	0.007
	(0.008)	(0.007)	(0.005)	(0.005)
Nearby Production Flag	0.039***	0.026***	0.018**	0.012***
	(0.013)	(0.005)	(0.008)	(0.004)
Firm/Nearby Prod Interaction	0.002	0.005	0.015	0.017
	(0.026)	(0.021)	(0.017)	(0.014)
intercept	0.010***	0.004	0.005**	-0.011***
	(0.003)	(0.011)	(0.002)	(0.002)
R squared	0.013	0.025	0.011	0.019
$E(Y_i \mid F_i = 0, NearbyProd_i=0)$	0.010	0.011	0.005	0.004
p-value: $\beta_1 + \beta_3 = 0$	0.677	0.388	0.204	0.086
Observations	1692	10191	1692	10191

Table B.15: This table examines the probability of drilling and production activity prior to initial reassignment. The dependent variable in Columns (1) and (2) is whether drilling ever happened on the lease prior to any trades. The dependent variable in Columns (3) and (4) is whether production ever happened on the lease prior to any trades. Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample. Regressions are limited to leases that can be matched to LR2000 data.

	Restricted Sample			Full Sample		
	All	No NP	NP	All	No NP	NP
	(1)	(2)	(3)	(4)	(5)	(6)
All	0.75	0.72	0.77	0.83	0.83	0.84
Won by an individual	0.81	0.73	0.86	0.84	0.84	0.85
Sophisticated	0.74	0.50	0.85	0.82	0.79	0.84
Unsophisticated	0.88	0.90	0.87	0.86	0.86	0.85
Won by a firm	0.58	0.70	0.43	0.58	0.67	0.46

Table B.16: Probability that trade happened before drilling – conditional on drilling happening. Columns (1)-(3) use the restricted sample; Columns (4)-(6) use the full sample. Columns (2) and (5) are limited to leases that are not close to existing production ("No NP" means "no nearby production"); Columns (3) and (6) are limited to leases that are close to existing production ("NP" means "nearby production"). Sophisticated versus unsophisticated is defined as in Section B.5 of the Online Appendix.

However, for leases won by firms, our dependent variable is whether they were drilled or not, regardless of trade. This modified dependent variable is designed to exclude the likely cases where individuals drilled overoptimistically.

Our regression results are in Table B.17. Columns (1) and (2) are run on the restricted and full samples, respectively. Columns (3) and (4) further restrict the dependent variable such that trade by an individual only counts if it was *to a firm*. Results are similar to those in Table 5: Even with this measure of drilling where we do not count some drilling in leases won by individuals, we find that for leases close to existing production, those won by individuals still had a higher probability of drilling than those won by firms.

	(1)	(2)	(3)	(4)
	Drilled (and transferred if individual)			
Firm Winner	0.018	0.017	0.022*	0.021**
	(0.013)	(0.011)	(0.013)	(0.011)
Nearby Production Flag	0.125***	0.099***	0.107***	0.086***
	(0.027)	(0.025)	(0.027)	(0.023)
Firm/Nearby Prod Interaction	-0.098***	-0.073***	-0.080**	-0.060**
	(0.032)	(0.026)	(0.032)	(0.025)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.112	0.076	0.107	0.072
$E(Y_i \mid F_i = 0, NearbyProd_i{=}0)$	0.043	0.054	0.039	0.050
p-value: $\beta_1 + \beta_3 = 0$	0.004	0.015	0.042	0.070
Observations	1800	10762	1800	10762

Table B.17: Regressions where when firms win, the dependent variable is whether there is drilling within 12 years of the lease start date; when individuals win, the dependent variable is whether there drilling and a transfer to a firm within 12 years of the start of the lease. Columns (3) and (4) further restrict the dependent variable for individuals to only equal 1 if there was drilling and a transfer *to a firm* within 12 years of the start of the lease. Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample.

B.4 Additional Drilling and Production Results

This section includes robustness and other results related to drilling and production outcomes. We first replicate drilling and production graphs from Figures 2 and 3, adding confidence intervals. Next, we examine our drilling and production results for leases close to nearby production and use a range of alternative definitions for nearby production. Third and fourth, we conduct robustness results for our drilling and production regressions where we incorporate county and township fixed effects and nearby producing well group fixed effects. Fifth, we examine other outcome variables related to drilling and production. Finally, we examine the correlation of trade outcomes with drilling and production outcomes, showing further evidence that results for leases close to nearby production are driven by the nearby producing firm.

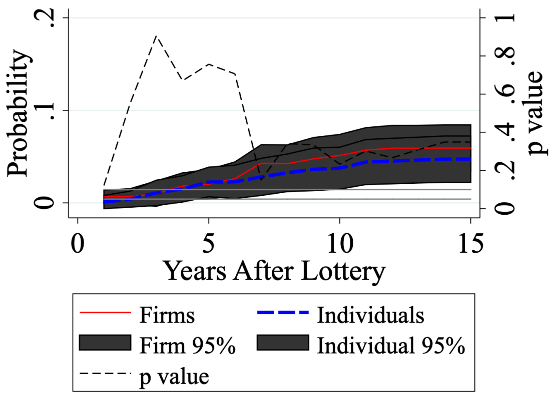
Graphs with confidence intervals

Figures B.4 and B.5 replicate Figures 2 and 3, respectively, adding in 95% confidence intervals around the point estimates. We include these confidence intervals to illustrate the precision of estimates, but not as a proxy for whether the treatment effect is statistically significant. Indeed, confidence intervals for two variables can overlap even when means of two variables are statistically different from each other (see, for example, https://www.cscu.cornell.edu/news/statnews/73_ci.pdf).

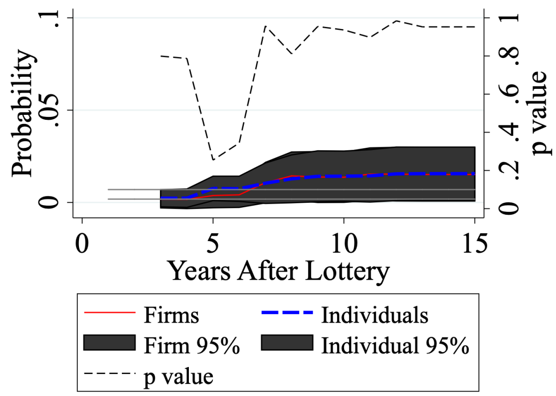
Alternative Nearby Production Specifications

We now examine drilling and production for leases close to existing production, varying the definition of "nearby production." Figure B.6 defines nearby production as occurring over the last 10 years rather than the last 5 years. Figure B.7 defines it as occurring within 1.7 miles rather than 2.6 miles; Figure B.8 defines it as occurring within 3.2 miles rather than 2.6 miles. For all three modifications, results are broadly consistent with our main specification in Figure 3.

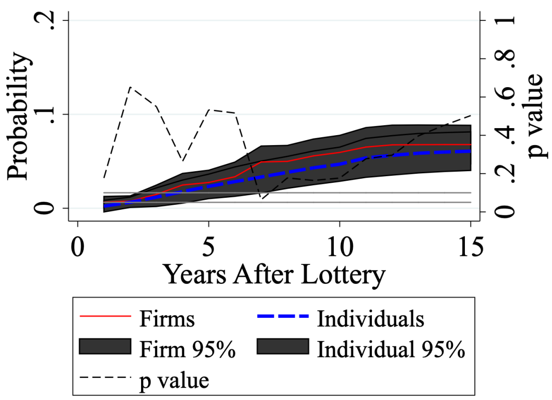
Drilling and Production Results: Leases without Nearby Production



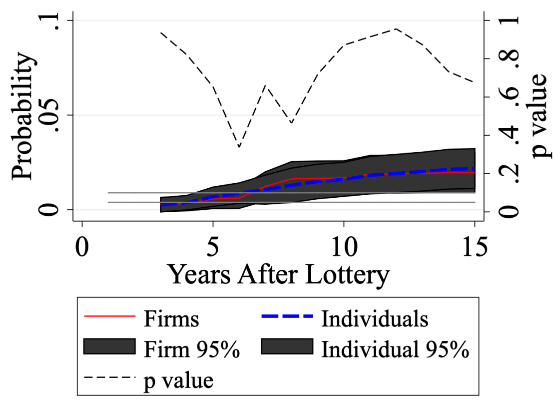
(a) Drilling



(b) Production



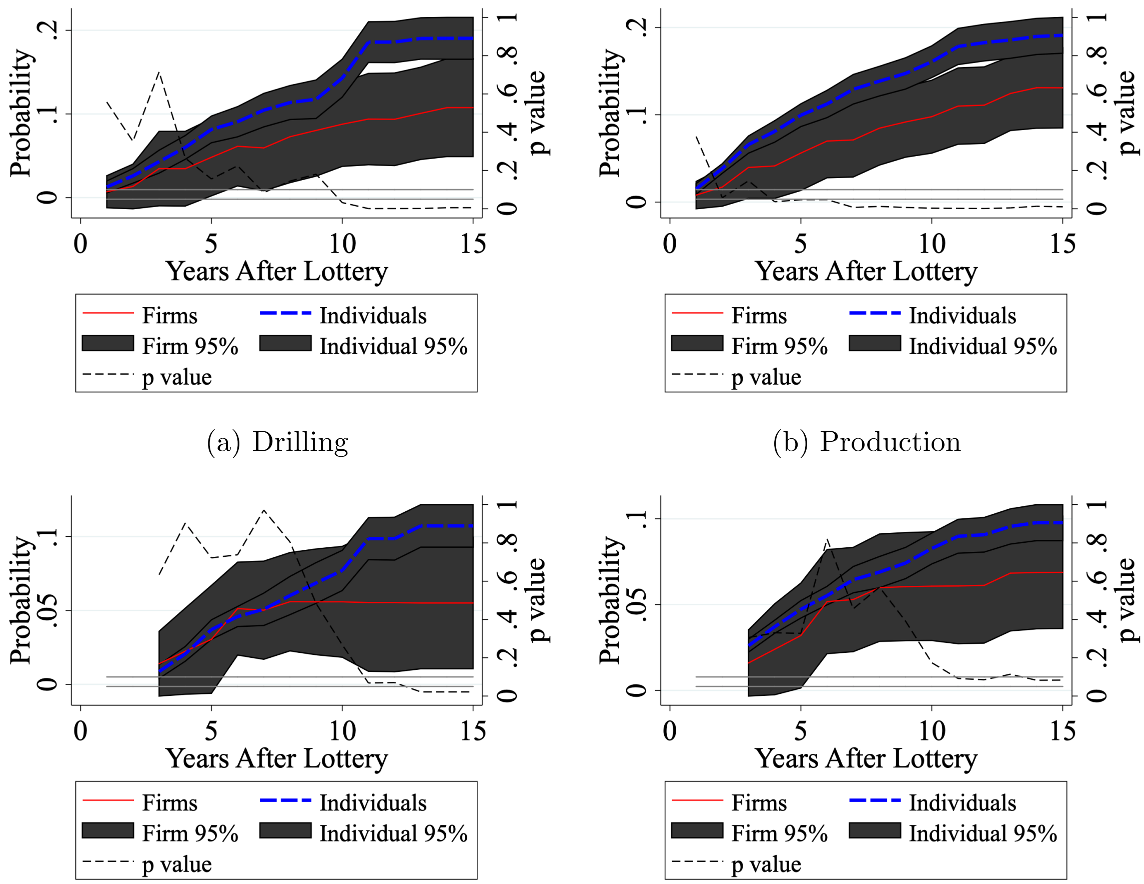
(c) Drilling



(d) Production

Figure B.4: Leases far from existing production, comparing those won by individuals and those won by firms. Panels (a) and (b) use the restricted sample with one firm winner. Panels (c) and (d) use the full sample and control for endogenous entry using observables. Probabilities for leases won by individuals are $E(Y_i|F_i = 0, NearbyProd_i = 0)$; predicted probabilities for leases won by firms are $E(Y_i|F_i = 0, NearbyProd_i = 0) + \beta_1$. The right vertical axis gives the p value of a test that the two means are not equal. Includes 95% confidence intervals of the point estimates.

Restricted Sample Results: Leases with Nearby Production



(a) Drilling

(b) Production

(c) Drilling

(d) Production

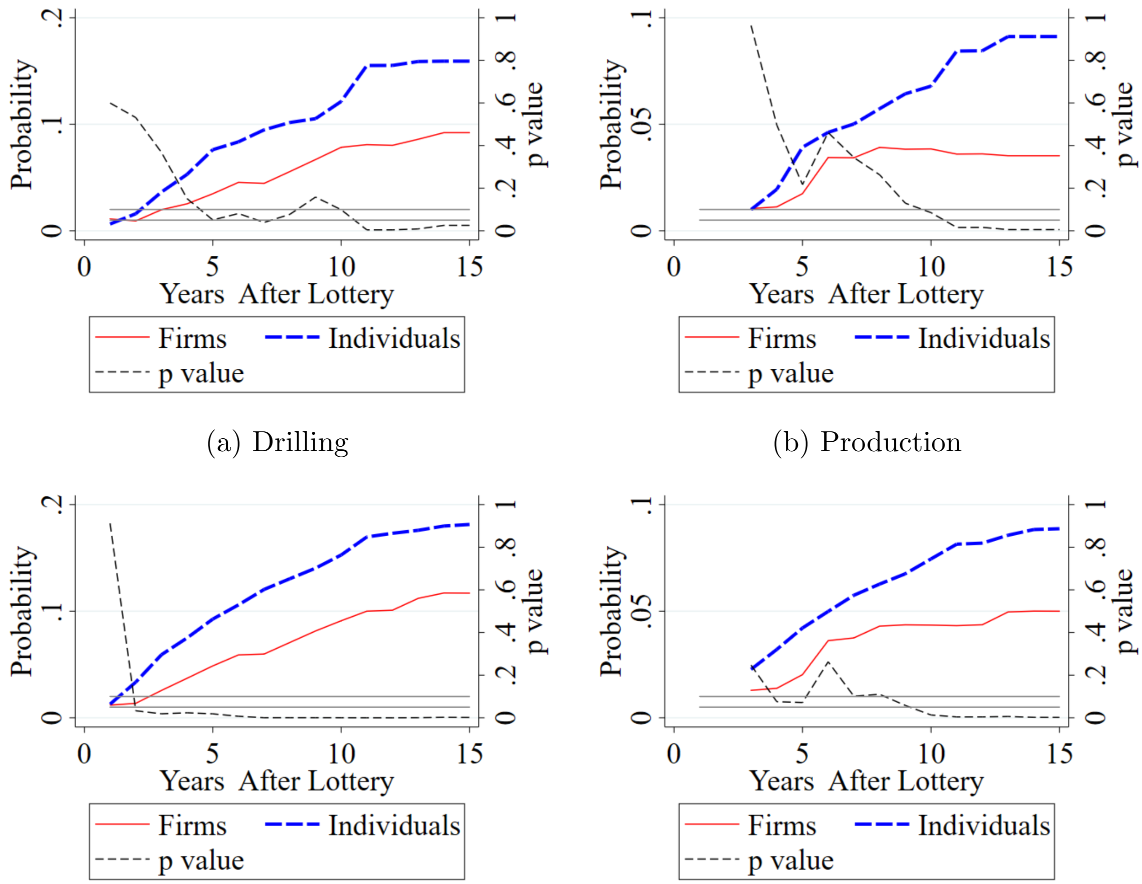
Figure B.5: Leases close to existing production, comparing those won by individuals and those won by firms. Panels (a) and (b) use the restricted sample with one firm winner. In those panels, probabilities for leases won by individuals are $E(Y_i|F_i = 0, NearbyProd_i = 0)$; predicted probabilities for leases won by firms are $E(Y_i|F_i = 0, NearbyProd_i = 0) + \hat{\beta}_1$. Panels (c) and (d) use the full sample and control for endogenous entry using observables. In those panels, probabilities for leases won by individuals are $E(Y_i|F_i = 0, NearbyProd_i = 0) + \hat{\beta}_2$; predicted probabilities for leases won by firms are $E(Y_i|F_i = 0, NearbyProd_i = 0) + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$. The right vertical axis gives the p value of a test that the two probabilities are not equal. Includes 95% confidence intervals of the point estimates.

Figure B.9 modifies the definition of nearby production to also include natural gas wells.[11] Natural gas wells were significantly less profitable than oil during this time period. Results are broadly similar, but with less stark differences between firms and individuals. For cases other than drilling in the restricted sample (panel a), the differences approach but do not reach conventional levels of statistical significance.

Figure B.10 examines cases with nearby drilling (rather than nearby production). Here, nearby is defined to only include the closest 1.7 miles of drilled wells because over half of leases had wells drilled within 2.6 miles. Results are much less stark and the differences are never statistically significant. Given that most wells, and especially exploratory wells, are not productive and that unproductive leases are more likely to be abandoned, proximity to a well is unlikely to be a good predictor of whether there is a nearby operating firm with private information about the underlying value of oil.

[11] We also modify the definition of recent to be defined by the first production year rather than the drilling year.

Longer Nearby Production Window

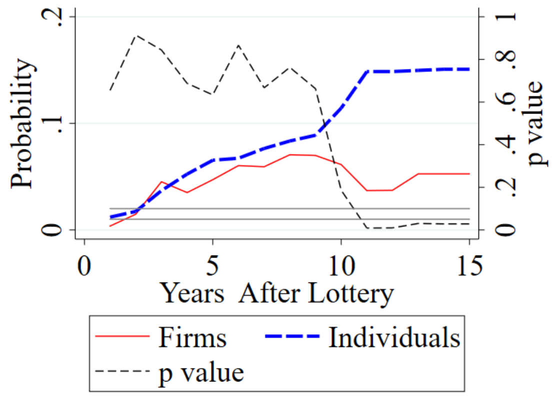


(a) Drilling (b) Production

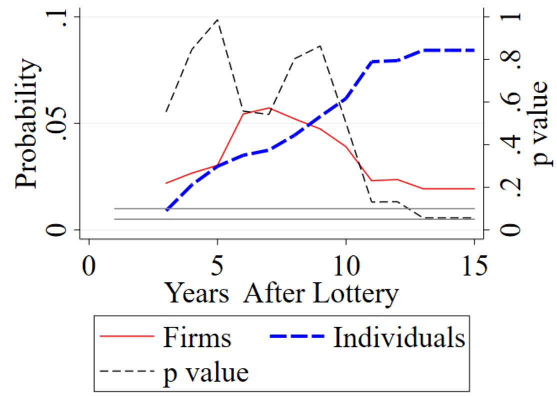
(c) Drilling (d) Production

Figure B.6: Robustness results where proximity to nearby production is defined as when the lease is within 2.6 miles of any productive well drilled in the last 10 (rather than 5) years. Panels (a) and (b) are the restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. Panels (c) and (d) use the full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal.

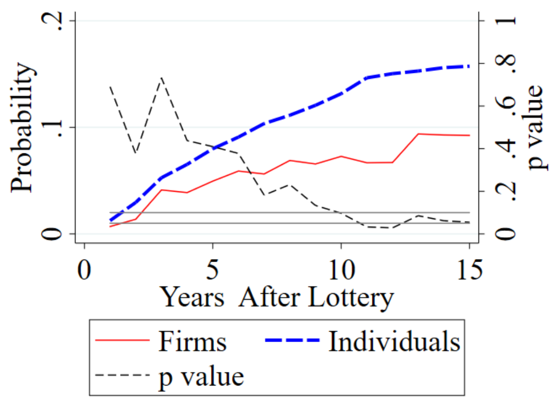
Tighter Nearby Definition



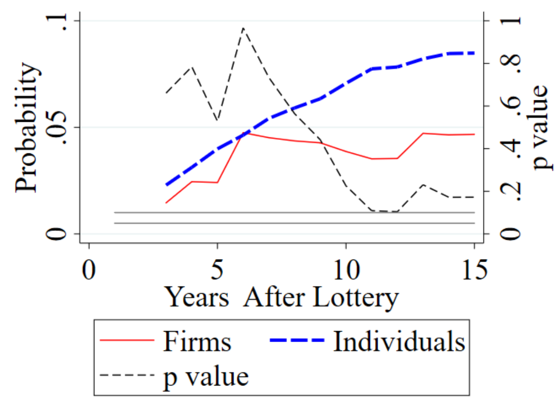
(a) Drilling



(b) Production



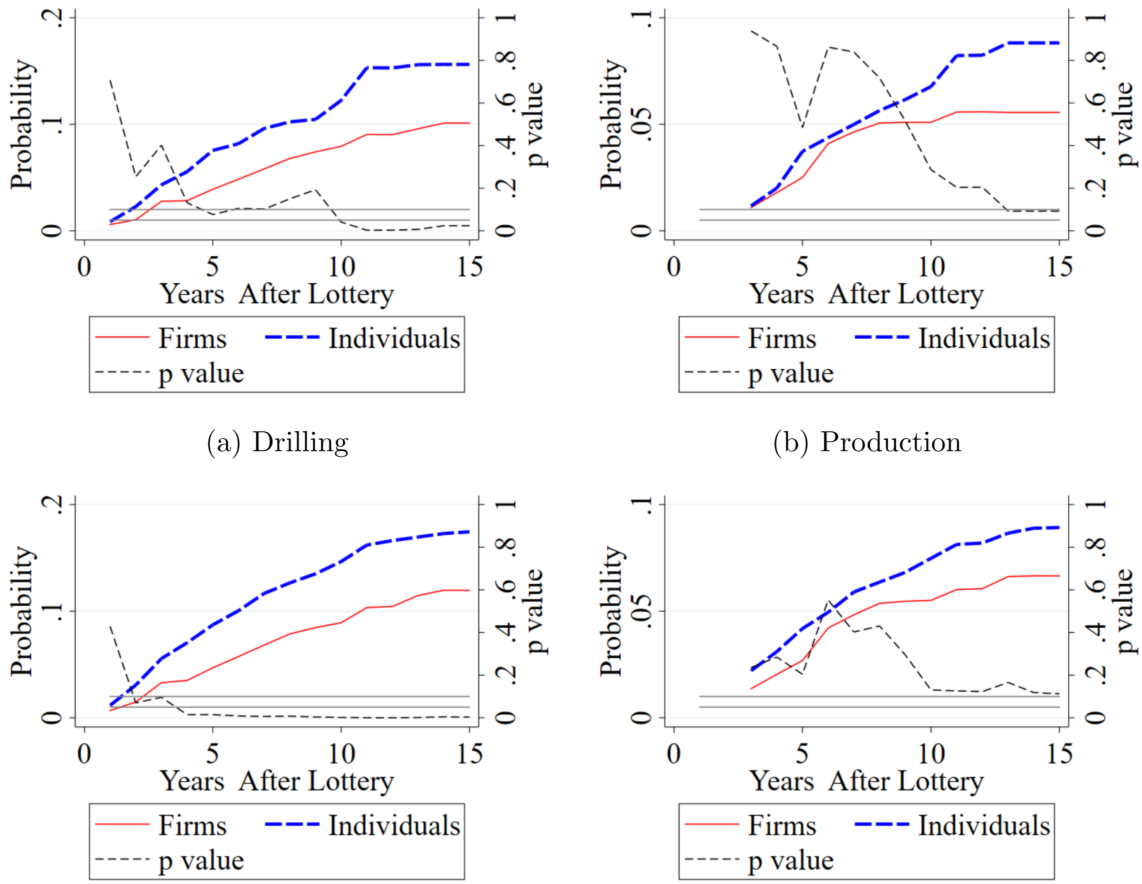
(c) Drilling



(d) Production

Figure B.7: Robustness results where proximity to nearby production is defined using the 1.7 mile cutoff (rather than the 2.6 mile cutoff). Panels (a) and (b) are the restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. Panels (c) and (d) use the full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal.

Looser Nearby Definition



(a) Drilling

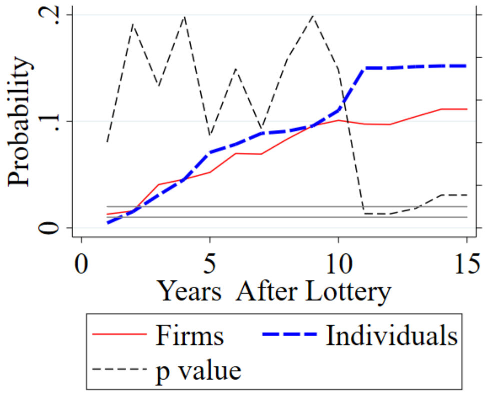
(b) Production

(c) Drilling

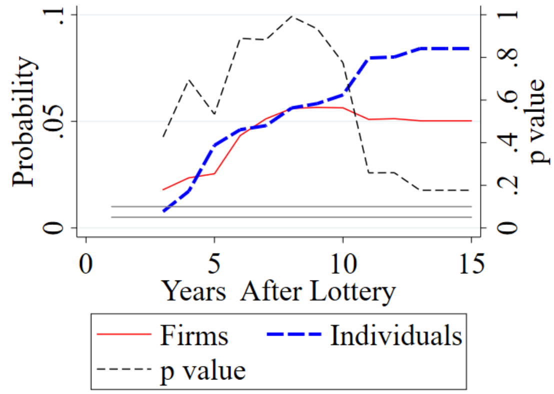
(d) Production

Figure B.8: Robustness results where proximity to nearby production is defined using the 3.2 mile cutoff (rather than the 2.6 mile cutoff). Panels (a) and (b) are the restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. Panels (c) and (d) use the full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal.

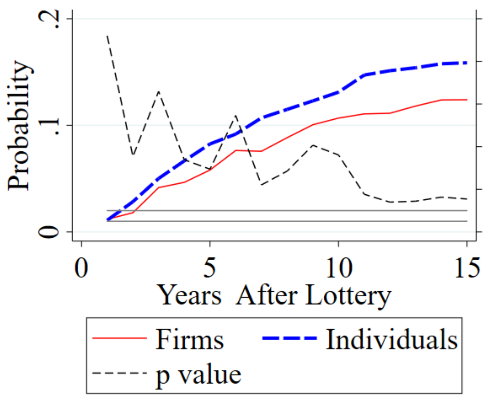
Including Gas Production



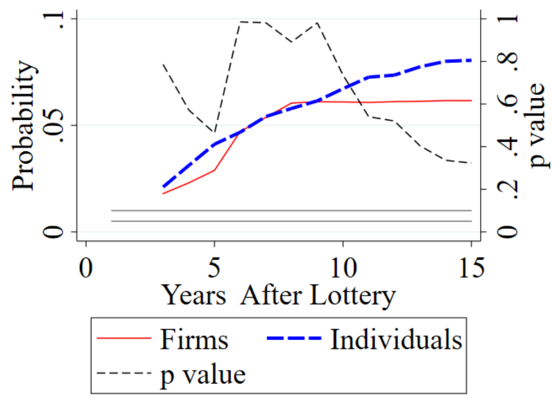
(a) Drilling



(b) Production



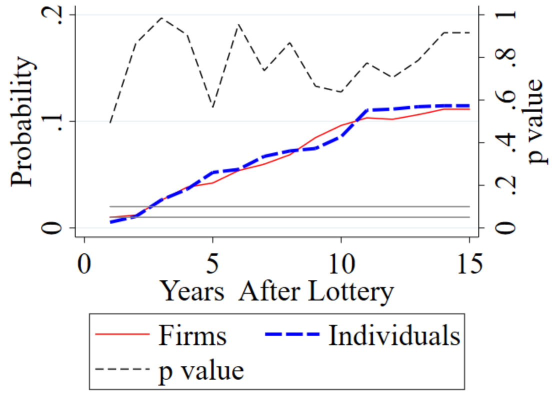
(c) Drilling



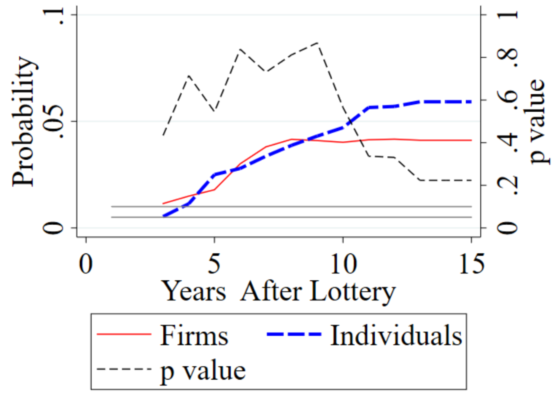
(d) Production

Figure B.9: Robustness results where proximity to nearby production also includes cases where nearby production was of natural gas. Here we define recent by the first production year rather than the drilling year. Panels (a) and (b) are the restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. Panels (c) and (d) use the full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal.

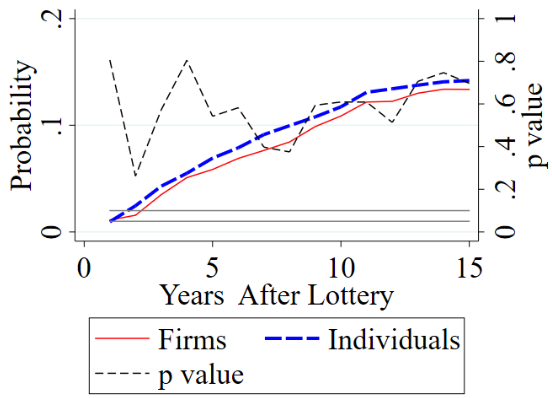
Nearby Drilling



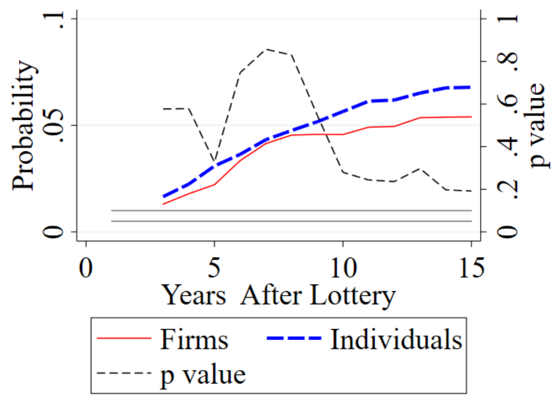
(a) Drilling



(b) Production



(c) Drilling



(d) Production

Figure B.10: Robustness results where we condition on there being previous drilling within 1.7 miles (rather than nearby production within 2.6 miles). Panels (a) and (b) are the restricted sample limited to cases where exactly one firm appeared among the first-, second-, and third-place winners. Panels (c) and (d) use the full sample, relying on controls to eliminate bias from endogenous entry. The firm effect is the coefficient on an indicator for whether the first place winner was a firm plus an interaction effect for nearby production and the firm effect. The right vertical axis gives the p value of a test that the two means are not equal. Panel (c) excludes month-of-lottery fixed effects due to the small sample size of this specification. This differs from our primary specification by looking at nearby drilling within 1.7 miles instead of nearby production within 2.6 miles.

Geographic Fixed Effects

This section includes geographic fixed effects. We include results that are analogous to our main results in Table 5. Table B.18 includes county-level fixed effects. In cases where a lease spans multiple counties, we include a separate county fixed effect for the multi-county combination. Our results with county fixed effects is very similar to our specifications without county fixed effects in Table 5.

Table B.19 includes township fixed effects.[12] Here, when a lease spans multiple townships, we randomly assign it to one of the townships in assigning fixed effects. Results are again very similar. Because in our restricted sample over 30% of townships in the sample include on lease, identification in that regression is driven by those townships with multiple leases.

[12] A township here is a contiguous set of sections, typically comprising a six mile by six mile square piece of land (for a total of 36 sections), although they are occasionally smaller due to the curvature of the earth. The word "township" can also be used in the PLSS as a measure of distance north or south of a baseline.

	(1)	(2)	(3)	(4)
	Drilling Probability		Production Probability	
Firm Winner	0.012	0.010	-0.001	-0.001
	(0.012)	(0.011)	(0.008)	(0.005)
Nearby Production Flag	0.115***	0.075***	0.063***	0.044***
	(0.025)	(0.018)	(0.019)	(0.011)
Firm/Nearby Prod Interaction	-0.096***	-0.068***	-0.034	-0.024
	(0.033)	(0.026)	(0.023)	(0.019)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.154	0.122	0.131	0.086
E($Y_i \mid F_i = 0, NearbyProd_i$=0)	0.051	0.068	0.019	0.025
p-value: $\beta_1 + \beta_3 = 0$	0.006	0.012	0.133	0.184
Observations	1800	10762	1800	10762

Table B.18: This table's dependent variables are the probability of drilling by twelve years (Columns 1 & 2), and the probability of production by twelve years (Columns 3 & 4). Columns (1) & (3) use our restricted sample, while Columns (2) & (4) use the full sample. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

	(1)	(2)	(3)	(4)
	Drilling Probability		Production Probability	
Firm Winner	0.006	0.013	0.007	0.001
	(0.023)	(0.011)	(0.018)	(0.006)
Nearby Production Flag	0.131^{**}	0.038^{***}	0.080^{*}	0.018^{**}
	(0.061)	(0.012)	(0.045)	(0.008)
Firm/Nearby Prod Interaction	-0.165^{*}	-0.068^{**}	-0.113^{*}	-0.018
	(0.086)	(0.030)	(0.068)	(0.020)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
Township Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.665	0.336	0.626	0.328
$E(Y_i \mid F_i = 0, NearbyProd_i{=}0)$	0.055	0.076	0.019	0.031
p-value: $\beta_1 + \beta_3 = 0$	0.037	0.044	0.080	0.361
Observations	1800	10762	1800	10762

Table B.19: This table's dependent variables are the probability of drilling by twelve years (Columns 1 & 2), and the probability of production by twelve years (Columns 3 & 4). Columns (1) & (3) use our restricted sample, while Columns (2) & (4) use the full sample. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

Nearby Producing Well Group Fixed Effects and Distance Bin Analysis

In this section, we incorporate nearby producing well group fixed effects. To define these nearby producing well group fixed effects, we first group together producing wells that are all within the same PLSS section into the same well group, and construct the groups to be PLSS sections. In cases where there are two or more nearby producing well groups which are tied for being the closest producing well group to a given lease, we only include a fixed effect for the most commonly appearing well group – e.g., the producing well group which appears most frequently as a nearby producing well group for the leases in the sample. This helps to reduce the number of fixed effects that need to be estimated. Even then, we still have a large number of nearby producing well group fixed effects to estimate: 461 in the restricted sample, and 1,046 in the full sample.

Tables B.20 and B.21 include our regression estimates where we limit our analysis to those leases that are within 5.2 miles of a nearby well group where we can reliably identify reasonable nearby producing wells. Columns (2) and (4) include the well group fixed effects, where we add fixed effects for the closest producing well group. We find similar estimates as when we do not include well group fixed effects, and in particular find that the estimate of $\beta_1 + \beta_3$ is negative and statistically significant.

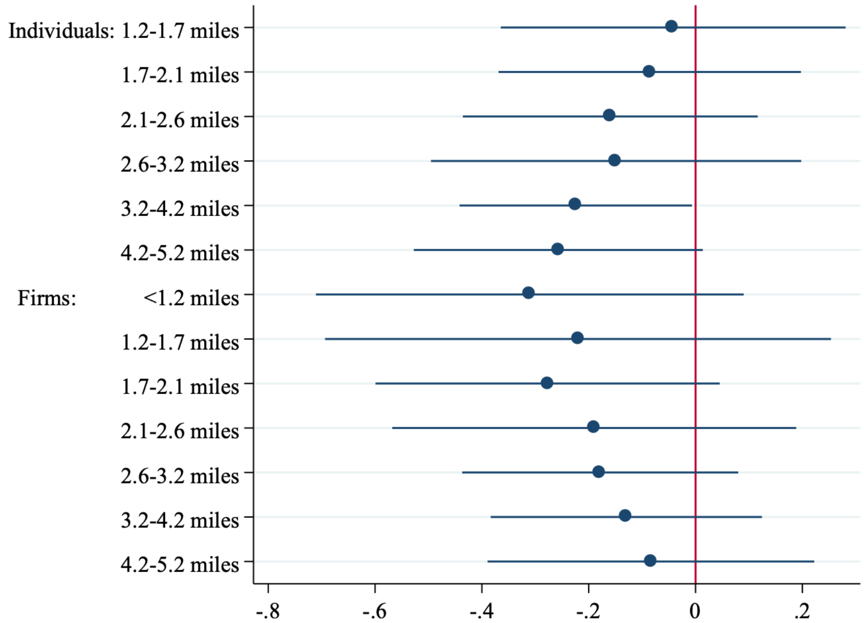
We also perform regressions with drilling and production outcomes where we allow for more granular distance bins, interacting each distance bin with whether the lease was won by a firm. Graphs of the regression estimates using the restricted sample for the various binned distance variables and interactions are in Figure B.11. The graph shows that greater proximity to nearby producing wells is associated with higher probability of drilling when the winner is an individual, but associated with lower probability of drilling when the winner is a firm. However, the large standard errors make it clear that the variation in the data is typically not rich enough to support analysis with such granular geographic bins.

	(1) Drill	(2) Drill	(3) Drill	(4) Drill
Firm Winner	0.052*	0.083	0.032	0.030
	(0.028)	(0.070)	(0.027)	(0.027)
Nearby Production Flag	0.140***	0.138*	0.082***	0.041***
	(0.023)	(0.075)	(0.022)	(0.013)
Firm/Nearby Prod Interaction	-0.146***	-0.258**	-0.099**	-0.073*
	(0.040)	(0.127)	(0.039)	(0.039)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Nearby production group FE	No	Yes	No	Yes
R squared	0.154	0.659	0.083	0.401
E($Y_i \mid F_i = 0, NearbyProd_i$=0)	0.052	0.061	0.095	0.114
p-value: $\beta_1 + \beta_3 = 0$	0.001	0.045	0.004	0.105
Observations	819	819	5043	5043

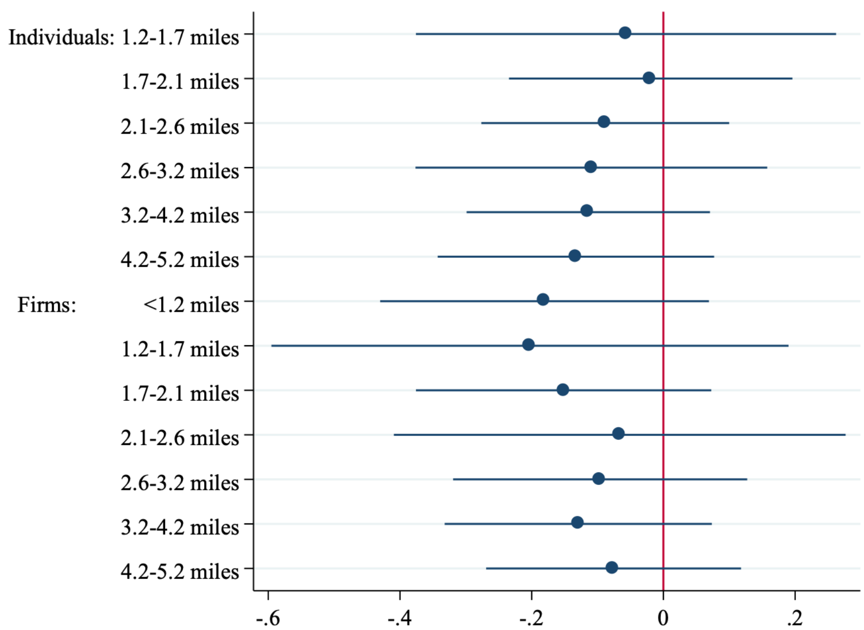
Table B.20: This table's dependent variables are the probability of drilling within twelve years of the start of the lease. Columns (1) and (2) use the restricted sample, while Columns (3) and (4) use the full sample. Columns (2) and (4) include nearby well group fixed effects. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

	(1) Prod	(2) Prod	(3) Prod	(4) Prod
Firm Winner	-0.000	0.019	-0.001	0.002
	(0.017)	(0.039)	(0.013)	(0.017)
Nearby Production Flag	0.073***	0.078	0.042***	0.020***
	(0.018)	(0.053)	(0.010)	(0.007)
Firm/Nearby Prod Interaction	-0.047*	-0.123	-0.024	-0.014
	(0.025)	(0.088)	(0.022)	(0.024)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Nearby production group FE	No	Yes	No	Yes
R squared	0.118	0.669	0.063	0.402
$E(Y_i \mid F_i = 0, NearbyProd_i{=}0)$	0.027	0.031	0.043	0.053
p-value: $\beta_1 + \beta_3 = 0$	0.043	0.132	0.142	0.419
Observations	819	819	5043	5043

Table B.21: This table's dependent variables are the probability of production within twelve years of the start of the lease. Columns (1) and (2) use the restricted sample, while Columns (3) and (4) use the full sample. Columns (2) and (4) include nearby well group fixed effects. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.



(a) Drilling



(b) Production

Figure B.11: Coefficient estimates and 95% confidence intervals graph for binned distance interacted with individual as well as binned distance interacted with firm. Dependent variable in (a) is whether there is drilling within 12 years; dependent variable in (b) is whether there is production within 12 years. The excluded group is individuals for where the nearby well group was less than 1.2 miles away. Regression specification also controls for acreage, polynomial of entries, and month-of-lottery fixed effects.

Drilling duration and production quantity results

Table B.22 shows regression results for two additional dependent variables: Time spent drilling and oil production over the first three years of a well's life. Because many drilling rigs are rented out by the day and workers are frequently paid by the hour, the number of days spent drilling is a rough proxy for drilling costs. Production quantities are a rough proxy for revenue. These regressions have small samples because they are conditional on drilling and production, respectively.

Columns (1) and (2) examine drilling time. We find that of the leases that are drilled, drilling durations for leases initially won by firms are statistically indistinguishable from those initially won by individuals. This holds for both leases that are close to and far from nearby production. Columns (3) and (4) examine production quantities. We find that of the leases that produce oil, those initially won by firms have production quantities that are statistically indistinguishable from those initially won by individuals. Again, this holds both for leases that are close to and far from nearby production.

	(1) Drill Time (Ln Days)	(2) Drill Time (Ln Days)	(3) Oil Prod. First 36 Months (Ln Barrels)	(4) Oil Prod. First 36 Months (Ln Barrels)
Firm Winner	-0.287	-0.060	-1.345	-1.080
	(0.251)	(0.231)	(1.391)	(0.834)
Nearby Production Flag	-0.207	-0.006	-0.244	0.274
	(0.159)	(0.088)	(0.810)	(0.306)
Firm/Nearby Prod Interaction	0.051	0.009	1.174	0.405
	(0.323)	(0.355)	(1.404)	(1.064)
Full Sample	No	Yes	No	Yes
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	No	Yes	No	Yes
R squared	0.089	0.113	0.195	0.237
E($Y_i \mid F_i = 0, NearbyProd_i$=0)	3.701	3.382	9.121	9.057
p-value: $\beta_1 + \beta_3 = 0$	0.179	0.762	0.818	0.151
Observations	156	1241	91	635

Table B.22: Columns (1) and (2) look at the length of time it takes to drill an oil well. Durations longer than 180 days or less than 1 day are excluded because they indicate a data error or that there was likely a stoppage in the middle of drilling. Columns (3) and (4) look at oil production on productive wells. Columns (1) and (3) use our restricted sample with exactly one firm winner. Columns (2) and (4) use our full sample, relying on our controls to correct for endogenous entry. All specifications look at wells drilled within twelve years of the lease date.

Correlations between trade outcomes and drilling and production outcomes

This section examines the links between trade and the asset utilization outcomes – drilling and production. In particular, we evaluate the correlation between whether trade happens and whether drilling and production happen. This section serves as auxiliary evidence that the nearby producing firm plays an important role in driving our empirical results.

In Table B.23 we examine how different measures of trade correlate with drilling and production outcomes. Here we regress drilling (Columns 1 through 4) or production outcomes (Columns 5 through 8) on whether there is trade, whether the lease is close to existing production, and the interaction of the two. We examine how the correlation between trade and asset utilization outcomes (drilling and production) depends on the measure of trade and whether the lease is close to existing production. We use three measures of trade: Any trade, trade with a firm, and trade with the nearby producing firm. We limit the sample to those observations within 5.2 miles of existing production (such that we can reasonably identify nearby producing firms).

Table B.23 demonstrates that trade and drilling outcomes and trade and production outcomes tend to be more strongly correlated as we shift from more general trade measures (any trade) to more specific measures (trade with a nearby producing firm). This pattern is especially strong for leases close to existing production. When we run a "horse race" using all three measures (Columns 4 and 8), we find that trade with a nearby producing firm most strongly predicts drilling and production, with trade to other firms only having a small predictive power that is only marginally significant in the drilling regression (Column 4). These results are suggestive that trade with the nearby producing firm is an important determinant of whether there is drilling and production, and is especially important for leases close to existing production.

In Table B.24 we regress asset utilization outcomes (drilling and production) on three variables and their interactions: Whether the lease was won by a firm, whether the lease was traded to the nearby producing firm, and whether the lease is within 2.6 miles of existing production. We use the following regression specification:

$$
\begin{aligned}
Y_i = \beta_1 F_i + \beta_2 TradeNPF_i + \beta_3 F_i * TradeNPF_i + \\
\beta_4 NearbyProd_i + \beta_5 F_i * NearbyProd_i + X_i \Omega + \varepsilon_i \quad (43)
\end{aligned}
$$

We find that trade with a nearby producing firm is positively correlated with drilling and production outcomes when the winner is an individual (β_3 is positive and usually statistically significant). However, we find that for firms, trade with the nearby producing firm is not as strongly correlated with drilling and production outcomes (estimates of β_3 negative, usually statistically significant, and of the same magnitude as β_2). This is suggestive that for individuals, trading with the nearby producing firm tends to serve to reallocate productive leases to the nearby producing firm. In contrast, for firm winners, trade with the nearby producing firm is much less able to predict the drilling and production outcomes.

In Table B.25, we examine log production amounts, which are only defined if the lease had a producing well.[13] Because the number of cases with production are small, we examine all leases within 5.2 miles of nearby production, and use the full sample results. In Column (1), we include as a covariate whether the lease is traded to the nearby producing firm, comparing production of leases that were traded with those that were not. There we find that trade with the nearby producing firm is associated with a 101.2 log point increase in production. In other words, leases that are traded to a nearby producing firm tend to be about twice as productive as those

[13] Here the dependent variable is the log of the first 36 months of production. Because wells typically follow an approximately exponential decline, production within the first 36 months will be approximately proportional to total production.

that are not, conditional on production. Given that the probability of drilling is higher for leases that are traded to a nearby producing firm than those that are not, and leases that are not drilled will have lower productivity on average than those that are, this estimate of 100% probably underestimates the true differences in productivity θ for leases that are traded to a nearby producing firm and those that are not.

Column (2) of Table B.25 examines how productivity of leases traded to the NPF differs depending on whether the lease was initially won by an individual or a firm. There we find that conditional on trade to the nearby producing firm, those leases initially won by a firm have higher productivity than those won by individuals. Although the small sample and overall noisiness of production implies that the point estimate is not statistically significant, it is large, with leases won by firms having a 27.4 log point higher production than leases won by individuals. This is consistent with the intuition of the informed buyer model in Section 3: When the winner is a firm (and has lower drilling costs), trade will only happen in the cases where the true value of θ was especially high, which translates into higher average productivity conditional on trade to the nearby producing firm.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Drilling				Production			
Traded	0.016			-0.015	0.005			-0.008
	(0.012)			(0.020)	(0.011)			(0.013)
Traded to firm		0.027***		0.032*		0.010		0.010
		(0.010)		(0.017)		(0.009)		(0.010)
Traded to NPF			0.091**	0.088**			0.098***	0.098***
			(0.039)	(0.039)			(0.033)	(0.032)
Nearby production flag	0.036**	0.046***	0.064***	0.035**	0.023*	0.021**	0.030***	0.022*
	(0.016)	(0.015)	(0.018)	(0.016)	(0.013)	(0.010)	(0.006)	(0.013)
Traded x Nearby production flag	0.051*			0.018	0.022			-0.017
	(0.027)			(0.023)	(0.019)			(0.016)
Traded to firm x Nearby production flag		0.045**		0.024		0.030		0.032*
		(0.022)		(0.017)		(0.019)		(0.017)
Traded to NPF x Nearby production flag			0.089	0.082			0.077*	0.073
			(0.055)	(0.053)			(0.046)	(0.045)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.084	0.086	0.097	0.100	0.063	0.065	0.092	0.094
E(Y | No Trade, Nearby Production = 0)	0.084	0.079	0.090	0.104	0.038	0.035	0.035	0.039
Observations	5043	5043	5043	5043	5043	5043	5043	5043

Table B.23: Dependent variables are probability of drilling in Columns (1) through (4) and probability of production in Columns (5) through (8). The sample is the full sample but limited to observations within 5.2 miles of existing production.

	(1)	(2)	(3)	(4)
	Drilling		Production	
Firm Winner	0.059**	0.042	0.007	0.009
	(0.027)	(0.026)	(0.016)	(0.012)
Traded to NPF	0.093**	0.146***	0.080	0.148***
	(0.046)	(0.034)	(0.050)	(0.034)
Firm/Traded to NPF Interaction	-0.094	-0.170**	-0.110**	-0.183***
	(0.107)	(0.081)	(0.052)	(0.035)
Nearby Production Flag	0.137***	0.077***	0.070***	0.037***
	(0.022)	(0.021)	(0.018)	(0.009)
Firm/Nearby Prod Interaction	-0.143***	-0.092**	-0.045*	-0.018
	(0.039)	(0.038)	(0.025)	(0.021)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.158	0.097	0.124	0.092
Observations	819	5043	819	5043

Table B.24: Dependent variables are probability of drilling in Columns (1) and (2) and probability of production in Columns (3) and (4). Sample is limited to those leases within 5.2 miles of existing production. Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample.

	(1)	(2)
	Log first 36 months oil production	
Trade with NPF	1.012***	1.005***
	(0.294)	(0.294)
Trade with NPF x Firm Winner		0.274
		(0.516)
Nearby production flag	0.611**	0.615**
	(0.288)	(0.292)
Number of Entries & Acreage Controls	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes
R squared	0.166	0.166
E(Y| Not traded to NPF)	8.351	8.353
Observations	579	579

Table B.25: Regressions where the dependent variable is log first 36 months of oil production, conditional on production. Sample limited to leases within 5.2 miles of existing production and that which had production.

B.5 Different Entrant Types

This section discusses evidence related to potential heterogeneity among individuals and potential heterogeneity among firms – e.g. individuals with more versus less industry experience, and major versus minor oil firms. We first discuss how we measure these types. We then examine whether individual or firm type is correlated with whether a lease is close to nearby production. Third, we examine the extent to which trade, drilling, and production outcomes vary depending on the firm's or individual's type. Finally, we do an exercise to explore the extent to which for leases close to existing production, our drilling and production differentials are explained by type composition. We do not find evidence that different entrant types explain our results.

Defining types of individuals and firms

Our lottery winner data gives very limited information about the identity of the first-, second-, and third-place winners. Typically, only the name is recorded. In the case of first-place winners, the address is also reported. To examine variation in different types of individuals or firms, we use this limited information to categorize individuals and winners into different types.

For individuals, our first approach is to examine the frequency with which any given individual appears in our data. Individuals that appear 14 or more times in our data are categorized as "sophisticated"; individuals that appear 13 times or fewer are categorized as "unsophisticated". This cutoff is chosen to maximize statistical power: Of the leases in our restricted sample that were won by an individual, approximately half of them were won by a "sophisticated" individual and the other half were won by an "unsophisticated" individual.

Our second approach for individuals is to use information on the address of the individual. In particular, we examine whether any given first-place individual's address is similar to the address for a firm who appears at any time in our data as a first-place winner. This approach is designed especially for cases where individuals may have submitted entries in behalf of firms. To examine similarity of matches, we use a bigram method to construct a score of how similar a firm's address and an individual's address are, conditional on them appearing in the same zip code.[14] Figure B.12 shows a smoothed kernel estimate of the maximum fuzzy score for an individual over all firms in the sample that share the same zip code. We find that the distribution is bimodal, with a score of approximately 0.72 separating the two modal humps, and a score of approximately 0.85 at the peak of the right-side hump. Therefore, we construct two definitions of whether an individual shares the address of a firm, one using 0.72 as the cutoff, and the other using 0.85 as the cutoff.

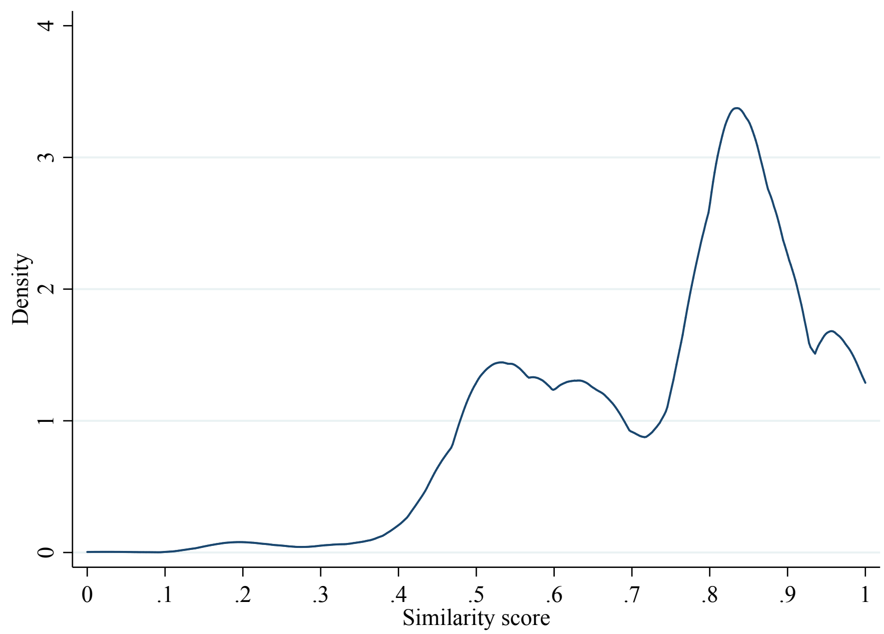


Figure B.12: Maximum fuzzy name matching similarity scores for individuals taken over all firms in the same zip code.

For firms, we also take two approaches. The first approach is similar to that which we use to define sophisticated and unsophisticated individuals, and uses the

[14]We use the "matchit" package for Stata.

total number of times that a firm appears. Firms that appear 19 or more times in our data are categorized as "major" firms; firms that appear 18 times or less are categorized as "minor" firms. This cutoff is chosen to maximize power within our restricted sample: Of the leases in our restricted sample that were won by a firm, approximately half of them were won by a major firm and half were won by a minor firm.

Our second approach for firms involves examining whether the firm is one of the major oil producing firms of the 1970s known as the "seven sisters". The seven sisters consisted of British Petroleum, Shell, Chevron, Gulf Oil, Texaco, Exxon, and Mobil. Overall, we find very few appearances of seven sister firms in our data. Within our entire sample, we find that only 46 of the 2,047 entries with firm names are those of seven sisters. There are 22 appearances of Gulf Oil, 15 appearances of Shell, 6 appearances of Texaco, 2 of Amoco, and 1 of Exxon.

Correlations between type and nearby production

Is there correlation between whether a lease is close to existing production and the type of individual or firm that appears as a first-, second-, or third-place winner? If so, then an alternative explanation for why leases close to versus far from existing production differ in their firm-individual drilling differentials might be differences in the composition of individual and/or firm types.

We first examine sophisticated versus unsophisticated individuals. We limit the sample to those observations of first-, second-, and third-place entrees who are individuals. The dependent variable is whether an individual is a "sophisticated" individual as defined above. Dependent variables include whether the lease is close to existing production, log acreage, month of lottery fixed effects, and a polynomial of offers, using the following regression specification, where S_i is an indicator for whether the individual is a sophisticated individual:

$$S_i = \alpha_0^S + \alpha_1^S NearbyProd_i + \Omega X_i + \varepsilon \tag{44}$$

In Table B.26 we display our results. We find that proximity to nearby production is inversely correlated with the probability that a sophisticated individual appears in the data, showing that unsophisticated individuals are more likely to crowd out sophisticated individuals for leases close to existing production. The coefficient on nearby production is statistically significant in all specifications, both with and without additional controls, as well as both for the restricted sample and the full sample.

Tables B.27 and B.28 take a similar approach, but instead examine whether control variables predict whether an entrant has an address similar to that of a firm. We use the 0.72 threshold definition of similar address in Table B.27 and the 0.85 threshold definition of similar address in Table B.28. Because we only have addresses for first-place winners, these regressions are limited to observations of first-place winners where the first-place winner was an individual. Here we find that whether an individual has a similar address to a firm is not statistically significantly correlated with proximity to nearby production. While Column (3) of Table B.27 shows that in the full sample, proximity to nearby production is negatively correlated with whether an individual has a similar address to a firm, the correlation is no longer statistically significant when controlling for the other suite of controls.

We now turn to examining the distribution of firm types. Tables B.29 and B.30 examine similar regressions, but instead limit the observations to the first-, second-, or third-place winners in the data that are firms. Table B.29 reports results where we examine whether a firm is a "major" firm, as defined above, on a suite of controls, including whether the lease is close to existing production, using the following regression specification, where Maj_i is an indicator for whether the firm is a major firm:

$$Maj_i = \alpha_0^{Maj} + \alpha_1^{Maj} NearbyProd_i + \Omega X_i + \varepsilon \tag{45}$$

In Table B.30, we use a very similar approach, but regress whether a firm is a "Seven Sisters" supermajor on a suite of controls. We find that proximity to nearby production is not statistically significantly correlated with whether the firm is a "major" firm and is not statistically significantly correlated with whether the firm is a supermajor.

Table B.31 displays the first word of corporate winners' names for leases in our restricted sample. The second column shows frequencies for all winners who won at least 5 leases with nearby production, representing roughly half of the 376 leases without nearby production. Column (3) shows all winners who won at least 14 leases with nearby production, representing roughly half of the 1,424 leases with nearby production. The two lists greatly overlap, suggesting that firms that won leases with nearby production are similar to firms that won leases without nearby production.

With the exception of sophisticated versus unsophisticated individuals in Table B.26, the remaining tables show no statistically significant correlation between type of entrant and proximity to nearby production. However, the point estimates are almost all negative. This is weak evidence for leases close to existing production, those individuals or firms who have less industry experience and/or fewer industry ties are slightly more common.

	(1)	(2)	(3)	(4)
		Sophisticated		
Nearby Production Flag	-0.077***	-0.063***	-0.057***	-0.024***
	(0.020)	(0.020)	(0.007)	(0.007)
Constant	0.527***	0.763***	0.471***	0.687***
	(0.009)	(0.053)	(0.003)	(0.023)
Number of Entries & Acreage Controls	No	Yes	No	Yes
Month of Lottery Fixed Effects	No	Yes	No	Yes
R squared	0.004	0.145	0.002	0.141
Observations	3600	3600	30239	30239

Table B.26: We examine the probability that an individual appearing as a first-, second-, or third-place winner is a "sophisticated" individual. The sample is limited to entrants in the data that are individuals. Columns (1) and (2) use the restricted sample; Columns (3) and (4) use the full sample. Columns (1) and (3) use no controls other than proximity to nearby production; Columns (2) and (4) also control for number of entries, acreage, and month-of-lottery fixed effects. Regressions use conventional OLS standard errors.

	(1)	(2)	(3)	(4)
	Similar address as firm (0.72)			
Nearby Production Flag	0.011	0.006	-0.020*	-0.005
	(0.033)	(0.034)	(0.011)	(0.011)
Constant	0.273***	0.359***	0.265***	0.291***
	(0.015)	(0.093)	(0.005)	(0.037)
Number of Entries & Acreage Controls	No	Yes	No	Yes
Month of Lottery Fixed Effects	No	Yes	No	Yes
R squared	0.000	0.073	0.000	0.014
Observations	1181	1181	10065	10065

Table B.27: We examine the probability that an individual appearing as a first-place winner has a similar address to a firm, where we use the 0.72 threshold for address similarity as discussed above. The sample is limited to entrants in the data that are individuals. Columns (1) and (2) use the restricted sample; Columns (3) and 4 use the full sample. Columns (1) and (3) use no controls other than proximity to nearby production; Columns (2) and (4) also control for number of entries, acreage, and month-of-lottery fixed effects. Regressions use conventional OLS standard errors.

	(1)	(2)	(3)	(4)
	Similar address as firm (0.85)			
Nearby Production Flag	0.007	-0.001	-0.004	0.004
	(0.026)	(0.027)	(0.009)	(0.009)
Constant	0.146***	0.225***	0.150***	0.097***
	(0.012)	(0.074)	(0.004)	(0.030)
Number of Entries & Acreage Controls	No	Yes	No	Yes
Month of Lottery Fixed Effects	No	Yes	No	Yes
R squared	0.000	0.057	0.000	0.011
Observations	1181	1181	10065	10065

Table B.28: We examine the probability that an individual appearing as a first-place winner has a similar address to a firm, where we use the 0.85 threshold for address similarity as discussed above. The sample is limited to entrants in the data that are individuals. Columns (1) and (2) use the restricted sample; Columns (3) and (4) use the full sample. Columns (1) and (3) use no controls other than proximity to nearby production; Columns (2) and (4) also control for number of entries, acreage, and month-of-lottery fixed effects. Regressions use conventional OLS standard errors.

	(1)	(2)	(3)	(4)
	Major Firm			
Nearby Production Flag	-0.028	-0.025	-0.039	-0.037
	(0.029)	(0.029)	(0.027)	(0.027)
Constant	0.485***	0.792***	0.493***	0.787***
	(0.013)	(0.077)	(0.012)	(0.076)
Number of Entries & Acreage Controls	No	Yes	No	Yes
Month of Lottery Fixed Effects	No	Yes	No	Yes
R squared	0.001	0.109	0.001	0.108
Observations	1800	1800	2047	2047

Table B.29: We examine the probability that a firm appearing as a first-, second-, or third-place winner is a "major" firm. The sample is limited to entrants in the data that are firms. Columns (1) and (2) use the restricted sample; Columns (3) and (4) use the full sample. Columns (1) and (3) use no controls other than proximity to nearby production; Columns (2) and (4) also control for number of entries, acreage, and month-of-lottery fixed effects. Regressions use conventional OLS standard errors.

	Seven Sisters Supermajor			
Nearby Production Flag	-0.009	-0.010	-0.010	-0.011
	(0.009)	(0.009)	(0.008)	(0.008)
Constant	0.025***	0.039	0.025***	0.039*
	(0.004)	(0.024)	(0.004)	(0.023)
Number of Entries & Acreage Controls	No	Yes	No	Yes
Month of Lottery Fixed Effects	No	Yes	No	Yes
R squared	0.001	0.054	0.001	0.050
Observations	1800	1800	2047	2047

Table B.30: We examine the probability that a firm appearing as a first-, second-, or third-place winner is a "seven sisters" supermajor firm. The sample is limited to entrants in the data that are firms. Columns (1) and (2) use the restricted sample; Columns (3) and (4) use the full sample. Columns (1) and (3) use no controls other than proximity to nearby production; Columns (2) and (4) also control for number of entries, acreage, and month-of-lottery fixed effects. Regressions use conventional OLS standard errors.

First Word of Firm Name	Frequency in Leases with Nearby Production (%)	Frequency in Leases without Nearby Production (%)
Beard	5.1	1.8
Mar-Win	4.3	3.0
Syracuse	4.3	4.8
Allied	3.2	3.6
Husky	3.2	1.5
Yates	2.9	2.9
Reserve	2.1	1.8
Murphy	1.9	
UV	1.9	2.8
Creslenn	1.6	
Rainbow	1.6	1.7
Read	1.6	1.6
Western	1.6	1.0
Aquitaine	1.3	2.1
Barlow	1.3	2.5
Hawthorn	1.3	1.1
Kanorado	1.3	1.1
Marshall	1.3	1.1
Pennzoil	1.3	1.4
Petroleum	1.3	1.0
Skyline	1.3	2.2
Texas	1.3	1.0
Banta		1.9
Beren		1.4
Flag-Redfern		1.4
Gulf		1.3
Mountain		1.3
Oklahoma		1.2
Short		1.2
Kansas-Nebraska		1.0
Kissinger		1.0

Table B.31: In our restricted sample of 1,800 leases there are 376 leases with nearby production and 1,424 leases without nearby production. This table displays the most common first word of the firm name for frequent winners. Roughly half of the leases are represented in each column.

Correlations between type and outcome measures

How do lease trading, drilling and production outcomes differ depending on type of individual or firm? Here we examine our two categorizations of individuals (sophisticated versus unsophisticated and whether individuals have a similar address to firms) and our two categorizations of firms (major versus minor as well as whether the firm is a seven sisters supermajor.)

Sophisticated vs. Unsophisticated Individuals: To examine how outcomes vary depending on whether an individual is "sophisticated" or not, we run a regression similar to Equation 1, but where we also include an indicator for whether the winner is a sophisticated individual, as well as an interaction between sophisticated and nearby production. Sophisticated individuals are defined as discussed above. The estimate of β_1^S determines whether lease outcomes for leases far from existing production differ depending on whether the winner was a sophisticated versus unsophisticated individual; the estimate of $\beta_1^S + \beta_3^S$ determines whether lease outcomes for leases close to existing production differ depending on whether the winner was a sophisticated versus unsophisticated individual. Our results for trade outcomes are in Table B.32; our results for drilling and production outcomes are in Table B.33.

$$\begin{aligned} Y_i = \beta_0 + \beta_1^F F_i + \beta_1^S S_i + \beta_2 NearbyProd_i \\ + \beta_3^F NearbyProd_i * F_i + \beta_3^S NearbyProd_i * S_i + \Omega X_i + \varepsilon_i \end{aligned} \tag{46}$$

We use two types of samples for these regressions. The first is the full sample, included in the odd–numbered regression columns. In the even-numbered columns we use a sample of leases where exactly one of the lease's three winners is a firm, one winner is a sophisticated individual, and one is an unsophisticated individual. This ensures we identify the causal effect of a sophisticated winner within this sample.

In Table B.32, Column (2), we find some evidence that sophisticated individuals are more likely to trade their leases than unsophisticated individuals when the lease is far from existing production (but this is not reflected in Column 1). Columns (3) and (4) shows that conditional on trade, sophisticated individuals tend to take somewhat longer to trade. Columns (5) and (6) show that sophisticated individuals may be less likely to trade with nearby producing firms (NPF) than unsophisticated individuals are, exhibiting similar trading patterns as firms.

Table B.33 shows some evidence that sophisticated individuals behave similarly to firms. For leases far from production, we do not find that assignment to sophisticated individuals has an impact on drilling and production outcomes. However, for leases close to existing production, we find outcomes for parcels won by sophisticated winners are similar to those of parcels won by firms, with both lower drilling and production probabilities. Although the estimates of $\beta_1^S + \beta_3^S$ in Columns (2) and (4) are not statistically significant, they are of similar magnitude as the estimates in Columns (1) and (3).

	(1)	(2)	(3)	(4)	(5)	(6)
	Reassign Probability		Log Time to Reassign		Trade with NPF	
Firm Winner	-0.212***	-0.122***	0.690***	0.660***	-0.002	0.073
	(0.026)	(0.044)	(0.070)	(0.127)	(0.020)	(0.058)
Sophisticated Individual	0.002	0.125***	0.248***	0.163	0.008	0.013
	(0.012)	(0.045)	(0.038)	(0.133)	(0.011)	(0.033)
Nearby Production Flag	0.016	0.135	-0.008	-0.146	0.056***	0.090*
	(0.014)	(0.087)	(0.060)	(0.257)	(0.013)	(0.054)
Firm/Nearby Prod Interaction	-0.053	-0.164	-0.073	0.411	-0.095***	-0.191**
	(0.057)	(0.105)	(0.154)	(0.256)	(0.025)	(0.097)
Sophisticated/Nearby Prod Interaction	0.002	-0.169*	-0.010	0.212	-0.046***	0.007
	(0.019)	(0.087)	(0.047)	(0.266)	(0.017)	(0.070)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.131	0.200	0.114	0.237	0.049	0.189
$E(Y_i \mid F_i = S_i = NearbyProd_i{=}0)$	0.761	0.627	6.343	6.446	0.073	0.016
p-value: $\beta_1^F + \beta_3^F = 0$	0.000	0.002	0.000	0.000	0.000	0.023
p-value: $\beta_1^S + \beta_3^S = 0$	0.773	0.588	0.000	0.108	0.002	0.724
Observations	10762	761	6920	406	5043	352

Table B.32: This table's dependent variables are the probability of reassignment by twelve years (Columns 1 & 2), the log length of time to reassignment (conditional on reassignment) (Columns 3 & 4), and whether traded to the nearby producing firm (Columns 5 & 6). Columns (1), (3), and (5) use our restricted sample (the winners include one sophisticated individual, one unsophisticated individual, and one firm), while Columns (2), (4), and (6) use the full sample. Columns (5) and (6) are limited to leases that are within 5.2 miles of production. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

	(1)	(2)	(3)	(4)
	Drilling Probability		Production Probability	
Firm Winner	0.010	0.016	0.002	0.013
	(0.012)	(0.017)	(0.006)	(0.010)
Sophisticated Individual	-0.001	-0.015	0.005	-0.003
	(0.005)	(0.021)	(0.004)	(0.011)
Nearby Production Flag	0.127***	0.131***	0.073***	0.065*
	(0.030)	(0.048)	(0.017)	(0.034)
Firm/Nearby Prod Interaction	-0.099***	-0.154**	-0.044**	-0.062
	(0.029)	(0.068)	(0.021)	(0.050)
Sophisticated/Nearby Prod Interaction	-0.037**	-0.031	-0.035**	-0.010
	(0.016)	(0.059)	(0.015)	(0.052)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.078	0.138	0.062	0.103
$E(Y_i \mid F_i = S_i = NearbyProd_i{=}0)$	0.060	0.052	0.020	0.011
p-value: $\beta_1^F + \beta_3^F = 0$	0.000	0.036	0.026	0.313
p-value: $\beta_1^S + \beta_3^S = 0$	0.013	0.437	0.037	0.812
Observations	10762	761	10762	761

Table B.33: This table's dependent variables are the probability of drilling within twelve years (Columns 1 and 2) and the probability of production within twelve years (Columns 3 and 4). Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

Individuals: Whether shares a firm's address: We then use our full sample to examine two alternative regression results. In the first, we separately examine firms ($F_i = 1$) versus individuals whose addresses close match those of firms ($A_i = 1$) versus individuals who do not. This is shown in Equation 47:

$$\begin{aligned} Y_i = \beta_0 + \beta_1^F F_i + \beta_1^A A_i + \beta_2 NearbyProd_i \\ + \beta_3^F NearbyProd_i * F_i + \beta_3^A NearbyProd_i * A_i + \Omega X_i + \varepsilon_i \end{aligned} \tag{47}$$

In the second specification, we effectively recategorize as firms those individuals whose addresses closely match those of firms and use a regression specification similar to that in Equation 1:

$$\begin{aligned} Y_i =& \beta_0 + \beta_1^{FA}(F_i + A_i) + \beta_2 NearbyProd_i \\ &+ \beta_3^{FA} NearbyProd_i * (F_i + A_i) + \Omega X_i + \varepsilon_i \end{aligned} \tag{48}$$

Here $F_i = 1$ only for firms and $A_i = 1$ only for individuals with addresses similar to firms, such that $F_i + A_i$ is an indicator variable for whether an entrant is either a firm or an individual with an address similar to a firm. In the following regression tables, we denote $F_i + A_i$ as "redefined firm", meaning either firms or individuals with addresses sufficiently close to firms such that we might redefine them as firms.

Tables B.34 through B.38 include our regression results, with one table for each outcome. For each outcome variable, we examine results using the two different address thresholds, with the 0.72 threshold in Columns (1) and (3) and the 0.85 threshold in Columns (2) and (4). We also use both regression specifications, with Equation 47 in Columns (1) and (2) and Equation 48 in Columns (3) and (4). In these regressions, we only use the full sample. We are unable to construct a restricted sample that ensures exogeneity for this address exercise because we only observe address for the first-place winner.

When examining trade, we find in Tables B.34 and B.35 that individuals whose addresses are similar to firms have different reassignment patterns, and that such effects differ depending on whether we use the 0.85 threshold (Columns 2 and 4 of each table), which is more restrictive in determining which individuals have addresses sufficiently close to firms – or the looser 0.72 threshold (Columns 1 and 3 of each table). In Columns (1) and (2) of Table B.34, we find that relative to other individuals, individuals with addresses similar to firms may be more likely to sell their leases when their leases are far from existing production – but less likely to sell their leases when their lease is close to existing production. In Columns (1) and (2) of Table B.35, we find that individuals with addresses similar to firms tend to take a longer amount of time to trade their leases.

When examining trade with the nearby producing firm, we find in Table B.36 that for leases further from existing production, those won by individuals with addresses similar to firms do not have significantly different probabilities of trade with the nearby producing firm than other individuals (Columns 1 and 2). However, when we redefine firms to include either firms or individuals with sufficiently similar addresses, we find some evidence that even for leases further from production, those won by firms are less likely to sell to the nearby producing firm. For leases close to existing production, we do not find statistically significant differences between individuals with addresses similar to those of firms and other individuals. We also find that when we redefine firms to include individuals with similar addresses, our estimates of $\beta_1^{FA} + \beta_3^{FA}$ show firms are less likely to trade with the nearby producing firm, and that this is especially true for the stricter address threshold (Column 4).

Drilling regression results in Table B.37, Columns (1) and (2), shows that for leases far from existing production, individuals who are at a firm address are slightly less likely to drill than individuals who are not. For leases that are close to existing production, individuals who are at a firm address are less likely to drill than those individuals who are not at a firm address, with the difference statistically significant when using the 0.85 similarity score cutoff. This is suggestive that individuals who list addresses as firms behave somewhat similarly to firms.

Columns (3) and (4) of Table B.37 examines drilling results where we recategorize as firms those individuals who addresses are close to firms. We find that for leases far from existing production, we have similar results to our main table in Table 5: Leases won by individuals and firms have similar probabilities of drilling. For leases close to existing production, we find that those won by firms have a statistically significant lower probability of drilling than those won by individuals when we use the stricter 0.85 threshold, but not the 0.72 threshold. This is suggestive of the idea that using a 0.72 threshold is too lax of a threshold, with the estimate of $\beta_1^{FA} + \beta_3^{FA}$ being dampened by the inclusion of individuals whose leases tend to have a high probability of drilling when close to existing production.

Similar findings hold for production results in Table B.38, but the results are dampened, possibly due to decreased power because our production regressions have fewer cases with production. Therefore, we do find point estimates that suggest that for leases close to existing production, individuals with addresses similar to firms are less likely to drill than individuals with other addresses (Columns 1 and 2). However, the estimates are not statistically significant. When recategorizing individuals as firms if their addresses are sufficiently close to that of firms (Columns 3 and 4), we find that for leases close to existing production, those won by firms have lower probability of production than those won by individuals – but that this finding only holds when we use the stricter 0.85 threshold (Column 4), and does not hold with the looser 0.72 threshold (Column 3).

Altogether, these results suggest that our main results are not due to individuals acting on behalf of firms. Using our stricter 0.85 threshold to recategorize some individuals as firms, we find similar results for drilling and production as in our main tables in Table 5: For leases far from existing production, leases won by individuals and firms have similar probabilities of drilling and production; for leases close to existing production, leases won by individuals are more likely to have drilling and production than leases won by firms. Our smaller point estimates for $\beta_1 + \beta_3$ suggest that such recategorization is also leading to measurement error and attenuation bias, where some individuals are being recategorized as firms in this specification but are behaving more similarly to other individuals that are not being recategorized as firms.

	(1)	(2)	(3)	(4)
	Reassign Probability			
Firm Winner	-0.216***	-0.209***		
	(0.026)	(0.026)		
Individual at Firm Address	-0.011	0.029**		
	(0.011)	(0.013)		
Redefined Firm			-0.053***	-0.046***
			(0.011)	(0.013)
Nearby Production Flag	0.028**	0.024*	0.027**	0.024*
	(0.012)	(0.013)	(0.012)	(0.013)
Firm/Nearby Prod Interaction	-0.064	-0.061		
	(0.057)	(0.056)		
Indiv. at Firm Address x Nearby Production	-0.042*	-0.047*		
	(0.024)	(0.028)		
Redefined Firm x Nearby Production			-0.050*	-0.053*
			(0.025)	(0.028)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.131	0.131	0.119	0.117
$\mathrm{E}(Y_i \mid F_i = 0, A_i = 0, NearbyProd_i{=}0)$	0.765	0.757	0.765	0.757
p-value: $\beta_1^F + \beta_3^F = 0$	0.000	0.000		
p-value: $\beta_1^A + \beta_3^A = 0$	0.017	0.454		
p-value: $\beta_1^{FA} + \beta_3^{FA} = 0$			0.000	0.000
Observations	10762	10762	10762	10762

Table B.34: Regressions where the dependent variable is whether the lease was sold within 12 years. Columns (1) and (2) use the regression specification in Equation 47. Columns (3) and (4) use the regression specification in Equation 48. Columns (1) and (3) define individuals at a firm address (A_i) as those where the similarity score is 0.72 or above; Columns (2) and (4) define individuals at a firm address (A_i) as those where the similarity score is 0.85 or above.

	(1)	(2)	(3)	(4)
	Log Time to Reassign			
Firm Winner	0.696***	0.665***		
	(0.065)	(0.066)		
Individual at Firm Address	0.207***	0.154***		
	(0.025)	(0.031)		
Redefined Firm			0.282***	0.273***
			(0.027)	(0.034)
Nearby Production Flag	-0.028	-0.048	-0.028	-0.048
	(0.048)	(0.049)	(0.048)	(0.049)
Firm/Nearby Prod Interaction	0.111	0.133		
	(0.131)	(0.134)		
Indiv. at Firm Address x Nearby Production	-0.036	0.065		
	(0.065)	(0.086)		
Redefined Firm x Nearby Production			-0.004	0.087
			(0.060)	(0.084)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.121	0.118	0.112	0.110
$E(Y_i \mid F_i = 0, A_i = 0, NearbyProd_i{=}0)$	6.084	6.114	6.084	6.114
p-value: $\beta_1^F + \beta_3^F = 0$	0.000	0.000		
p-value: $\beta_1^A + \beta_3^A = 0$	0.004	0.004		
p-value: $\beta_1^{FA} + \beta_3^{FA} = 0$			0.000	0.000
Observations	8120	8120	8120	8120

Table B.35: Regressions where the dependent variable is log time until trade (conditional on trade). Columns (1) and (2) use the regression specification in Equation 47. Columns (3) and (4) use the regression specification in Equation 48. Columns (1) and (3) define individuals at a firm address (A_i) as those where the similarity score is 0.72 or above; Columns (2) and (4) define individuals at a firm address (A_i) as those where the similarity score is 0.85 or above.

	(1)	(2)	(3)	(4)
	Trade with Nearby Producing Firm			
Firm Winner	-0.008	-0.009		
	(0.019)	(0.019)		
Individual at Firm Address	-0.011	-0.024		
	(0.013)	(0.015)		
Redefined Firm			-0.011	-0.019*
			(0.012)	(0.011)
Nearby Production Flag	0.033***	0.034***	0.033***	0.034***
	(0.010)	(0.008)	(0.010)	(0.008)
Firm/Nearby Prod Interaction	-0.073***	-0.073***		
	(0.027)	(0.026)		
Indiv. at Firm Address x Nearby Production	0.008	0.014		
	(0.015)	(0.018)		
Redefined Firm x Nearby Production			-0.010	-0.014
			(0.015)	(0.014)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.047	0.047	0.045	0.046
$E(Y_i \mid F_i = 0, A_i = 0, NearbyProd_i{=}0)$	0.080	0.080	0.080	0.080
p-value: $\beta_1^F + \beta_3^F = 0$	0.000	0.000		
p-value: $\beta_1^A + \beta_3^A = 0$	0.753	0.546		
p-value: $\beta_1^{FA} + \beta_3^{FA} = 0$			0.084	0.013
Observations	5043	5043	5043	5043

Table B.36: Regressions where the dependent variable is whether there is trade with the nearby producing firm. Columns (1) and (2) use the regression specification in Equation 47. Columns (3) and (4) use the regression specification in Equation 48. Columns (1) and (3) define individuals at a firm address (A_i) as those where the similarity score is 0.72 or above; Columns (2) and (4) define individuals at a firm address (A_i) as those where the similarity score is 0.85 or above. Sample is limited to those leases that are within 5.2 miles for existing production.

	(1)	(2)	(3)	(4)
		Drilling Probability		
Firm Winner	0.009	0.009		
	(0.011)	(0.011)		
Individual at Firm Address	-0.009*	-0.015**		
	(0.005)	(0.007)		
Redefined Firm			-0.005	-0.007
			(0.005)	(0.006)
Nearby Production Flag	0.110***	0.113***	0.110***	0.113***
	(0.029)	(0.027)	(0.029)	(0.027)
Firm/Nearby Prod Interaction	-0.082***	-0.085***		
	(0.027)	(0.026)		
Indiv. at Firm Address x Nearby Production	0.003	-0.014		
	(0.021)	(0.015)		
Redefined Firm x Nearby Production			-0.015	-0.037***
			(0.020)	(0.014)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.078	0.078	0.077	0.078
$E(Y_i \mid F_i = 0, A_i = 0, NearbyProd_i{=}0)$	0.062	0.062	0.062	0.062
p-value: $\beta_1^F + \beta_3^F = 0$	0.002	0.001		
p-value: $\beta_1^A + \beta_3^A = 0$	0.783	0.029		
p-value: $\beta_1^{FA} + \beta_3^{FA} = 0$			0.264	0.000
Observations	10762	10762	10762	10762

Table B.37: Regressions where the dependent variable is whether the lease was drilled on within 12 years. Columns (1) and (2) use the regression specification in Equation 47. Columns (3) and (4) use the regression specification in Equation 48. Columns (1) and (3) define individuals at a firm address (A_i) as those where the similarity score is 0.72 or above; Columns (2) and (4) define individuals at a firm address (A_i) as those where the similarity score is 0.85 or above.

	(1)	(2)	(3)	(4)
	Production Probability			
Firm Winner	0.000	-0.001		
	(0.005)	(0.005)		
Individual at Firm Address	0.002	-0.001		
	(0.003)	(0.005)		
Redefined Firm			0.002	-0.001
			(0.003)	(0.004)
Nearby Production Flag	0.059***	0.061***	0.059***	0.061***
	(0.012)	(0.012)	(0.012)	(0.012)
Firm/Nearby Prod Interaction	-0.030	-0.032*		
	(0.019)	(0.019)		
Indiv. at Firm Address x Nearby Production	-0.004	-0.020		
	(0.015)	(0.016)		
Redefined Firm x Nearby Production			-0.010	-0.024*
			(0.013)	(0.014)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.061	0.061	0.060	0.061
$E(Y_i \mid F_i = 0, A_i = 0, NearbyProd_i{=}0)$	0.022	0.022	0.022	0.022
p-value: $\beta_1^F + \beta_3^F = 0$	0.080	0.062		
p-value: $\beta_1^A + \beta_3^A = 0$	0.899	0.113		
p-value: $\beta_1^{FA} + \beta_3^{FA} = 0$			0.490	0.033
Observations	10762	10762	10762	10762

Table B.38: Regressions where the dependent variable is whether the lease was sold within 12 years. Columns (1) and (2) use the regression specification in Equation 47. Columns (3) and (4) use the regression specification in Equation 48. Columns (1) and (3) define individuals at a firm address (A_i) as those where the similarity score is 0.72 or above; Columns (2) and (4) define individuals at a firm address (A_i) as those where the similarity score is 0.85 or above.

Major and minor firms: To examine how various outcomes differ between major and minor firms, we run the following regression

$$\begin{aligned} Y_i = \; & \beta_0 + \beta_1^F F_i + \beta_1^{Maj} Maj_i + \beta_2 NearbyProd_i \\ & + \beta_3^F NearbyProd_i * F_i + \beta_3^{Maj} NearbyProd_i * Maj_i + \Omega X_i + \varepsilon_i \end{aligned} \tag{49}$$

where F_i is an indicator that the winner is a firm (either major or minor) and Maj_i is an indicator that the winner is a *major* firm. To examine whether outcomes differ between major and minor firms, we examine whether β_1^{Maj} and β_3^{Maj} are statistically different from zero.

We first examine trade outcomes, finding that for leases that are traded, major firms take about 20% longer until first trade. However, for all other outcomes we find no statistically significant differences between major and minor firms. In Table B.39, Columns (1) and (2) examine whether a lease is traded within 12 years, Columns (3) and (4) examine time until trade (if trade happened), and Columns (5) and (6) examine whether the lease was traded to the nearby producing firm.

Then, Table B.40 examines drilling and production outcomes, finding that for leases far from existing production, leases won by major firms have a slightly higher probability of production than those won by minor firms (significant at the 10% level). However, we find no statistically significant differences in the probability of drilling. For leases close to existing production, those won by major firms have a weakly significant lower probability of drilling and production than those won by minor firms.

Our results provide weak evidence that major firms have a higher value of leases than minor firms. Therefore, information asymmetries caused by an informed buyer have a larger adverse effect on drilling and production when the initial winner is a major, rather than a minor, firm. This pattern is similar to the overall pattern we observe for all firms versus individuals.

	(1)	(2)	(3)	(4)	(5)	(6)
	Reassign Probability		Log Time to Reassign		Trade with NPF	
Firm Winner	-0.221***	-0.217***	0.549***	0.544***	-0.041	-0.014
	(0.035)	(0.035)	(0.106)	(0.086)	(0.032)	(0.020)
Major Firm	-0.004	0.008	0.202*	0.210**	0.031	0.018
	(0.053)	(0.054)	(0.112)	(0.094)	(0.046)	(0.038)
Nearby Production Flag	0.024	0.017	0.019	-0.039	0.030	0.036***
	(0.038)	(0.014)	(0.105)	(0.048)	(0.026)	(0.009)
Firm/Nearby Prod Interaction	-0.035	-0.024	0.042	0.088	-0.040	-0.064*
	(0.072)	(0.066)	(0.224)	(0.178)	(0.055)	(0.038)
Major Firm/Nearby Prod Interaction	-0.036	-0.062	0.059	0.094	-0.024	-0.022
	(0.083)	(0.092)	(0.248)	(0.250)	(0.052)	(0.044)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.158	0.131	0.163	0.116	0.115	0.047
$\mathrm{E}(Y_i \mid F_i = NearbyProd_i{=}0)$	0.736	0.762	6.210	6.137	0.078	0.077
Observations	1800	10762	1200	8120	819	5043

Table B.39: This table's dependent variables are the probability of reassignment by twelve years in Columns (1) and (2), log length of time until reassignment conditional on reassignment in Columns (3) and (4), and whether the lease was traded to the nearby producing firm in Columns (5) and (6). Columns (1), (3), and (5) use the restricted sample; Columns (2), (4) and (6) use the full sample. Columns (5) and (6) are limited to leases within 5.2 miles of existing production. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

	(1)	(2)	(3)	(4)
	Drilling Probability		Production Probability	
Firm Winner	0.013	0.007	-0.010	-0.009
	(0.015)	(0.014)	(0.008)	(0.006)
Major Firm	0.001	0.010	0.021*	0.018*
	(0.019)	(0.017)	(0.012)	(0.010)
Nearby Production Flag	0.136***	0.111***	0.074***	0.058***
	(0.026)	(0.026)	(0.020)	(0.011)
Firm/Nearby Prod Interaction	-0.085*	-0.055	-0.008	-0.000
	(0.050)	(0.041)	(0.038)	(0.031)
Major Firm/Nearby Prod Interaction	-0.044	-0.058	-0.071*	-0.059
	(0.056)	(0.055)	(0.041)	(0.038)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.115	0.078	0.089	0.061
$E(Y_i \mid F_i = NearbyProd_i{=}0)$	0.046	0.060	0.017	0.022
Observations	1800	10762	1800	10762

Table B.40: This table's dependent variables are the probability of drilling within twelve years in Columns (1) and (2) and the probability of production within twelve years in Columns (3) and (4). Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

Supermajors: To examine how various outcomes differ between supermajor firms (the seven sisters) and other firms, we run the following regression

$$\begin{aligned} Y_i = \beta_0 + \beta_1^F F_i + \beta_1^{Sup} Sup_i + \beta_2 NearbyProd_i \\ + \beta_3^F NearbyProd_i * F_i + \beta_3^{Sup} NearbyProd_i * Sup_i + \Omega X_i + \varepsilon_i \end{aligned} \quad (50)$$

where Sup_i is an indicator for whether the firm is a major firm. To examine whether outcomes differ between supermajor firms and other firms, we examine whether β_1^{Sup} and β_3^{Sup} are statistically different from zero. Because of the very small number of leases won by supermajors, all of these regression results should be treated with caution.

We find suggestive evidence that supermajors may be more able to trade to overcome inefficient initial assignment when there is a better agent to utilize the lease. Table B.41 examines trade outcomes, finding that supermajors are significantly less likely to reassign leases to another party and also less likely to sell the lease to the nearby producing firm. We also find that for leases close to existing production, supermajors are much more likely than other firms to reassign their leases. Column (5) suggests that supermajors are more likely to assign to the nearby producing firm, although Column (6) does not show such an effect.

In Table B.42 we examine whether drilling (Columns 1 and 2) and production (Columns 3 and 4) outcomes are different for supermajors than other firms. We find that the estimates of β_1^{Sup} are typically positive but statistically insignificant, which is suggestive that for leases far from existing production, supermajors may be better at drilling and production. On the other hand, for leases close to existing production, the estimates of β_3^{Sup} are typically negative but not close to statistical significance. The small number of supermajors that appear in our data, and especially the few that appear for leases close to production, makes it challenging to make any inference about the effect of supermajors, especially for leases close to existing production.

	(1)	(2)	(3)	(4)	(5)	(6)
	Reassign Probability		Log Time to Reassign		Trade with NPF	
Firm Winner	-0.214***	-0.206***	0.633***	0.631***	-0.024	-0.004
	(0.027)	(0.026)	(0.077)	(0.064)	(0.027)	(0.019)
Supermajor	-0.299**	-0.243**	0.499	0.539	-0.079***	-0.057**
	(0.121)	(0.115)	(0.557)	(0.400)	(0.021)	(0.029)
Nearby Production Flag	0.025	0.017	0.018	-0.039	0.030	0.036***
	(0.038)	(0.014)	(0.106)	(0.048)	(0.026)	(0.009)
Firm/Nearby Prod Interaction	-0.065	-0.064	0.070	0.130	-0.053	-0.075***
	(0.067)	(0.057)	(0.177)	(0.131)	(0.040)	(0.026)
Supermajor/Nearby Prod Interaction	0.677***	0.702***	0.078	-0.091	0.134***	-0.002
	(0.156)	(0.134)	(0.583)	(0.416)	(0.047)	(0.049)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.161	0.131	0.161	0.115	0.114	0.047
Observations	1800	10762	1200	8120	819	5043

Table B.41: This table's dependent variables are the probability of reassignment by twelve years in Columns (1) and (2), log length of time until reassignment conditional on reassignment in Columns (3) and (4), and whether the lease was traded to the nearby producing firm in Columns (5) and (6). Columns (1), (3), and (5) use the restricted sample; Columns (2), (4), and (6) use the full sample. Columns (5) and (6) are limited to leases within 5.2 miles of existing production. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

	(1)	(2)	(3)	(4)
	Drilling Probability		Production Probability	
Firm Winner	0.013	0.010	-0.002	-0.002
	(0.013)	(0.011)	(0.007)	(0.005)
Supermajor	0.025	0.028	0.070	0.065
	(0.072)	(0.065)	(0.069)	(0.062)
Nearby Production Flag	0.136***	0.111***	0.074***	0.058***
	(0.026)	(0.026)	(0.020)	(0.011)
Firm/Nearby Prod Interaction	-0.105***	-0.082***	-0.041*	-0.027
	(0.031)	(0.026)	(0.023)	(0.018)
Supermajor/Nearby Prod Interaction	-0.084	-0.083	-0.106	-0.086
	(0.079)	(0.070)	(0.075)	(0.062)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.115	0.078	0.088	0.061
Observations	1800	10762	1800	10762

Table B.42: This table's dependent variables are the probability of drilling within twelve years in Columns (1) and (2) and the probability of production within twelve years in Columns (3) and (4). Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

Can type explain our findings for leases close to existing production?

We now examine whether outcomes are being driven by changes in the composition of types of individuals and firms. In our main drilling and production results (Table 5), we find that especially for drilling outcomes, $\beta_1 + \beta_3$ (in Equation 1) is negative and statistically different from zero, meaning that for leases close to existing production, those won by firms are much less likely to have drilling than leases won by individuals. Results are similar but less strong for production outcomes. We now test whether these results may be driven by composition of type.

To do this, we first begin with baseline probabilities of drilling for leases far from production – e.g., the baseline probability of drilling for individuals and firms for leases far from production. We then incorporate measures of how these probabilities of drilling differ by type – e.g., how probability of drilling differs for sophisticated versus unsophisticated individuals in those leases that are far from production. We then incorporate information on how type changes with proximity to production – e.g., how the likelihood that an individual is sophisticated changes for leases close to production. Finally, we combine these numbers to calculate estimates of drilling probability for leases close to production, assuming that the only difference between leases close to and far from existing production is the changing composition of type.

Computing these measures for both individuals and firms, we then use these measures to test if the changing composition of type is sufficient to explain the drilling probability differentials for leases close to production that we observe in the data – e.g., sufficient to explain why for leases close to production, we find that $\beta_1 + \beta_3 < 0$. Does the extent to which type composition changes with proximity to production explain why we find leases won by individuals are more likely to be drilled than leases won by firms?

In exploring the role of type, we focus on the sophisticated/unsophisticated individual type rather than the address match because sophisticated/unsophisticated shows a stronger correlation with nearby production (Section B.5).[15] Similarly, for firms, we focus on major/minor firms rather than supermajors because of the very small number of leases where a supermajor wins.

Below we explain the elements for constructing these estimates:

- Baseline regression specification for drilling and production probabilities for leases far from existing production are the estimates of $E(Y|F_i = 0, NearbyProd_i = 0)$ and β_1 estimated using the specification in Equation 1 and shown in Table 5.

- Estimates of γ_1^S and γ_1^{Maj} of how drilling and production outcomes vary with initial winner type, as given in the regression specification in Equation 51 below and as shown in Table B.43.[16]

$$
\begin{aligned}
Y_i = \; & \gamma_0 + \gamma_1^S S_i + \gamma_1^F F_i + \gamma_1^{Maj} Maj_i + \gamma_2 NearbyProd_i \\
& + \gamma_3^S NearbyProd_i * S_i + \gamma_3^F NearbyProd_i * F_i \\
& + \gamma_3^{Maj} NearbyProd_i * Maj_i + \Omega X_i + \varepsilon_i \quad (51)
\end{aligned}
$$

- Estimates of how sophisticated versus unsophisticated type changes for leases close to existing production is the estimate of α_1^S estimated using the specification in Equation 44 and shown in Table B.26.

- Estimates of how major versus minor firm type changes for leases close to existing production is the estimate of α_1^{Maj} estimated using the specification in Equation 45 and shown in Table B.29.

For example, for individuals and drilling, we take the baseline predicted probability of drilling for leases far from existing production, and adjust it by both how proximity to nearby production correlates to type and how type affects the probability of drilling.

[15] For example, compare Tables B.26 with Tables B.27 and B.28.

[16] These are essentially identical to estimates for sophisticated/unsophisticated and major/minor estimates in Tables B.33 and B.40, respectively, only here we estimate the effects of sophisticated/unsophisticated and major/minor within a single regression specification.

	(1)	(2)	(3)	(4)
	Drilling Probability		Production Probability	
Sophisticated Individual	-0.003	-0.001	0.005	0.005
	(0.014)	(0.005)	(0.007)	(0.004)
Firm Winner	0.012	0.006	-0.008	-0.007
	(0.018)	(0.014)	(0.007)	(0.006)
Major Firm	0.001	0.009	0.021*	0.018*
	(0.018)	(0.017)	(0.012)	(0.010)
Nearby Production Flag	0.144***	0.127***	0.095***	0.073***
	(0.032)	(0.030)	(0.031)	(0.017)
Sophisticated/Nearby Prod Interaction	-0.016	-0.037**	-0.041	-0.035**
	(0.045)	(0.016)	(0.040)	(0.015)
Firm/Nearby Prod Interaction	-0.092*	-0.070*	-0.028	-0.015
	(0.050)	(0.042)	(0.044)	(0.033)
Major Firm/Nearby Prod Interaction	-0.044	-0.058	-0.072*	-0.060
	(0.056)	(0.055)	(0.041)	(0.038)
Number of Entries & Acreage Controls	Yes	Yes	Yes	Yes
Month of Lottery Fixed Effects	Yes	Yes	Yes	Yes
R squared	0.115	0.079	0.091	0.062
E($Y_i \mid F_i = S_i = NearbyProd_i$=0)	0.048	0.060	0.015	0.020
p-value: $\beta_1^S + \beta_3^S = 0$	0.695	0.013	0.385	0.036
p-value: $\beta_1^F + \beta_3^F = 0$	0.083	0.091	0.407	0.474
p-value: $\beta_1^F + \beta_1^{Maj} + \beta_3^F + \beta_3^{Maj} = 0$	0.006	0.001	0.008	0.003
Observations	1800	10762	1800	10762

Table B.43: This table uses the regression specification in Equation 51. This table's dependent variables are the probability of drilling within twelve years in Columns (1) and (2) and the probability of production within twelve years in Columns (3) and (4). Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample. Nearby production is a binary indicator for any production within 2.6 miles of the section(s) the lease is located on.

The equation for this is:

$$E(Y_i|\text{individual, adjusted}) = E(Y|F_i = 0, NearbyProd_i = 0) + \alpha_1^S \gamma_1^S \tag{52}$$

In other words, it is the baseline rate for individuals when far from existing production ($E(Y|F_i = 0, NearbyProd_i = 0)$), adjusted by the change in the probability that the individual is sophisticated for leases that are close to existing production (α_1^S) multiplied by the change in the expected outcome Y_i when the individual is a sophisticated individual for leases far from existing production (γ_1^S).

Similarly, for firms, the predicted probability of the outcome Y_i adjusted for changing type composition is:

$$E(Y_i|\text{firms, adjusted}) = E(Y|F_i = 0, NearbyProd_i = 0) + \beta_1 + \gamma_1^{Maj} \alpha_1^{Maj} \tag{53}$$

In other words, it is the baseline rate for firms when far from existing production ($E(Y|F_i = 0, NearbyProd_i = 0) + \beta_1$), adjusted by the change in the probability that the firm is a major firm for leases that are close to existing production (α_1^{Maj}) multiplied by the change in the expected outcome Y_i when the firm is a major firm for leases that are far from existing production (γ_1^{Maj}).

The difference between the two is the difference between Equations 53 and 52 and is given as:

$$E(Y_i|\text{firms, adjusted}) - E(Y_i|\text{individual, adjusted}) = \beta_1 + \gamma_1^{Maj} \alpha_1^{Maj} - \gamma_1^S \alpha_1^S \tag{54}$$

To test whether type composition explains the drilling probability differences we observe for leases close to existing production, we examine whether the difference between firms and individuals for leases close to production ($\beta_1 + \beta_3$) versus the predicted difference due to changing type (in Equation 54) are equal by computing a test statistic:

$$\text{Test statistic} = (\beta_3 + \beta_1) - (\beta_1 + \gamma_1^{Maj}\alpha_1^{Maj} - \gamma_1^S\alpha_1^S) \tag{55}$$

We use our regression specification in Equation 1 to test whether the test statistic equals zero. Results are in Table B.44. The row reporting the test statistic shows that the test statistic is very similar to the estimates of β_3: Although the test statistic is somewhat closer to zero than β_3 is, this test statistic only differs from β_3 by about 0.0005, suggesting that type composition can overall explain very little of the results. We also find that the p-value of a test that the test statistic is equal to zero is rejected at the 1% level for drilling outcomes. For the production outcomes, we find less strong statistical results: We find p values below 0.07 for the restricted sample and 0.11 for the full sample. These p values are of similar magnitudes as the p values of the test that $\beta_1 + \beta_3 = 0$ in our main production results in Table 5.

Because leases close to existing production have a higher probability of drilling and production overall, we also use an alternative specification where instead of examining differences in probabilities, we instead examine the ratio of probabilities. We construct the ratio of the predicted probability of drilling for firms relative to individuals, both using our baseline drilling estimates from Table 5, as well as our adjusted-for-type estimates from Equations 52 and 53, and test whether the two ratios are equal. In other words, we use our regression estimates from Table 5 and test whether the following holds using a non-linear test of equality that utilizes the delta method:

$$\left[\frac{E(Y|F_i = 0, NearbyProd_i = 0) + \beta_1 + \gamma_1^{Maj}\alpha_1^{Maj}}{E(Y|F_i = 0, NearbyProd_i = 0) + \gamma_1^S\alpha_1^S}\right] = \left[\frac{E(Y|F_i = 0, NearbyProd_i = 0) + \beta_1 + \beta_2 + \beta_3}{E(Y|F_i = 0, NearbyProd_i = 0) + \beta_2}\right] \tag{56}$$

The p-value for this test is in the last row of Table B.44. Using this specification, we are able to reject at the 1% level that the two are equal for the drilling specifications. In other words, even accounting for the fact that leases close to production should have higher probabilities of drilling anyway, we do not find that our predicted drilling probabilities can be explained by the changing composition of type for leases that are close to existing production. However, for production outcomes, we are not able to reject that they can be driven by type. This is consistent with our findings that in general due to the lower likelihood of production relative to drilling, there is less power in our production regressions than there is in our drilling regressions.

Outcome:	Drilling		Production	
Sample Type:	Restricted	Full	Restricted	Full
α_1^S	-0.0766	-0.0571	-0.0766	-0.0571
α_1^{Maj}	-0.0278	-0.0389	-0.0278	-0.0389
γ_1^S	-0.0028	-0.0007	0.0053	0.0047
γ_1^{Maj}	0.0010	0.0090	0.0209	0.0178
β_3	-0.1063	-0.0828	-0.0431	-0.0292
Test statistic	-0.1060	-0.0824	-0.0429	-0.0288
p-value: Test statistic = 0	0.0000	0.0007	0.0689	0.1091
p-value: Nonlinear test statistic	0.0056	0.0044	0.2988	0.2680

Table B.44: This table reports coefficient estimates as part of a test to determine the extent to which the composition of types (sophisticated versus unsophisticated individuals and major versus minor firms) drives the findings for drilling and production findings for leases close to existing production. The rows reporting α replicate the coefficients reported in Tables B.26 and B.29; the rows reporting γ replicate coefficient estimate results from Table B.43, and the row reporting β_3 gives coefficient estimates from Table 5. The test statistic is computed as described in Equation 55.

B.6 Oil Prices

In this section we discuss the role of oil prices. We first explore the relationship between oil prices and drilling timing. We show in a survival analysis that our drilling results are robust to controlling for crude oil prices.

Changing oil prices, trading, and drilling patterns

Figure B.13 graphs the evolution of Wyoming oil wellhead prices and the date of first lease transfer over time. Prices were relatively constant until 1979, when they rose rapidly to an early 1981 peak. This was followed by declining prices, with a particularly sharp drop in 1986. They were then mostly lower, though volatile, until the 1990s. Figure B.13 demonstrates that much of the trade that happened was prior to the price increase.

Figure B.14 shows the same oil price data as well as histograms of the date that the first well's drilling began (spud date). We differentiate between whether drilling happened between zero and eight years of the lease start or between eight and twelve years of the lease start date. It shows that the price peak in 1981 is accompanied by a spike in drilling. There is a similarly sized drilling peak in 1985-1986 when oil prices were declining but when many leases were close to expiring (eight to twelve years from lease start date). This is consistent with evidence from Hendricks and Porter (1996), Agerton (2020) and Herrnstadt, Kellogg, and Lewis (2020), who show in other drilling settings that a significant fraction of drilling happens right before lease expiration.

Hazard Analysis

We next examine how our main trade and drilling results in Tables 3 and 5 change when accounting for changing oil prices.[17] To do this, we use a survival analysis where we construct a panel data set of lease by month observations from the start of the lease until twelve years after the lottery date. The two dependent variables we examine are whether trade happens and whether drilling begins. Because we are interested in predicting the first date of trade and drilling, we exclude observations that happen after the first trade (or first drilling). We use the following linear probability regression

[17] We do not examine production in the light of changing oil prices because we do not have first production dates for wells that produced before 1978.

specification:

$$1(Y_{it} = 1 | Y_{is} = 0 \; \forall \; s < t) = \beta_0 + \beta_1 F_i + \beta_2 NearbyProd_i$$
$$+ \beta_3 NearbyProd_i * F_i + \Omega X_{it} + \varepsilon_i \quad (57)$$

Here Y_{it} is an indicator variable for lease i indicating whether the outcome (either first trade or first drilling) happened in month t. The term $1(Y_{it} = 1 \mid Y_{is} = 0 \; \forall \; s < t)$ represents the fact that the sample is limited to those observations up to the first date of drilling.

The right-hand-side covariates include our standard parcel-specific variables, including whether the winner was a firm (β_1), whether the parcel was close to existing production (β_2), and the interaction of the previous two (β_3). Other non-time-varying covariates include acreage of the lease and number of entries. There are also time-varying covariates: One time-varying variable is the wellhead price at date t. The other is a five-knot spline in the age of the lease, which allows the probabilities to vary depending on how close the lease is to expiration.

Our trade results are in Table B.45; our drilling results are in Table B.46. For each table, Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample. Columns (1) and (2) control linearly for the age of the lease; Columns (3) and (4) control for a five-knot spline in the age of the lease.

Table B.45 shows that oil prices are negatively associated with trade. This may be due to initial lessees being more optimistic about the profitability of retaining and drilling when prices are higher. Proximity to nearby production is not associated with different baseline trade probabilities when winners are individuals. We find that firms baseline probability of trading is lower than individuals in general, and the full sample results in Columns (2) and (4) suggest that leases won by firms are especially unlikely to be traded if they are close to existing production.

Table B.46 shows that oil prices significantly affect the timing of drilling: A

$1 increase in oil wellhead prices increasing the probability of first drilling in a given month by between 0.0013% and 0.0021%. We find that, as expected, proximity to nearby production is associated with higher drilling probabilities. For leases far from existing production, we do not find that assignment to a firm has any effect on the baseline probability of drilling. But for leases close to existing production, assignment to a firm decreases the baseline probability.

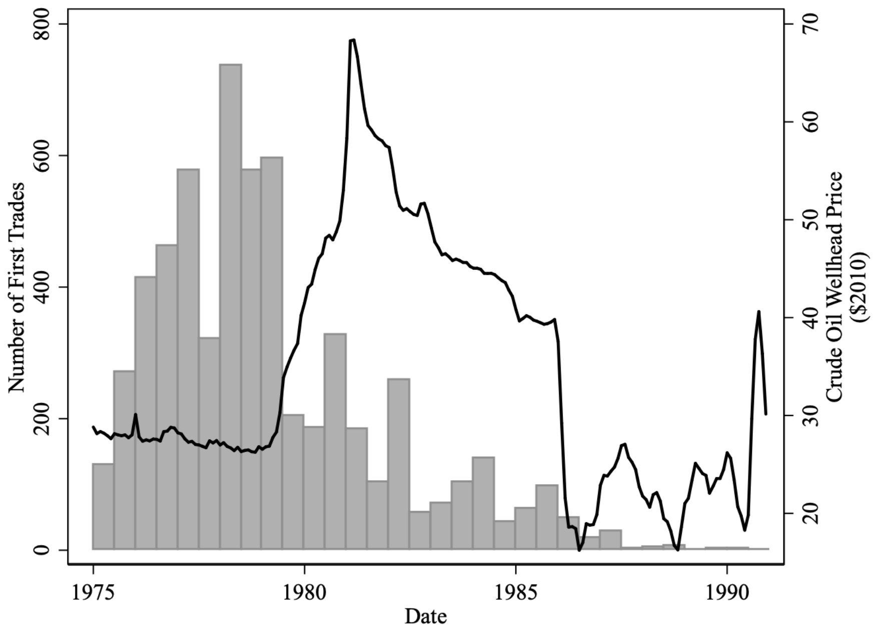


Figure B.13: Oil prices and first trade dates

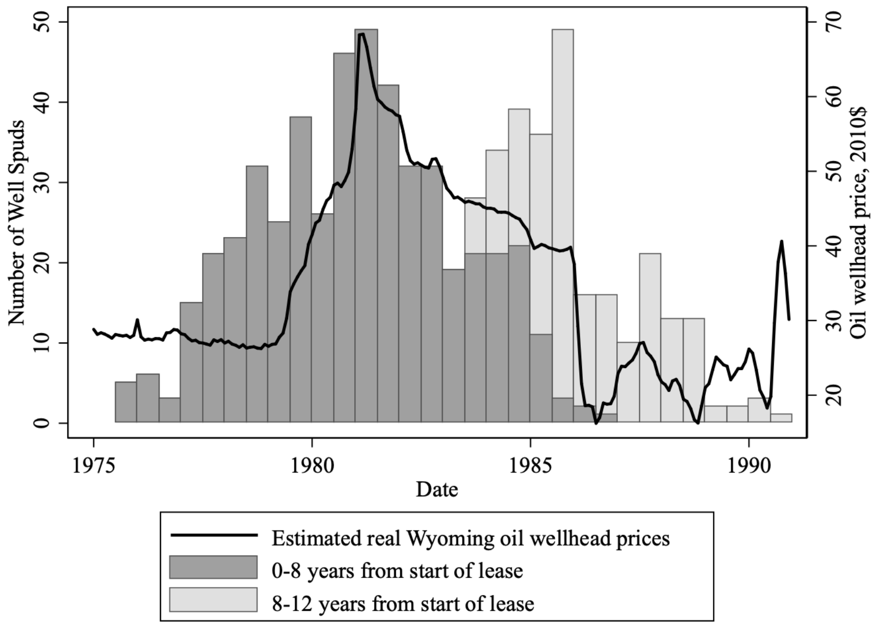


Figure B.14: Oil prices and first drilling dates, decomposed into timing of drilling date relative to start of lease.

References

Agerton, M. (2020). Learning where to drill: Drilling decisions and geological quality in the Haynesville shale. USAEE Working Paper No. 20-439.

Dixit, A. K. and R. S. Pindyck (1994). *Investment under Uncertainty.* New Jersey: Princeton University Press.

Hendricks, K. and R. H. Porter (1996). The timing and incidence of exploratory drilling on offshore wildcat tracts. *The American Economic Review 86*(3), 388–407.

Herrnstadt, E., R. Kellogg, and E. Lewis (2020). The economics of time-limited development options: The case of oil and gas leases. Working Paper 27165, NBER.

	(1)	(2)	(3)	(4)
		Probability of First Trade		
Oil Wellhead Price (2010 Dollars)	-0.000099***	-0.000214***	0.000077**	0.000099***
	(0.000030)	(0.000034)	(0.000032)	(0.000019)
Firm Winner	-0.006188***	-0.005673***	-0.005953***	-0.005348***
	(0.000649)	(0.000508)	(0.000617)	(0.000474)
Nearby Production Flag	0.001370	0.001227	0.001321	0.001184
	(0.001440)	(0.000780)	(0.001393)	(0.000739)
Firm/Nearby Prod Interaction	-0.002330	-0.002497***	-0.002237	-0.002363***
	(0.001580)	(0.000968)	(0.001525)	(0.000910)
R squared	0.007621	0.011030	0.008685	0.012826
p-value: $\beta_1 + \beta_3 = 0$	0.000000	0.000000	0.000000	0.000000
Observations	143850	676848	143850	676848

Table B.45: This table uses a linear probability specification and a survival data framework to examine the probability that the leases's first trade happens in a given month. Columns (1) and (2) control for age of lease in a linear manner. Columns (3) and (4) control for age of lease using a spline specification. Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample.

	(1)	(2)	(3)	(4)
		Probability of First Drilling		
Crude Oil Wellhead Price				
(2010 Dollars)	0.000014***	0.000013***	0.000021***	0.000018***
	(0.000004)	(0.000003)	(0.000008)	(0.000004)
Firm Winner	0.000112	0.000080	0.000112	0.000080
	(0.000083)	(0.000068)	(0.000083)	(0.000068)
Nearby Production				
Flag	0.000916***	0.000810***	0.000917***	0.000810***
	(0.000194)	(0.000207)	(0.000194)	(0.000207)
Firm/Nearby Prod				
Interaction	-0.000610***	-0.000496***	-0.000612***	-0.000496***
	(0.000217)	(0.000170)	(0.000218)	(0.000169)
R squared	0.000584	0.000551	0.000626	0.000578
p-value: $\beta_1 + \beta_3 = 0$	0.018616	0.011762	0.018380	0.011492
Observations	291680	1726619	291680	1726619

Table B.46: This table uses a linear probability specification and a survival data framework to examine the probability that drilling happens in a given month. Columns (1) and (2) control for age of lease in a linear manner. Columns (3) and (4) control for age of lease using a spline specification. Columns (1) and (3) use the restricted sample; Columns (2) and (4) use the full sample.