

# Patchwork Policies, Spillovers, and the Search for Oil and Gas

Eric Lewis

## A Online Appendix

### A.1 Extensions and Proofs for Section 3

#### A.1.1 Proofs for Figure 3

I first discuss proofs for Figure 3. I first show why the shape of the indifference curves ( $\pi_1 = \pi_2$ ,  $\pi_1 = 0$ , and  $\pi_2 = 0$ ) do not depend on  $G(\mu_1, \mu_2)$ . Then I show why the indifference curves have the slope they do and why a shift in  $C_1$  down leads to shifts as shown in the graph.

**Proposition:** The shape of indifference lines do not depend on  $G(\mu_1, \mu_2)$ .

*Proof.* A firm begins knowing  $\mu_1$  and  $\mu_2$  as well as  $F(R_1, R_2|\mu_1, \mu_2)$ . Note that profits  $\pi_1$  and  $\pi_2$  in equations 1 and 2 are a function of  $\mu_1$ ,  $\mu_2$ , and  $F$  but not of the distribution  $G$ . Therefore the set of points  $\mu_1$  and  $\mu_2$  where the firm is indifferent between either drilling 1 first, drilling 2 first, and/or not drilling at all will not depend on  $G$ .  $\square$

**Proposition:** The slope of  $\partial\mu_2/\partial\mu_1$  for the indifference curve  $\pi_1 = 0$  has a slope less than  $-1$ .

*Proof.* Because  $E(R_2|\mu_1, \mu_2) = E(R_2|\mu_2) = \mu_2$ , the value of  $\mu_1$  has no effect on the value of  $E(R_2|\mu_1, \mu_2)$ . Therefore, holding  $\mu_2$  fixed,  $\pi_1$  only depends on  $\mu_1$  via the profits of drilling plot 1, and not through the option value of drilling plot 2 second. Therefore

$$(A1) \quad \frac{\partial\pi_1}{\partial\mu_1} = 1$$

Because the probability that plot 2 is drilled is less than 1 but greater than zero:

$$(A2) \quad \frac{\partial\pi_1}{\partial\mu_2} \in (0, 1)$$

Therefore doing total differentiation of the function  $\pi_1(\mu_1, \mu_2(\mu_1), C_1, C_2) = 0$  and rearranging yields:

$$(A3) \quad \left. \frac{\partial \mu_2}{\partial \mu_1} \right|_{\pi_1=0} = \frac{-\partial \pi_1 / \partial \mu_1}{\partial \pi_1 / \partial \mu_2} < -1$$

□

**Proposition:** The slope of  $\partial \mu_2 / \partial \mu_1$  for the indifference curve  $\pi_2 = 0$  has a slope between  $-1$  and  $0$ :

*Proof.*

$$(A4) \quad \frac{\partial \pi_2}{\partial \mu_2} = 1$$

$$(A5) \quad \frac{\partial \pi_2}{\partial \mu_1} \in (0, 1)$$

Therefore doing total differentiation of the function  $\pi_2(\mu_1, \mu_2(\mu_1), C_1, C_2) = 0$  and rearranging yields:

$$(A6) \quad \left. \frac{\partial \mu_2}{\partial \mu_1} \right|_{\pi_2=0} = \frac{-\partial \pi_2 / \partial \mu_1}{\partial \pi_2 / \partial \mu_2} \in (-1, 0)$$

□

**Proposition:** The indifference curve  $\pi_1 = \pi_2$  slopes upward. When  $C_1 = C_2$ , it is the 45 degree line.

*Proof.* Using equations [A1](#), [A2](#), [A4](#), and [A5](#), we can do total differentiation of the equation  $\pi_1(\mu_1, \mu_2(\mu_1), C_1, C_2) = \pi_2(\mu_1, \mu_2(\mu_1), C_1, C_2)$ . Rearranging yields:

$$(A7) \quad \left. \frac{\partial \mu_2}{\partial \mu_1} \right|_{\pi_1=\pi_2} = \frac{[\partial \pi_1 / \partial \mu_1 - \partial \pi_2 / \partial \mu_1]}{[\partial \pi_2 / \partial \mu_2 - \partial \pi_1 / \partial \mu_2]} > 0$$

In the case where  $C_1 = C_2$ , then the functions are symmetric such that equality is achieved only when  $\mu_1 = \mu_2$ .

□

**Proposition:** A reduction in  $C_1$  moves the  $\pi_1 = 0$  indifference curve and the  $\pi_2 = 0$  indifference curve closer to the origin. Holding  $\mu_1$  fixed and

given a decrease in  $C_1$ , the  $\mu_2$  that sets  $\pi_1 = 0$  will increase more than the  $\mu_2$  that sets  $\pi_2 = 0$ .

*Proof.* Because drilling plot 1 first means that the firm always pays  $C_1$  and drilling plot 2 first means that the firm always pays  $C_2$

$$(A8) \quad \frac{\partial \pi_1}{\partial C_1} = -1$$

$$(A9) \quad \frac{\partial \pi_2}{\partial C_2} = -1$$

Because drilling plot 1 first means that there is only some probability that plot 2 is drilled, and because drilling plot 2 first means that there is only some probability that plot 1 is drilled,

$$(A10) \quad \frac{\partial \pi_1}{\partial C_2} \in (-1, 0)$$

$$(A11) \quad \frac{\partial \pi_2}{\partial C_1} \in (-1, 0)$$

Now we examine the shift in the  $\pi_1 = 0$  line as  $C_1$  changes. First the change in  $\mu_1$ , holding  $\mu_2$  fixed. Taking total derivatives of  $\pi_1(\mu_1(C_1), \mu_2, C_1, C_2) = 0$  and rearranging:

$$(A12) \quad \left. \frac{\partial \mu_1}{\partial C_1} \right|_{\pi_1=0, \mu_2} = \frac{-\partial \pi_1 / \partial C_1}{\partial \pi_1 / \partial \mu_1} = -1$$

Second, the change in  $\mu_2$ , holding  $\mu_1$  fixed. Taking total derivatives of  $\pi_1(\mu_1, \mu_2(C_1), C_1) = 0$  and rearranging:

$$(A13) \quad \left. \frac{\partial \mu_2}{\partial C_1} \right|_{\pi_1=0, \mu_1} = \frac{-\partial \pi_1 / \partial C_1}{\partial \pi_1 / \partial \mu_2} < -1$$

Now in contrast, we examine the shift in the  $\pi_2 = 0$  line as  $C_1$  changes. First the change in  $\mu_1$ , holding  $\mu_2$  fixed. Taking total derivatives of  $\pi_2(\mu_1(C_1), \mu_2, C_1, C_2) = 0$  and rearranging:

$$(A14) \quad \left. \frac{\partial \mu_1}{\partial C_1} \right|_{\pi_2=0, \mu_2} = \frac{-\partial \pi_2 / \partial C_1}{\partial \pi_2 / \partial \mu_1} = -1$$

Second, the change in  $\mu_2$ , holding  $\mu_1$  fixed. Taking total derivatives of  $\pi_2(\mu_1, \mu_2(C_1), C_1) = 0$  and rearranging:

$$(A15) \quad \left. \frac{\partial \mu_2}{\partial C_1} \right|_{\pi_2=0, \mu_1} = \frac{-\partial \pi_2 / \partial C_1}{\partial \pi_2 / \partial \mu_2} \in (-1, 0)$$

□

**Proposition:** A reduction in  $C_1$  shifts the  $\pi_1 = \pi_2$  line to the left.

*Proof.* First, holding  $\mu_2$  fixed and take total derivatives for  $\pi_1(\mu_1(C_1), \mu_2, C_1, C_2) = \pi_2(\mu_1(C_1), \mu_2, C_1, C_2)$  and rearranging:

$$(A16) \quad \left. \frac{\partial \mu_1}{\partial C_1} \right|_{\pi_1=\pi_2, \mu_2} = \frac{-[\partial \pi_1 / \partial C_1 - \partial \pi_2 / \partial C_1]}{[\partial \pi_1 / \partial \mu_1 - \partial \pi_2 / \partial \mu_1]} > 0$$

Second, holding  $\mu_1$  fixed and take total derivatives for  $\pi_1(\mu_1, \mu_2(C_1), C_1, C_2) = \pi_2(\mu_1, \mu_2(C_1), C_1, C_2)$  and rearranging:

$$(A17) \quad \left. \frac{\partial \mu_2}{\partial C_1} \right|_{\pi_1=\pi_2, \mu_1} = \frac{[\partial \pi_1 / \partial C_1 - \partial \pi_2 / \partial C_1]}{[\partial \pi_2 / \partial \mu_2 - \partial \pi_1 / \partial \mu_2]} < 0$$

□

### A.1.2 Proofs for ever-drill, production, and production conditional on drilling

For proofs about whether a plot is ever drilled, how much each plot produces (section-level production), and production conditional on drilling (well-level production), I again show that spillover effects are theoretically ambiguous.

I make the restriction that  $\mu_1 = \mu_2$  and  $r_1 = r_2$ . I make this assumption because it significantly simplifies the math. Under this assumption, I show that the spillover effects are not ambiguous when examining whether a site is the site of the initial exploratory well: Proximity to a low-cost plot will decrease the probability that exploratory drilling is done on a high cost plot. However under this assumption, it is still the case that the spillover effects for the other three outcomes—whether a plot is ever drilled, how much a plot produces, and how much a plot produces conditional on drilling—are still theoretically ambiguous.

With a single variable  $\mu$  and a single variable  $R$ , we can write  $\mu \sim G$  and  $F(R|\mu)$ . I assume that  $F$  exhibits first order stochastic dominance with respect to  $\mu$ :  $\partial F(r|\mu) / \partial \mu \leq 0 \forall \mu, r$ . In other words, for any given threshold,

an increase in the signal  $\mu$  implies a decrease in the probability that that true reserves are beneath the threshold.

Re-writing equations [1](#) and [2](#) for this simple case, we have:

$$(A18) \quad \pi_1(\mu, C_1, C_2) = \mu - C_1 + \delta E_{R|\mu}[\max\{R - C_2, 0\}],$$

$$(A19) \quad \pi_2(\mu, C_1, C_2) = \mu - C_2 + \delta E_{R|\mu}[\max\{R - C_1, 0\}],$$

**Proposition:** If  $C_1 < C_2$  then  $\pi_1(\mu, C_1, C_2) > \pi_2(\mu, C_1, C_2)$ :

*Proof.* We can write  $\pi_1$  using indicator functions rather than using the maximum function:

$$\begin{aligned} \pi_1 &= \mu + \int [-C_1 + 1(R - C_2 \geq 0)\delta(R - C_2)]\partial F(R|\mu) \\ &= \mu + \int [-C_2 + (C_2 - C_1) \\ &\quad + [1(R - C_1 \geq 0) - 1(C_1 \leq R < C_2)]\delta(R - C_1 + C_1 - C_2)]\partial F(R|\mu) \\ &= \pi_2 + \int [(C_2 - C_1) + 1(R \geq C_1)\delta(C_1 - C_2) \\ &\quad - 1(C_1 \leq R < C_2)\delta(R - C_2)]\partial F(R|\mu) \\ &= \pi_2 + \int [(C_2 - C_1)[1(R < C_1) + (1 - \delta)1(R \geq C_1)] \\ &\quad + 1(C_1 \leq R < C_2)\delta(C_2 - R)]\partial F(R|\mu) \end{aligned}$$

where the term under the integral is positive. □

**Proposition:** With this model, for any cost profile, there will be a unique threshold values of  $\mu$  that determine whether there is drilling:

*Proof.* To show this, it is only necessary to show that  $\pi_i$  is strictly increasing

in  $\mu$ . Without loss of generality, consider  $\pi_1$ :

$$(A20) \quad \pi_1 = \mu - C_1 + \delta \int_{C_2}^{\infty} (R - C_2) f(R|\mu) \partial R$$

$$(A21) \quad = \mu - C_1 + \delta \int_{C_2}^{\infty} \int_{C_2}^R f(R|\mu) \partial x \partial R$$

$$(A22) \quad = \mu - C_1 + \delta \int_{C_2}^{\infty} \int_x^{\infty} f(R|\mu) \partial R \partial x$$

$$(A23) \quad = \mu - C_1 + \delta \int_{C_2}^{\infty} [1 - F(x|\mu)] \partial x$$

$$(A24) \quad = \mu - C_1 + \delta \int_{C_2}^{\infty} [1 - F(R|\mu)] \partial R$$

Now taking derivatives of  $\pi_1$  with respect to  $\mu$ :

$$\frac{\partial \pi_1}{\partial \mu} = 1 - \int_{C_2}^{\infty} \frac{\partial F(R|\mu)}{\partial \mu} \partial R$$

Because we assumed that  $F(R|\mu)$  is decreasing in  $\mu$ ,  $\partial \pi_1 / \partial \mu$  is unambiguously positive. Therefore  $\pi_1$  is strictly increasing in  $\mu$  and there is a unique threshold value  $\mu_1^*$  that will set  $\pi_1$  equal to zero.  $\square$

**Proposition:**  $\pi_i$  is decreasing in  $C_1$  and  $C_2$  so that therefore  $\mu_i^*$  is increasing in both  $C_1$  and  $C_2$ .

*Proof.* Without loss of generality, consider  $\pi_1$ . Taking derivatives with respect to either  $C_1$  or  $C_2$  in the equation  $\pi_1(\mu^*(C_1, C_2), C_1, C_2) = 0$ , we have:

$$\frac{\partial \mu_1^*}{\partial C_i} = \frac{-\partial \pi_1 / \partial C_i}{\partial \pi_1 / \partial \mu}$$

$\partial \pi_1 / \partial C_i < 0$  for both  $i = 1$  and  $2$ .  $\partial \pi_1 / \partial \mu > 0$ . Therefore  $\mu^*$  is increasing in both  $C_1$  and  $C_2$ .  $\square$

Therefore the threshold value  $\mu^*$  for the  $C_S, C_F$  case will be less than the threshold value for the  $C_F, C_F$ , case.

**Probability of ever drilling:** Now I examine predicted probability of ever drilling.

If both plots are federal land, the threshold is  $\mu^*(C_F, C_F)$ . In that case, the firm will drill if  $\mu \geq \mu^*(C_F, C_F)$ . As it is indifferent between the two plots, it will randomize the probability that it drills each one with probability .5.

If plot 1 is in state ownership and plot 2 is in federal ownership, then the threshold value is  $\mu^*(C_S, C_F)$ . In such a case, the firm will first drill on state land and then decide whether to drill a follow-up well on federal land.

If plot 1 is in state ownership and plot 2 is in federal ownership, then plot 2 can only be drilled second, so that the probability that plot 2 is ever drilled is:

$$(A25) \quad p_2(C_S, C_F) = \int_{\mu_1^*(C_S, C_F)}^{\infty} (1 - F(C_F|\mu)) \partial G(\mu)$$

If both plots are in federal ownership, the probability that plot 2 is ever drilled is:

$$(A26) \quad p_2(C_F, C_F) = 0.5[1 - G(\mu^*(C_F, C_F))] + 0.5 \int_{\mu^*(C_F, C_F)}^{\infty} (1 - F(C_F|\mu)) \partial G(\mu)$$

The first term is the probability that plot 2 is the initial exploratory well; the second term is the probability that plot 2 is drilled in period 2.

**Proposition:** The spillover effect is ambiguous.

*Proof.* The difference in probabilities between equation [A25](#) and equation [A26](#) is:

$$(A27) \quad p_2(C_S, C_F) - p_2(C_F, C_F) = \int_{\mu_1^*(C_S, C_F)}^{\mu_1^*(C_F, C_F)} (1 - F(C_F|\mu)) \partial G(\mu) - \int_{\mu_1^*(C_F, C_F)}^{\infty} \frac{1}{2} F(C_F|\mu) \partial G(\mu)$$

The first term is clearly positive. It represents the total expected cost effect. Proximity to state land means that the threshold  $\mu^*$  is lower, which means for values of  $\mu$  between  $\mu^*(C_S, C_F)$  and  $\mu^*(C_F, C_F)$ , there is now a probability that plot 2 is drilled because plot 1 will be the site of the initial exploratory well and plot 2 may have a follow-up well.

The second term is clearly negative. This is the substitution effect. For values of  $\mu$  above  $\mu^*(C_F, C_F)$ , the plot 2 is less likely to be drilled if the neighbor is state land because it will not be the site of the exploratory well.  $\square$

**Production (unconditional on drilling):** Now I examine production, unconditional on drilling. **Proposition:** Average production of state plots will

be highest. It is ambiguous whether federal plots close to state land or far from state land will have higher production.

*Proof.* For state land, expected production is:

$$(A28) \quad \int_{\mu_1^*(C_S, C_F)}^{\infty} \mu \partial G(\mu)$$

For federal land close to state land, expected production is:

$$(A29) \quad \int_{\mu_1^*(C_S, C_F)}^{\infty} \int_{C_F}^{\infty} R \partial F(R|\mu) \partial G(\mu)$$

It is clear that [A29](#) is less than [A28](#) because the inner integral in [A29](#) is less than  $\mu$ .

Expected production for federal land far from state land is an average of the above two equations but using the higher signal  $\mu_1^*(C_F, C_F)$ :

$$(A30) \quad 0.5 \cdot \int_{\mu_1^*(C_F, C_F)}^{\infty} \mu \partial G(\mu) + 0.5 \cdot \int_{\mu_1^*(C_F, C_F)}^{\infty} \int_{C_F}^{\infty} R \partial F(R|\mu) \partial G(\mu)$$

[A30](#) is smaller than [A28](#) because the federal land far from state land has a higher  $\mu_1^*$  and also for the same reason that [A29](#) is less than [A28](#).

It is ambiguous whether expected production for federal land close to or far from state land will be higher. The difference between [A29](#) and [A30](#) is:

$$(A31) \quad \int_{\mu_1^*(C_S, C_F)}^{\mu_1^*(C_F, C_F)} \int_{C_F}^{\infty} R \partial F(R|\mu) \partial G(\mu) + 0.5 \int_{\mu_1^*(C_S, C_F)}^{\mu_1^*(C_F, C_F)} \left[ -\mu + \int_{C_F}^{\infty} R \partial F(R|\mu) \right] \partial G(\mu)$$

The second term is negative because  $\int_{C_F}^{\infty} R \partial F(R|\mu) < \mu$ . As the distribution  $G$  is unknown, we cannot determine whether [A31](#) is positive or negative. The intuition is similar to that of equation [A27](#) which examines whether federal land close or far from state land will have more drilling. The substitution effect means that federal land close to state land will have a lower  $\mu^*$  and so greater likelihood of drilling low productivity wells. But the substitution effect decreases the drilling and therefore the production on federal land close to state land.

□

**Helpful lemma:** Before examining production conditional on drilling (e.g., well-level production), I first prove a helpful lemma.

Lemma: If  $0 \leq a(\mu) \leq \min\{b(\mu), c(\mu)\} \leq \max\{b(\mu), c(\mu)\} < 1$ ,  $a(\mu) < b(\mu)c(\mu)$ ,  $b'(\mu) \geq 0$ , and  $c'(\mu) \leq 0$ , then:

$$(A32) \quad \frac{\int_x^y a(\mu) \partial G(\mu)}{\int_x^y b(\mu) \partial G(\mu)} < \frac{\int_x^y c(\mu) \partial G(\mu)}{\int_x^y \partial G(\mu)}$$

*Proof.* Because  $a(\mu) < b(\mu)c(\mu)$ , we have:

$$(A33) \quad \int_x^y a(\mu) \partial G(\mu) < \int_x^y b(\mu)c(\mu) \partial G(\mu)$$

$$(A34) \quad \leq \left[ \int_x^y b(\mu) \partial G(\mu) \right] \left[ \int_x^y c(\mu) \partial G(\mu) \right]$$

$$(A35) \quad \leq \frac{\left[ \int_x^y b(\mu) \partial G(\mu) \right] \left[ \int_x^y c(\mu) \partial G(\mu) \right]}{\int_x^y \partial G(\mu)}$$

The first inequality follows by definition. The second inequality follows because  $b(\mu)$  is increasing in  $\mu$  while  $c(\mu)$  is decreasing. As the two are negatively correlated, the expectation of the product is less than the product of the expectations (see [Tao and Vu \(2006\)](#), page 20 for a formal proof). Finally, the last line follows because  $G(\mu)$  is a cdf—increasing in  $\mu$  and bounded above by one and below by zero.

Note: This proof also holds if  $b'(\mu) \leq 0$ , and  $c'(\mu) \geq 0$  □

**Production conditional on drilling:** Given this lemma, average well production (production conditional on drilling) will be lowest on state land. It is ambiguous whether wells on federal land close to state land or federal wells far from state land will have higher production.

*Proof.* When examining the distribution of production conditional on drilling, it is helpful to look specifically at the distribution of production conditional on a well being the first to be drilled versus the second to be drilled. And I use costs  $C_L$  and  $C_H$ , where one plot has cost  $C_L$  and the other has cost  $C_H$ , where the plot with cost  $C_L$  is drilled first and  $C_L \leq C_H$ . This accommodates both the  $\{C_S, C_F\}$  and the  $\{C_F, C_F\}$  cases. I will simply use  $\mu^*$  to refer to the threshold. And rather than solving for expected production, I instead focus on the cdf.

First, consider the cdf of production for the first well drilled:

$$(A36) \quad p(R \leq x | \text{first drilled}) =$$

$$(A37) \quad \frac{p(R \leq x, \text{first drilled})}{p(\text{first drilled})} =$$

$$(A38) \quad \frac{\int_{\mu^*}^{\infty} F(x|\mu) \partial G(\mu)}{1 - G(\mu^*)}$$

Now consider the cdf of production for the second well drilled:

$$\begin{aligned} p(R \leq x | \text{second drilled}) &= \\ \frac{p(R \leq x, \text{second drilled})}{p(\text{second drilled})} &= \\ \frac{\int_{\mu^*}^{\infty} 1(x \geq C_H) [F(x|\mu) - F(C_H|\mu)] \partial G(\mu)}{\int_{\mu^*}^{\infty} (1 - F(C_H|\mu)) \partial G(\mu)} & \end{aligned}$$

From here we can use the lemma from above. Defining:

$$\begin{aligned} a(\mu) &= 1(x \geq C_F) [F(x|\mu) - F(C_F|\mu)] \\ b(\mu) &= 1 - F(C_F|\mu) \\ c(\mu) &= F(x|\mu) \end{aligned}$$

and setting the lower bound of integration as  $x = \mu^*$  and the upper bound of integration as  $y = \infty$ . It is easy to see that the conditions of the Lemma in [A.1.2](#) hold. Therefore

$$(A39) \quad p(R \leq X | \text{first drilled}) > p(R \leq X | \text{second drilled})$$

This implies that the expected production for state wells will be lower than expected production for federal wells that are close to state land. It also implies that for the case of federal land far from state land, federal wells that are drilled first will have lower expected production than federal wells that are drilled second.

Also, we can compare what happens with expected production for the initial well conditional on drilling as we increase  $\mu^*$ , e.g., as we move from the  $\{C_S, C_F\}$  to the  $\{C_F, C_F\}$  case. Taking the derivative of the log of the term

in equation [A38](#) with respect to  $\mu^*$  gives:

$$(A40) \quad \frac{\partial \log p(R \leq x | \text{drilled first})}{\partial \mu^*} = \frac{g(\mu^*) \int_{\mu^*}^{\infty} [F(x|\mu) - F(x|\mu^*)] \partial G(\mu)}{\left[ \int_{\mu^*}^{\infty} F(x|\mu) \partial G(\mu) \right] \left[ \int_{\mu^*}^{\infty} \partial G(\mu) \right]} < 0$$

The inequality follows because  $F(x|\mu) < F(x|\mu^*)$  for all  $\mu > \mu^*$ .

This implies that the expected productivity of a well on state land, which is always exploratory, is lower than expected productivity of exploratory wells on federal land that is far from state land. Furthermore, as the expected productivity of an exploratory well on federal land far from state land is less than the expected productivity of a secondary well on federal land that is far from state land, this means that expected productivity of a well on state land is less than the expected productivity of a secondary well that is drilled on federal land far from state land.

Finally, the cdf of productivity of a federal well on federal land that is far from state land is going to be a convex combination of the cdf of productivity if the well is the initial exploratory well and if the well is the second well:

$$(A41) \quad p(R \leq x | \text{drilled}) = \frac{p(R \leq x, \text{drilled})}{p(\text{drilled})} =$$

$$(A42) \quad \frac{p(R \leq x, \text{is first plot drilled}) + p(R \leq x, \text{is second plot drilled})}{p(\text{is first plot drilled}) + p(\text{is second plot drilled})} =$$

$$(A43) \quad \frac{0.5 \left[ \int_{\mu_1^*(C_F, C_F)}^{\infty} F(x|\mu) \partial G(\mu) \right] + 0.5 \left[ \int_{\mu_1^*(C_F, C_F)}^{\infty} 1(x \geq C_F) [F(x|\mu) - F(C_F|\mu)] \partial G(\mu) \right]}{0.5 \left[ (1 - G(\mu_1^*(C_F, C_F))) \right] + 0.5 \left[ \int_{\mu_1^*(C_F, C_F)}^{\infty} (1 - F(C_F|\mu)) \partial G(\mu) \right]}$$

Therefore expected productivity of a state well is less than the expected productivity of a federal well that is far from state land.

However it is ambiguous whether a well on federal land that is close to state land or a well on federal land that is far from state land has higher expected production. Taking the ratio of the cdf for secondary wells on federal land close to state land in equation [A39](#) with the cdf of federal wells far from state land in equation [A43](#), for the case where  $x > C_F$  gives:

$$(A44) \quad \frac{p(R \leq X|\text{fed}, \{C_S, C_F\})}{p(R \leq X|\text{fed}, \{C_F, C_F\})} =$$

$$(A45) \quad \frac{\int_{\bar{\mu}^*}^{\infty} [1 - \frac{1}{2}F(C_F|\mu)] \partial G(\mu)}{\int_{\underline{\mu}^*}^{\infty} [1 - F(C_F|\mu)] \partial G(\mu)} \cdot \frac{\int_{\underline{\mu}^*}^{\infty} [F(x|\mu) - F(C_F|\mu)] \partial G(\mu)}{\int_{\bar{\mu}^*}^{\infty} [2F(x|\mu) - F(C_F|\mu)] \partial G(\mu)}$$

where  $\bar{\mu}^* = \mu_1^*(C_F, C_F)$  and  $\underline{\mu}^* = \mu_1^*(C_S, C_F)$ .

It is easy to see why it is ambiguous whether this term is less than or greater than 1. The first term in [A45](#) is ambiguous because  $1 - \frac{1}{2}F(C_F|\mu) > 1 - F(C_F|\mu)$  but  $\bar{\mu}^* > \underline{\mu}^*$ . Similarly,  $2F(x|\mu) > F(x|\mu)$  but  $\underline{\mu}^* < \bar{\mu}^*$ . As we don't know the distribution of  $G$ , we cannot determine which dominates.  $\square$

## A.2 Primary Terms and Delays: Theory

Here I discuss the role of heterogeneous primary terms and delays. I show that if state land has shorter primary terms than federal land, proximity of state land to federal land will tend to decrease drilling on federal land. I next show that if federal land is more likely to have delays, proximity to state land will tend to reduce drilling on federal land. Finally, I examine a model with both primary terms and delays to show that they can result in an effect similar to the total expected cost effect.

### A.2.1 Primary Terms

Because federal leases have primary terms of 10 years whereas state leases have primary terms of 5 years, it is more likely that a state lease will expire before a federal lease, and less likely that a federal lease will expire before a state lease. Therefore in this model of primary terms, there will be a longer period of time for a firm to drill on federal land and a shorter period of time for a firm to drill on state land.

To model the spillover effects of shorter state primary terms, I assume that state and federal land have identical fixed costs. This assumption is important for demonstrating the effect of primary terms while holding all else constant. To model primary terms as simply as possible, I assume that state leases can only be drilled in period 1, whereas federal leases can be drilled in either period 1 or period 2.<sup>30</sup>

<sup>30</sup>The reason for this modeling choice is that because firms discount the future, and

Figure [A1](#) shows drilling choices in the federal-federal and state-federal regime. The lines  $\overline{ab}$ ,  $\overline{bc}$ , and  $\overline{bd}$  are identical to the lines  $\overline{ab}$ ,  $\overline{bc}$ , and  $\overline{bd}$  in Figure [3](#), and represent indifference curves for the federal-federal case. In the federal-federal case, the firm will drill first on plot 1 if signals are in B, will drill plot 2 first if the signals are in D, E, or F, and will not drill at all if the signals are in A.

The lines  $\overline{af}$  and  $\overline{ec}$  represent indifference lines for the state-federal case, where plot 1 is in state ownership and can only be drilled in the first period. The line  $\overline{ec}$  represents the indifference curve  $\pi_1 = 0$  (where  $\pi_1$  is defined as in Equation [1](#)). If the signals are to the right of this line (in B or D), the firm drills on state land first. The line  $\overline{af}$  represents the indifference curve  $\mu_2 - \text{cost} = 0$ . The firm will drill plot 2 first if the signals are above this line (in E). The firm will not drill at all if the signals are in A or F.

The spillover effects of state land are two fold, and both reduce the likelihood of drilling on federal land. First, proximity to state land can may lead the firm to drill first on federal land rather than drilling first on state land, reducing the likelihood that federal land is drilled (region D). Second, proximity to state land means that there is no learning value to the federal plot first, as it cannot follow up with drilling the state plot in period 2. Therefore this further decreases the likelihood that the federal plot is drilled (region F).

## A.2.2 Delays

Now to model delays, suppose that if a firm wants to drill on federal land, it cannot drill on federal land until period 2 or later. Furthermore, suppose that there are three periods of time, such that the firm can drill on a state plot in periods 1, 2, and 3, and that the firm can drill on federal land in periods 2 and 3. Therefore in spite of the delays on federal land, the delays do not inhibit the firm's ability to drill sequentially and learn. As in the previous subsection, suppose that the fixed costs of drilling on federal and state land are identical.

Denoting  $\pi_i^{FF,D}$  as the profits for drilling plot  $i$  first given that we are in this delay framework and given the federal-federal regime. Denote  $\pi_i^{SF,D}$  as the profits for drilling plot  $i$  first given that we are in this delay framework

---

because there is no uncertainty, there is no value to wait to drill other than to wait to learn information from the outcome of drilling the other plot. Therefore assuming very limited number of periods is necessary for short state leases to have any effect on drilling outcomes. If the model was extended to allow for future price uncertainty, as in [Kellogg \(2014\)](#), the model could be extended to have longer periods of time. In such a model with price uncertainty, shorter primary terms for state leases would lead to a similar effect as in this basic lease length model that has no price uncertainty.

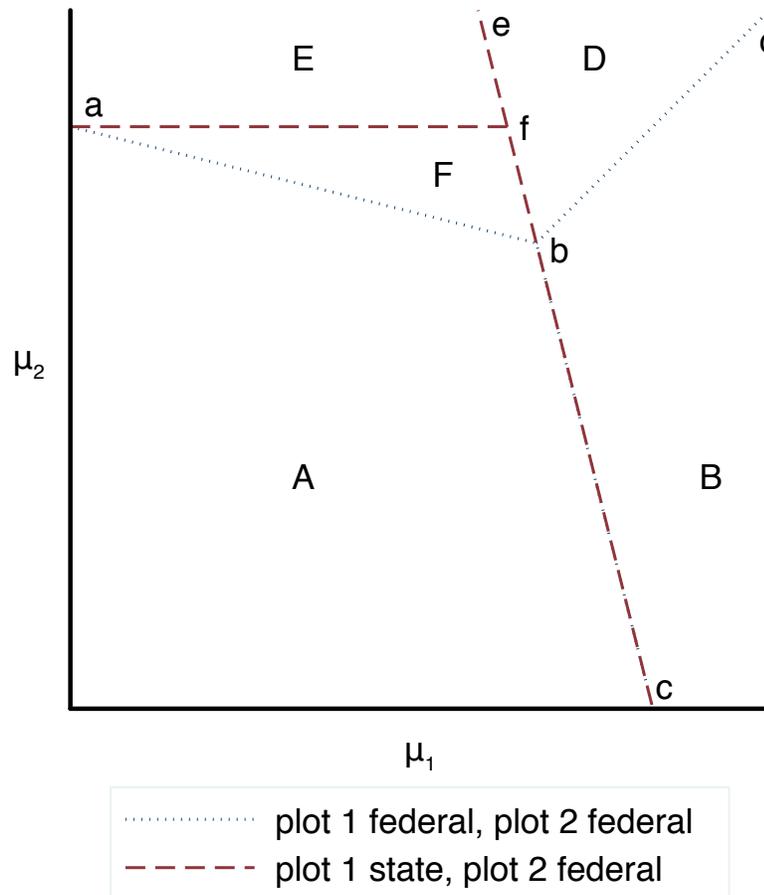


Figure A1: Diagram of optimal first-period drilling location choice as a function of expected production  $\mu_1$  and  $\mu_2$  under two different primary term regimes. The first regime is the case of federal-federal ownership, where both plots have 2 periods to drill. The second is the case where plot 1 is state-owned and the lease only lasts 1 period, while the other is the case with a federal lease lasting two periods. In region A, the firm never drills. In region B, the firm drills plot 1 first in both regimes. In region D, the firm drills plot 2 first in the federal-federal regime but plot 1 first in the state-federal regime. In region E, the firm drills plot 2 first and never drills plot 1. In region F, the firm drills plot 2 first in the federal-federal regime but does not drill at all in the state-federal regime.

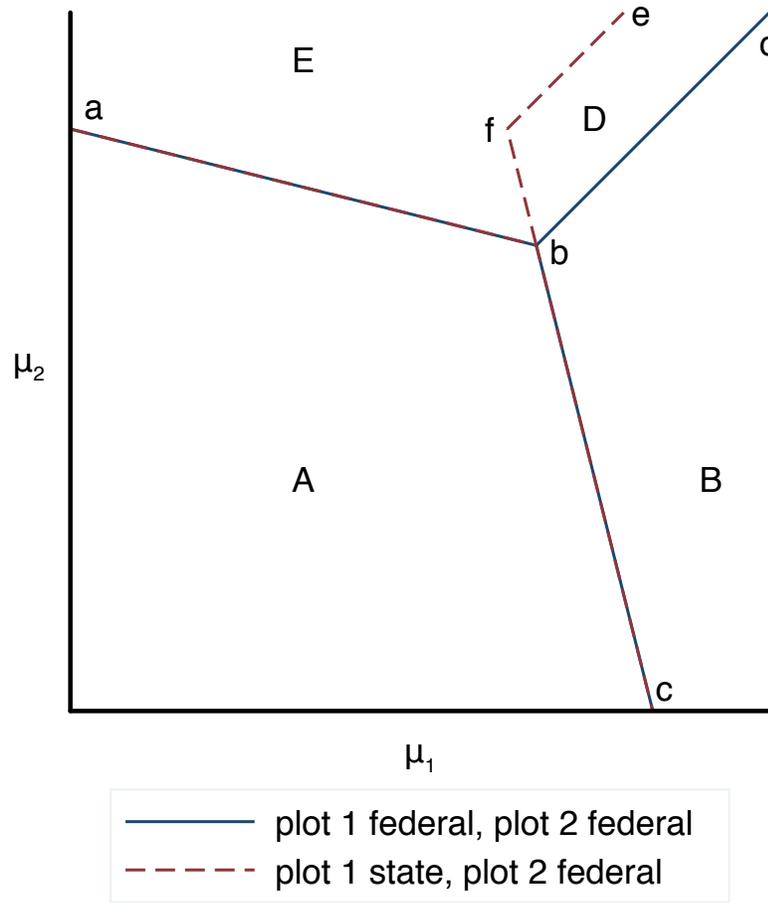


Figure A2: Diagram of optimal first-period drilling location choice as a function of expected production  $\mu_1$  and  $\mu_2$  under the case where federal land has delays. The first regime is the case of federal-federal ownership, where the firm cannot drill in period 1 for either plot. The second regime is the cases of state-federal ownership, where the firm cannot drill in period 1 for plot 2, but can drill plot 1 in any time period. In region A, the firm never drills. In region B, the firm drills plot 1 first in both regimes. In region D, the firm drills plot 2 first in the federal-federal regime but plot 1 first in the state-federal regime. In region E, the firm drills plot 2 first in both regimes.

and given the state-federal regime. We can relate these profits to the profits in the main model, as given in equations [1](#) and [2](#). For each of these, fixed costs are identical regardless of regime (e.g.,  $C_1 = C_2$ ). Therefore I exclude the cost arguments from the function.

$$(A46) \quad \pi_1^{FF,D}(\mu_1, \mu_2) = \delta\pi_1(\mu_1, \mu_2)$$

$$(A47) \quad \pi_2^{FF,D}(\mu_1, \mu_2) = \delta\pi_2(\mu_1, \mu_2)$$

$$(A48) \quad \pi_1^{SF,D}(\mu_1, \mu_2) = \pi_1(\mu_1, \mu_2)$$

$$(A49) \quad \pi_2^{SF,D}(\mu_1, \mu_2) = \delta\pi_2(\mu_1, \mu_2)$$

$$(A50)$$

Note that in the federal-federal case, there is no difference between the main model in Section [3](#) the model with delays here as the three indifference curves  $\pi_1^{FF,D} = 0$ ,  $\pi_2^{FF,D} = 0$ , and  $\pi_1^{FF,D} = \pi_2^{FF,D}$  do not change.

However, in the state-federal case, while the indifference lines  $\pi_1^{SF,D} = 0$  and  $\pi_2^{SF,D} = 0$  are the same as in the main model case, the indifference line  $\pi_1^{SF,D} = \pi_2^{SF,D}$  is shifted to the left relative to the  $\pi_1 = \pi_2$  indifference curve of the main model. This is because in cases where expected plot 1 reserves are slightly lower than expected plot 2 reserves, the firm prefers to drill earlier rather than getting the maximum expected profits one period in the future.

This is illustrated in Figure [A2](#) which graphs indifference curves. In region A the firm never drills, regardless of regime. In region B the firm always drills plot 1 first, regardless of regime. In region E the firm drills plot 2 first, regardless of regime. In region D, however, the firm drills plot 2 first if it is in the federal-federal regime but drills plot 1 first if it is in the state-federal regime.

Delays on federal land imply that proximity to state land has a substitution-like effect on federal land. Delays on federal land and proximity to state land imply that state land will get drilled first, which decreases the likelihood that federal land is ever drilled. There is no effect similar to the total expected cost effect in this model, however, as adding delays does not shift the  $\pi_1 = 0$  and the  $\pi_2 = 0$  lines.

### A.2.3 Delays with limited primary terms

In the previous model we assumed delays but that delays did not impede the firm from drilling sequentially and learning in whatever order it wanted. However, here I examine a setting where limited primary terms plus delays impede the ability of the firm to drill sequentially. In particular, I assume

the extreme case where for state land, the limited primary term on state land means the firm can only drill there in period 1. I also assume delays on federal land and a 2 period primary term for federal land such that the firm can only drill on federal land in period 2.

If the firm is in the federal-federal regime, the firm can only drill in period 2. Therefore it cannot drill sequentially to learn, and therefore it only drills on a plot if expected reserves exceed costs. It will drill both plots simultaneously in period 2 if  $\mu_1 \geq \text{cost}$  and  $\mu_2 \geq \text{costs}$ . It will drill only plot 1 in period 2 if  $\mu_1 \geq \text{cost}$  and  $\mu_2 < \text{costs}$ . Similarly, it will drill only plot 2 in period 2 if  $\mu_1 < \text{cost}$  and  $\mu_2 \geq \text{costs}$ . And finally it will not drill at all if both  $\mu_1$  and  $\mu_2$  are less than cost.

In contrast, if the firm is in the state-federal regime, the firm can drill plot 1 first and then learn about the profitability of drilling plot 2 in period 2. This has two effects on the likelihood that federal plot 2 is ever drilled. On the one hand, in cases where plot 2 would not be drilled in the federal-federal regime, the state-federal regime can lead to drilling on plot 1 which then leads to follow-up drilling on plot 2. This is similar to the expected cost effect. On the other hand, in cases where plot 2 would be drilled in the federal-federal regime, the increased profitability of sequential drilling implies that the firm will first drill plot 1 before deciding whether to drill plot 2—and in some cases plot 2 does not end up getting drilled. This is similar to the substitution effect. Therefore, proximity to state land has an ambiguous effect on whether federal plot 2 is ever drilled. However it unambiguously decreases the probability that federal plot 2 is drilled first.

### A.3 Common Pools: Theory

Here I discuss how to incorporate common pools into the theory model. Later on, in Appendix subsection [A.6](#), I discuss evidence on the extent to which common pools may have existed in the GGRB.

Incorporating common pools into the model can be done by adding a scalar  $\alpha \in [0, 1]$ , where  $\alpha$  represents the fraction of reserves from the nearby plot that the first well can extract. Larger values of  $\alpha$  imply larger common pool problems. Adding in common pools, we can modify equation [1](#) and write the profits of drilling on plot 1 first as

$$(A51) \quad \pi_1(\mu_1, \mu_2, C_1, C_2) = \mu_1 + \alpha\mu_2 - C_1 + \delta E_{R_1|\mu_1, \mu_2}[\max\{(1-\alpha)E_{R_2|r_1, \mu_1, \mu_2}(R_2) - C_2, 0\}]$$

Increasing the value of  $\alpha$  reduces the profitability of follow-up drilling,

which in turn decreases the likelihood that the second plot is drilled. In the extreme case of  $\alpha = 1$ , the substitution effect totally dominates: Proximity to state land leads to zero drilling and production on nearby federal land.

## A.4 Multiple firms: Theory

I examine a model of multiple firms. My model is based on [Hendricks and Kovenock \(1989\)](#) (hereafter HK), who examine a model where two firms have the rights to drill on two adjoining plots. If one firm drills at the beginning, it reveals the productivity of the region to the other firm, who can then drill. Firms face a trade-off between drilling today and waiting to see if the other firm will drill and reveal information about the profitability of drilling.

I extend HK by adding in asymmetric costs, such that I can compare the federal-federal case with the state-federal cost. I also make some stronger distributional assumptions on top of HK that allow me to more easily demonstrate why spillover effects are ambiguous.

There are two firms, 1 and 2, who have rights to drill on plots 1 and 2 respectively. Each firm  $i$  receives a signal of the value of underlying reserves  $\mu_i$ . While HK allows these to differ, I make the restriction that  $\mu_1 = \mu_2$ . This allows me to ignore one complicated feature of HK's model: Examining how each firm updates its beliefs about the other firm's signal based on the other firm's action. As with HK, I assume that the underlying reserves under each plot are identical. I use  $R$  to denote the unknown reserves from pre-drilling and  $r$  to denote the actual reserves which are revealed by drilling. Costs of drilling on each plot  $i \in \{1, 2\}$  are  $C_i$ , where  $C_i \in \{C_S, C_F\}$ , and  $C_S < C_F$ .

I make some additional distributional assumptions: First, I only allow for two possible signals,  $\underline{\mu}$  and  $\bar{\mu}$ , where  $\underline{\mu} < \bar{\mu}$ . The probability of each signal  $\mu$  is  $g(\mu)$ , where  $g(\underline{\mu}) + g(\bar{\mu}) = 1$ . Second, I restrict that  $R$  is a binary variable that either takes the value 0 or 1 and where  $E(R|\mu) = \mu$ . Third, I focus on the interesting case where  $C_S < \underline{\mu} < C_F < \bar{\mu}$ . With these assumptions, I can easily show that the spillover effect of state land will be positive in the  $\underline{\mu}$  case and negative in the  $\bar{\mu}$  case. Therefore the overall effect will depend on the function  $g(\cdot)$  which depends on geology and is unknown to the econometrician.

To solve for the Nash equilibrium, I solve the model separately for each possible value of  $\mu$ :

**The case where  $\mu = \underline{\mu}$ :** In examining the case where  $\mu = \underline{\mu}$ , first consider the federal-federal setting, where both firm 1 and firm 2 are on federal land. For both, drilling in the first period yields negative profits as  $\underline{\mu} - C_F < 0$ . Therefore neither firm will drill in the first period. Because neither firm will reveal the true productivity, neither firm will drill in the second period.

Therefore the probability of any drilling is zero.

Next, consider the state-federal setting, where firm 1 is on state land and firm 2 is on federal land. Firm 2 will not find it profitable to drill in the first period but firm 1 will, because it gets profits  $\underline{\mu} - C_S > 0$ . Firm 1 knows that firm 2 will not drill initially, and so it has no incentive to wait to drill, and so drills in the first period. With probability  $\underline{\mu}$ ,  $r = 1$ , and firm 2 finds it profitable to drill in the second period. Therefore the probability of drilling on federal plot 2 is  $\underline{\mu}$ .

Therefore in the case of  $\mu = \underline{\mu}$ , there is an effect similar to the total expected cost effect: Proximity to state land increases the likelihood of drilling on nearby federal land.

**The case where  $\mu = \bar{\mu}$ :** Now with this high signal, both firms are willing to drill in the first period but both know that there is some probability that the other firm will drill in the first period, revealing underlying productivity, which gives each an incentive to wait. Denote  $\gamma_j$  to be the probability that firm  $j$  drills in period 1. Firm  $i$  compares the profits of drilling today ( $\bar{\mu} - C_i$ ) with the profits of waiting, which are a weighted average of what happens if the other firm drills and reveals the true productivity and if it doesn't, in which case firm  $i$  will drill. Firm  $i$ 's profits for waiting rather than drilling in the first period are:

$$(A52) \quad \delta [\gamma_j \bar{\mu} (1 - C_i) + (1 - \gamma_j) (\bar{\mu} - C_i)]$$

Assuming that that  $C_S, C_F, \bar{\mu}, \delta$  are set such that there is an interior solution where the firm is indifferent between drilling today and waiting and therefore randomizes, the mixed strategy equilibrium will lead the firm  $i$  to randomize whether it drills in the first period or waits, with probabilities  $\gamma_i$ :

$$(A53) \quad \gamma_i(C_j) = \frac{C_j(1 - \bar{\mu})}{(1 - \delta)(\bar{\mu} - C_j)}$$

Notice that the probability that one drills in the first period is a function of the other's costs, but not of one's own costs. Taking derivatives of  $\gamma_i$  with respect to  $C_j$ :

$$(A54) \quad \frac{\partial \gamma_i}{\partial C_j} = \frac{(1 - \bar{\mu})\bar{\mu}}{(1 - \delta)(\bar{\mu} - C_j)^2} > 0$$

Therefore, lower costs for firm 1 ( $C_S$  rather than  $C_F$ ) will make firm 2 less likely to drill in the first period.

The probability that agent  $i$ 's plot ever has drilling on it is:

$$(A55) \quad \gamma_i + (1 - \gamma_i)\gamma_j\bar{\mu}$$

Therefore in the case of  $\mu = \bar{\mu}$ , a substitution-like effect dominates. Proximity to low-cost state land decreases the likelihood of drilling on high-cost federal land.

**Ambiguous effects:** Because proximity to state land increases the likelihood of drilling on federal land when the signal is low but decreases it when the signal is high, the overall effect of proximity to state land depends on the probabilities of the high and low signals.

In the federal-federal case, drilling can only happen in the high signal state. The probability that a federal plot is drilled in the federal-federal case is:

$$(A56) \quad g(\bar{\mu}) \cdot [\gamma(C_F) + (1 - \gamma(C_F))\gamma(C_F)\bar{\mu}]$$

In the state-federal case, drilling may happen on federal land in both the high and the low signal states. The probability that a federal plot is drilled in the state-federal case is:

$$(A57) \quad g(\underline{\mu}) \cdot \underline{\mu} + g(\bar{\mu}) \cdot [\gamma(C_S) + (1 - \gamma(C_S))\gamma(C_F)\bar{\mu}]$$

Whether the term in line [A56](#) or line [A57](#) is larger depends on the value of  $g(\bar{\mu})$ . If  $g(\bar{\mu})$  is large such that line [A56](#) is larger, a substitution-like effect will dominate. If  $g(\bar{\mu})$  is small such that line [A57](#) is larger, an effect similar to the total expected cost effect will dominate.

## A.5 Data and Data Construction

I take data from a number of sources:

- Wyoming Oil and Gas Conservation Commission (WOGCC, <http://wogcc.state.wy.us/>). This is the source of raw well data, well-level production data, and data on first production.
- USGS cleaned WOGCC well-level data. Cited in the Works Cited as [Biewick \(2011\)](#). Contains well data for up to 2010. I use this data because it determines drilling dates when the raw WOGCC well data is ambiguous. Available online at <http://certmapper.cr.usgs.gov/data/wlci/spatial/shape/WYwells2010shapefile.zip>. Documentation at <http://pubs.usgs.gov/ds/625/Documentation/wywells2010.htm>

- Bureau of Land Management: Source of federal surface and mineral ownership data. Available online at [http://www.blm.gov/wy/st/en/resources/public\\_room/gis/datagis/state/state-own.html](http://www.blm.gov/wy/st/en/resources/public_room/gis/datagis/state/state-own.html).
- State of Wyoming: Source of shapefile data on state mineral ownership. Not available online.
- Energy Information Administration (<http://www.eia.gov/>). Provides shapefiles of the boundaries of the Greater Green River Basin. Also source of oil and gas field shapefile data referenced in footnote 4 that give estimates of the fraction of oil and gas fields that are in federal ownership.
- Wyoming Geospatial Hub. Source of public land survey shapefile of township, ranges and sections. To download: <http://gospatialhub.org/geohubdata/data/location/sections24k.zip>. Also the source of data on municipalities, settlements, fences, crops, wetlands, irrigation, mining, etc.

I match wells to section by using the reported township, range and section in the well file. In some cases wells have horizontal segments where the bottomhole section is different from the tophole. In those cases I use the bottomhole location as the location of the well. This is because some places have restrictions on surface drilling, such that the tophole is on a different section than the section that the well is targeting. Drilling in the GGRB is dominated by vertical wells.

I limit the analysis to sections that are in Wyoming and that are in the Greater Green River Basin, where the Greater Green River Basin boundaries are defined by the EIA as in figure A3

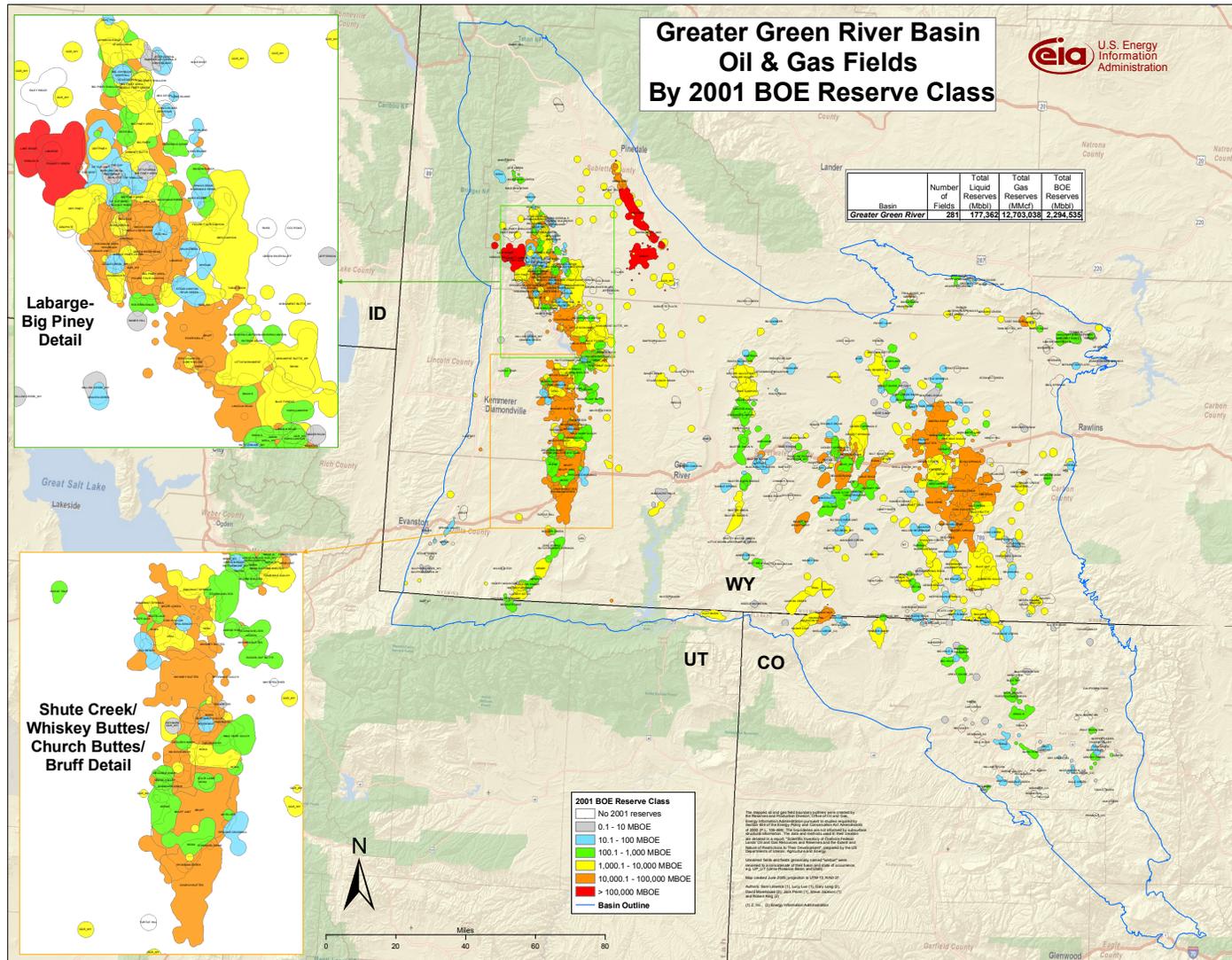


Figure A3: Map of the Greater Green River basin, from the Energy Information Administration. Available online at [https://www.eia.gov/maps/images/field\\_maps/GGR\\_BOE.pdf](https://www.eia.gov/maps/images/field_maps/GGR_BOE.pdf)

## A.6 Additional Empirical Results

This section contains additional empirical results.

### A.6.1 Regression Version of Figure 6

Table A1 is a seemingly unrelated regression of the probabilities of drilling by any year  $t$ , for  $t$  in 10 year increments from 1930 to 2010. It is the regression analog of Figure 6.

### A.6.2 Graphs of Coefficients

Figures A4, A5, A6, and A7 contain graphs of coefficient estimates that are shown in the tables of the main paper.

### A.6.3 Well-level Production Robustness

Tables A2 and A3 show that well-level production results in Table 8 are robust to the length of time that initial production is measured.

### A.6.4 Well Production Declines

Tables A4 and A5 show regressions examining how well production declines differ by location of well. They are regressions where the dependent variable is the log of one year percentage decline in production. These regressions find that decline curves for wells on 16/36 sections are about the same as for wells on sections 3 miles away from 16/36 sections. Declines are less steep for wells on sections that 1 mile from 16/36 sections. This implies that there are not faster rates of decline for wells that are on or close to 16/36 sections. That being said, these regression results find that decline curves vary significantly, with typically steeper decline on the sections that are  $\sqrt{2}$  and  $\sqrt{5}$  miles from 16/36 sections. Therefore this regression does not show evidence of common pools.

	(1) 16/36	(2) $\approx 1$ mile	(3) $\approx \sqrt{2}$ miles	(4) $\approx 2$ miles	(5) $\approx \sqrt{5}$ miles	(6) intercept	(7) joint	(8) non 16/36	(9) 16/36- $\sqrt{5}$
1930	0.002 (0.004)	-0.004 (0.005)	-0.005 (0.005)	-0.002 (0.005)	-0.003 (0.004)	0.018 (0.009)	0.213	0.177	0.000
1940	0.005 (0.005)	-0.003 (0.005)	-0.005 (0.005)	0.001 (0.005)	0.000 (0.004)	0.018 (0.009)	0.000	0.000	0.000
1950	0.005 (0.005)	-0.003 (0.004)	-0.006 (0.004)	0.001 (0.004)	0.001 (0.004)	0.023 (0.011)	0.094	0.256	0.055
1960	-0.002 (0.008)	0.004 (0.006)	-0.006 (0.008)	0.001 (0.007)	-0.005 (0.007)	0.054 (0.018)	0.330	0.221	0.230
1970	0.015 (0.011)	0.017 (0.007)	0.002 (0.006)	0.007 (0.006)	0.012 (0.007)	0.078 (0.025)	0.053	0.031	0.060
1980	0.002 (0.010)	0.003 (0.010)	-0.009 (0.012)	-0.006 (0.011)	-0.005 (0.009)	0.133 (0.031)	0.312	0.308	0.215
1990	-0.003 (0.012)	-0.014 (0.010)	-0.021 (0.011)	-0.021 (0.011)	-0.020 (0.011)	0.181 (0.039)	0.090	0.272	0.145
2000	-0.006 (0.013)	-0.024 (0.010)	-0.027 (0.012)	-0.028 (0.013)	-0.025 (0.013)	0.216 (0.042)	0.002	0.194	0.007
2010	0.016 (0.017)	-0.028 (0.011)	-0.028 (0.014)	-0.032 (0.012)	-0.026 (0.014)	0.249 (0.045)	0.002	0.074	0.017
R squared:		0.064							

Table A1: Seemingly unrelated regression of whether there is any drilling in a section by date T, for T=1930 to 2010, in 10 year increments. Sample size is 12,594 for each year. I use Conley standard errors with uniform weights and a maximum correlation distance of 20 miles. Column 7 gives the p value for a joint test that the coefficients in columns 1-5 in a given year are all equal to zero. Column 8 gives the p value for a joint test that the coefficients in columns 2-5 are all equal to zero. Column 9 gives the p value for a joint test the the coefficients in columns 1-5 are all equal to each other. The test that the difference between each coefficient in 2010 versus 1960 is equal to zero is rejected with a p value of  $2.6 \times 10^{-6}$ .

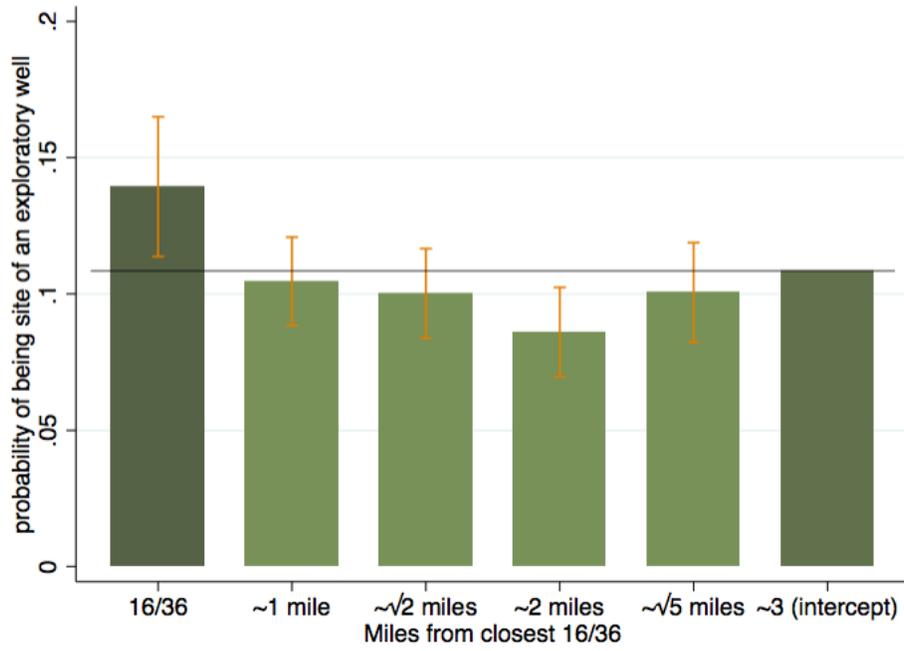


Figure A4: Graphical from of regression coefficients from column 2 of Table 6. Graph shows coefficient estimates for  $\alpha + \beta_d$  for the first 5 columns ( $\beta_{16/36}$  to  $\beta_{\sqrt{5}}$  miles) and  $\alpha$  for the last column (3 miles). The 95% confidence intervals are shown for each estimate of  $\beta_d$  relative to  $\alpha$ .

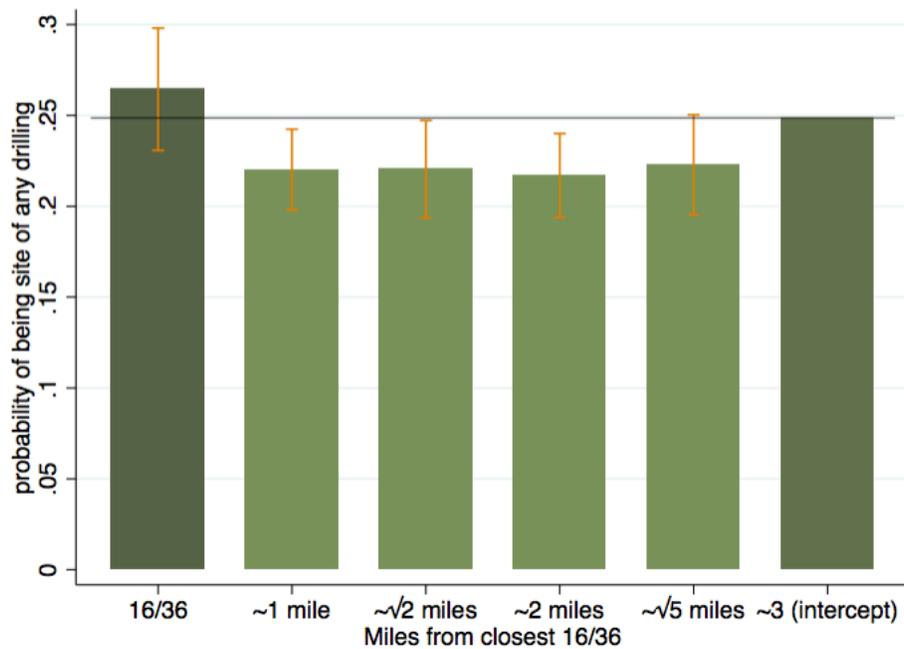


Figure A5: Graphical from of regression coefficients from column 6 of Table 6. Graph shows coefficient estimates for  $\alpha + \beta_d$  for the first 5 columns ( $\beta_{16/36}$  to  $\beta_{\sqrt{5}}$  miles) and  $\alpha$  for the last column (3 miles). The 95% confidence intervals are shown for each estimate of  $\beta_d$  relative to  $\alpha$ .

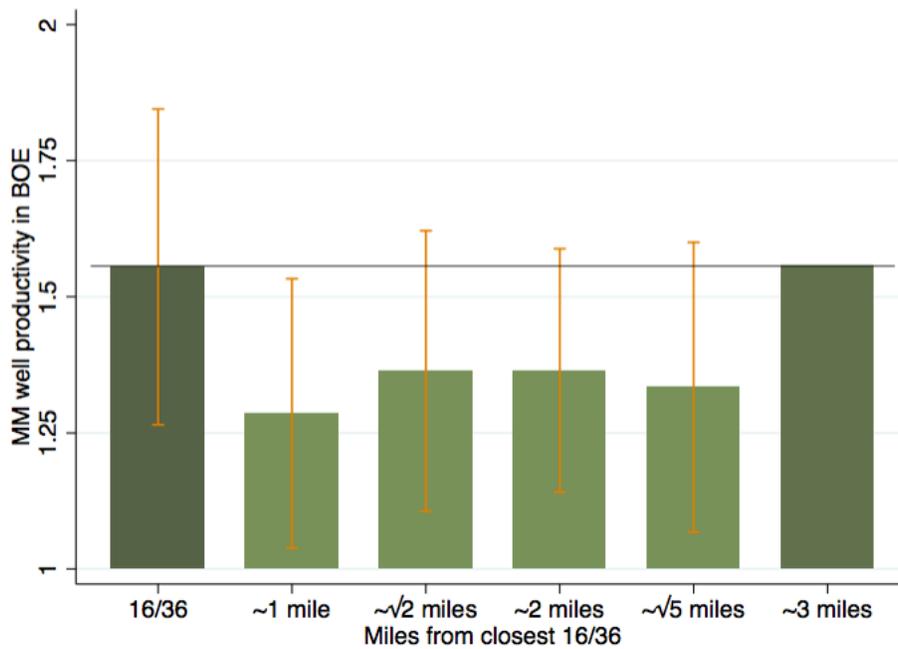


Figure A6: Graphical from of regression coefficients from column 2 of Table 7. Graph shows coefficient estimates for  $\beta_d$  for the first 5 columns ( $\beta_{16/36}$  to  $\beta_{\sqrt{5}}$  miles) relative to the average over the excluded category (3 miles). The 95% confidence intervals are shown for each estimate of  $\beta_d$  relative to the excluded category.

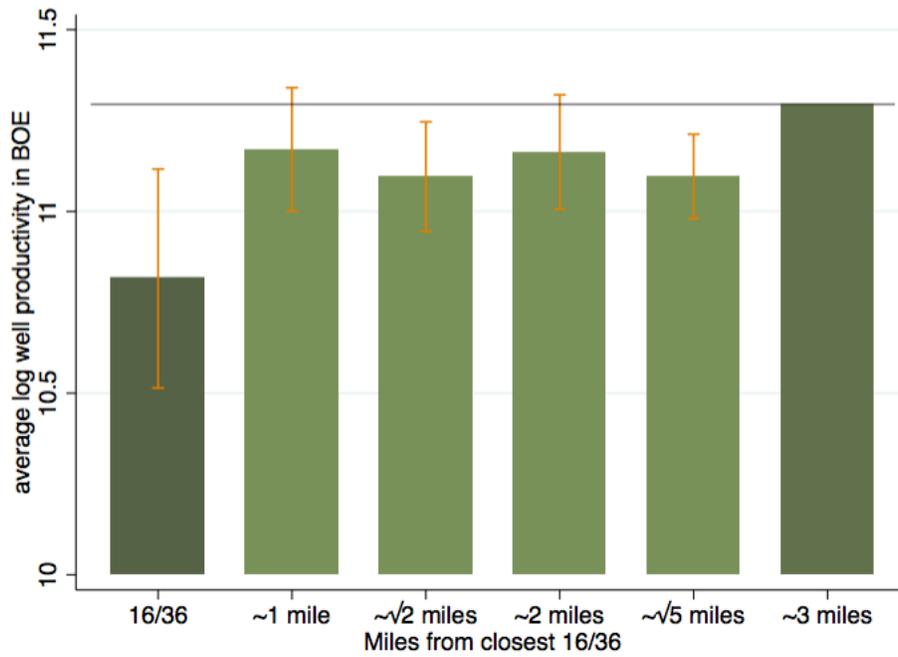


Figure A7: Graphical from of regression coefficients from column 2 of Table 8. Graph shows coefficient estimates for  $\beta_d$  for the first 5 columns ( $\beta_{16/36}$  to  $\beta_{\sqrt{5}}$ ) relative to the average over the excluded category (3 miles). The 95% confidence intervals are shown for each estimate of  $\beta_d$  relative to the average over the excluded group ( $\approx 3$  miles).

	(1)	(2)	(3)	(4)	(5)	(6)
	BOE	BOE	Gas	Gas	Oil	Oil
is 16/36	-0.48 (0.30)	-0.49 (0.13)	-0.61 (0.33)	-0.59 (0.17)	-0.21 (0.28)	-0.20 (0.06)
$\approx 1$ mile away	-0.22 (0.13)	-0.17 (0.07)	-0.23 (0.12)	-0.18 (0.04)	-0.12 (0.15)	-0.10 (0.05)
$\approx \sqrt{2}$ miles away	-0.16 (0.10)	-0.20 (0.07)	-0.20 (0.12)	-0.22 (0.04)	0.02 (0.12)	-0.14 (0.07)
$\approx 2$ miles away	-0.26 (0.15)	-0.16 (0.07)	-0.19 (0.14)	-0.10 (0.04)	-0.12 (0.17)	-0.10 (0.04)
$\approx \sqrt{5}$ miles away	-0.26 (0.08)	-0.21 (0.05)	-0.27 (0.12)	-0.18 (0.05)	-0.07 (0.12)	-0.12 (0.06)
constant	10.83 (0.52)		12.53 (0.56)		7.83 (0.53)	
township FE	No	Yes	No	Yes	No	Yes
R squared	0.00	0.51	0.00	0.54	0.00	0.49
p value joint test	0.21	0.00	0.17	0.00	0.86	0.00
p value non-16/36	0.00	0.00	0.14	0.00	0.39	0.04
Observations	7800	7800	7568	7568	7239	7239

Table A2: Regressions of log well-level production for the first 12 months of production for wells drilled between 1978 and 2010. BOE refers to barrel-of-oil equivalent production. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

	(1)	(2)	(3)	(4)	(5)	(6)
	BOE	BOE	Gas	Gas	Oil	Oil
is 16/36	-0.45 (0.30)	-0.44 (0.17)	-0.59 (0.33)	-0.55 (0.21)	-0.21 (0.27)	-0.18 (0.06)
$\approx 1$ mile away	-0.16 (0.12)	-0.10 (0.10)	-0.22 (0.11)	-0.13 (0.07)	-0.07 (0.17)	-0.06 (0.08)
$\approx \sqrt{2}$ miles away	-0.16 (0.08)	-0.18 (0.09)	-0.24 (0.11)	-0.21 (0.05)	0.02 (0.13)	-0.13 (0.08)
$\approx 2$ miles away	-0.21 (0.15)	-0.11 (0.09)	-0.17 (0.13)	-0.07 (0.07)	-0.10 (0.18)	-0.07 (0.07)
$\approx \sqrt{5}$ miles away	-0.24 (0.07)	-0.19 (0.07)	-0.35 (0.18)	-0.19 (0.06)	-0.07 (0.13)	-0.13 (0.08)
constant	11.51 (0.47)		13.24 (0.49)		8.43 (0.50)	
township FE	No	Yes	No	Yes	No	Yes
R squared	0.00	0.48	0.00	0.51	0.00	0.48
p value joint test	0.30	0.01	0.16	0.00	0.00	0.05
p value non-16/36	0.00	0.00	0.02	0.00	0.28	0.05
Observations	7800	7800	7620	7620	7341	7341

Table A3: Regressions of log well-level production for the first 36 months of production for wells drilled between 1978 and 2010. BOE refers to barrel-of-oil equivalent production. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

	(1)	(2)	(3)	(4)	(5)	(6)
	BOE	BOE	Gas	Gas	Oil	Oil
is 16/36	-0.03 (0.06)	0.00 (0.07)	-0.03 (0.06)	-0.00 (0.06)	0.00 (0.07)	0.01 (0.07)
$\approx 1$ mile away	0.03 (0.06)	0.04 (0.06)	0.03 (0.06)	0.03 (0.06)	0.02 (0.05)	0.01 (0.03)
$\approx \sqrt{2}$ miles away	-0.07 (0.06)	-0.05 (0.05)	-0.05 (0.06)	-0.04 (0.05)	-0.06 (0.07)	-0.06 (0.06)
$\approx 2$ miles away	0.02 (0.04)	0.01 (0.03)	-0.00 (0.05)	-0.01 (0.03)	0.06 (0.03)	0.04 (0.02)
$\approx \sqrt{5}$ miles away	-0.03 (0.03)	-0.03 (0.04)	-0.04 (0.03)	-0.03 (0.04)	-0.03 (0.03)	-0.05 (0.02)
constant	-0.54 (0.08)		-0.52 (0.08)		-0.64 (0.07)	
township FE	No	Yes	No	Yes	No	Yes
R squared	0.00	0.11	0.00	0.13	0.00	0.10
p value joint test	0.45	0.02	0.00	0.60	0.00	0.00
p value non-16/36	0.45	0.03	0.00	0.09	0.00	0.01
Observations	7432	7432	7236	7236	6823	6823

Table A4: Regressions of log well decline, measured as the  $\log(\text{production in year 2} - \text{production in year 1}) - \log(\text{production in year 1})$  For wells drilled between 1978 and 2010. BOE refers to barrel-of-oil equivalent production. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

	(1)	(2)	(3)	(4)	(5)	(6)
	BOE	BOE	Gas	Gas	Oil	Oil
is 16/36	-0.00 (0.03)	0.00 (0.04)	-0.01 (0.03)	-0.00 (0.04)	0.00 (0.01)	0.00 (0.05)
$\approx 1$ mile away	0.03 (0.01)	0.03 (0.05)	0.01 (.)	0.02 (0.05)	0.03 (0.01)	0.01 (0.02)
$\approx \sqrt{2}$ miles away	-0.01 (0.03)	-0.00 (0.06)	-0.02 (0.03)	-0.01 (0.05)	-0.02 (0.02)	-0.02 (0.02)
$\approx 2$ miles away	0.04 (0.02)	0.03 (0.04)	0.02 (0.03)	0.01 (0.04)	0.04 (0.03)	0.02 (0.03)
$\approx \sqrt{5}$ miles away	0.04 (0.03)	0.03 (0.03)	0.02 (0.03)	0.01 (0.04)	0.02 (0.05)	0.00 (0.04)
constant	-0.39 (0.05)		-0.37 (0.06)		-0.43 (0.03)	
township FE	No	Yes	No	Yes	No	Yes
R squared	0.00	0.15	0.00	0.16	0.00	0.07
p value joint test	0.07	0.15	0.00	0.72	0.11	0.57
p value non-16/36	0.07	0.03	0.45	0.69	0.11	0.24
Observations	6858	6858	6697	6697	6269	6269

Table A5: Regressions of log well decline, measured as the  $\log(\text{production in year 3} - \text{production in year 2}) - \log(\text{production in year 2})$  For wells drilled between 1978 and 2010. BOE refers to barrel-of-oil equivalent production. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

### A.6.5 Number of Operating Firms

Table [A6](#) examines whether there is evidence of companies being more heterogeneous in regions around 16/36 sections. It examines two types of section neighborhoods. One neighborhood is the group of sections that are the 16/36 section and the adjacent sections (e.g., 1 mile away). The other neighborhood is the group of sections that are the 13/33 section and the adjacent sections (e.g., 1 mile away). For this regression, I use firm names as of 2010. I compute for each group the total number of firms, the number of firms per well, and whether there is more than one firm listed. I find that while 16/36 neighborhoods have slightly more firms, more firms per well, and are more likely to have more than one firm, the differences do not appear to be either statistically nor economically important, with percentage differences less than 5% for each of these three measures. Therefore this regression does not show that ownership heterogeneity is likely to be significantly different. A significant drawback with this approach is that the data only include information on the name of the firm in 2010, not the name of the firm at the date the well was drilled. Because ownership of a well may change over time, these metrics probably have some measurement error.

	(1)	(2)	(3)
	Total firms	Firms per well	More than one
16/36 neighborhood	0.118 (0.105)	0.013 (0.020)	0.016 (0.036)
constant	2.446 (0.286)	0.714 (0.039)	0.565 (0.049)
R squared	0.001	0.000	0.000
Observations	685	685	685

Table A6: Regressions of company name information on location. The unit of observation is a collection of sections. There are two types of observations. The first type is made up of a 16/36 section and all sections within 1 mile of that section. The second type is made up of a 13/33 section and all sections within 1 mile of the section. There are three dependent variables: First, the total number of companies that appear in 2010. Second, the average number of companies per well. Third, whether there is more than one company. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles, basing the distance off of the location of either the 16/36 or the 13/33 sections.

### A.6.6 Examination of Roads

In this sub-subsection, I discuss evidence on roads. Oil and gas activity and roads are likely correlated but it is difficult to empirically determine why. On the one hand, presence of roads decreases the costs of accessing a drilling site. On the other hand, roads are often built as part of drilling operations. As the data on roads is only ex-post (roads as of 2009), it is difficult to determine which way the causality goes.

In this sub-subsection I focus on ruling out one story that is a threat to identification: If the instruments (16/36 sections and proximity to 16/36 sections) are correlated with road building which in turn affected drilling patterns, then the drilling results we see would be the result of roads rather than instruments.

In this sub-subsection, I begin by describing the data on roads, including the different categories of roads. Next I describe correlation of the instruments with road presence. I find that roads are more likely on the sections  $\approx 3$  miles from 16/36 sections, and that this pattern seems to be driven by county roads. Then, to rule out the story that roads are driving drilling patterns, I show that the drilling regressions are robust to controlling for roads. I also show that in a matched pair analysis where I compare sections  $\approx 3$  miles from 16/36 sections with those that are  $\approx \sqrt{5}$  miles from state land, that higher presence of county roads on one section relative to the other is not correlated with higher likelihood of drilling on one section relative to the other.

**Descriptive evidence on roads:** In examining roads, it is useful to account for different categories of roads, as the reasons for building a major highway are different from the reasons for building a four-wheel drive road. My road data uses road categorizations from the census:<sup>31</sup> Category A2 of roads refers to non-divided highways which are typically state highways. A3 refers to secondary and connecting roads, and in Wyoming are often local county roads. A4 refers to local, neighborhood, and rural roads. Categories A0, A5, A6, and A7 refer respectively to unclassified roads, vehicular trails (e.g., four-wheel drive), roads with special characteristics, and roads not used in public vehicular travel (like bike paths and private roads). In my analysis, I lump A0, A5, A6, and A7 together. My sample does not include any A1 roads (divided highways like interstate highways).

In the first and second panel of Figure [A8](#), I map the section centroids that have A2 and A3 roads. The figure shows that A2 (state highways) and A3 roads (county roads) are relatively sparse. In contrast the top panel of Figure [A9](#) shows that A4 roads are relatively widespread. To a lesser extent,

---

<sup>31</sup>This categorization is known as the Census Feature Class Codes (CFCC).

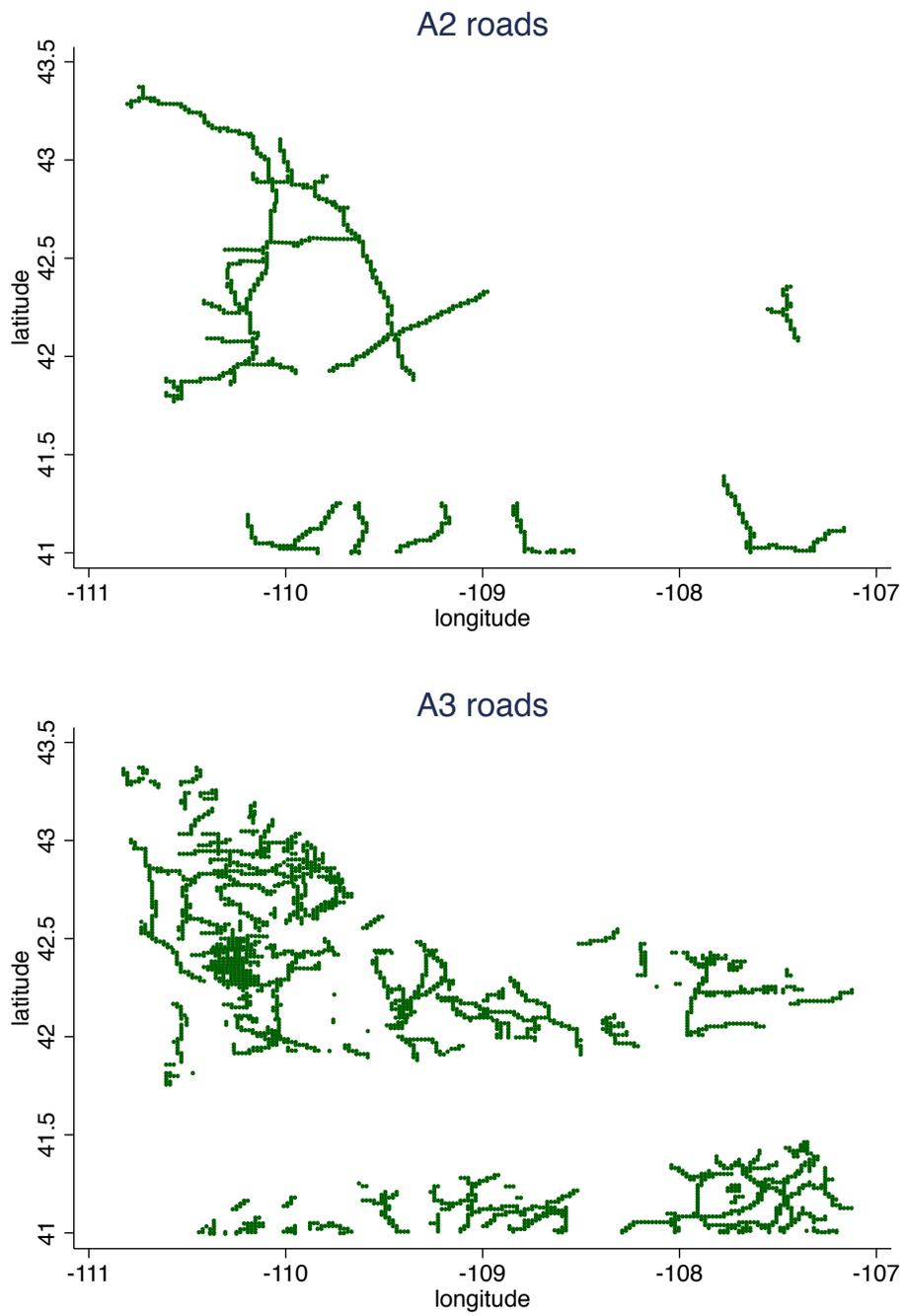


Figure A8: Maps of section centroids with A2 and A3 road presence, respectively.

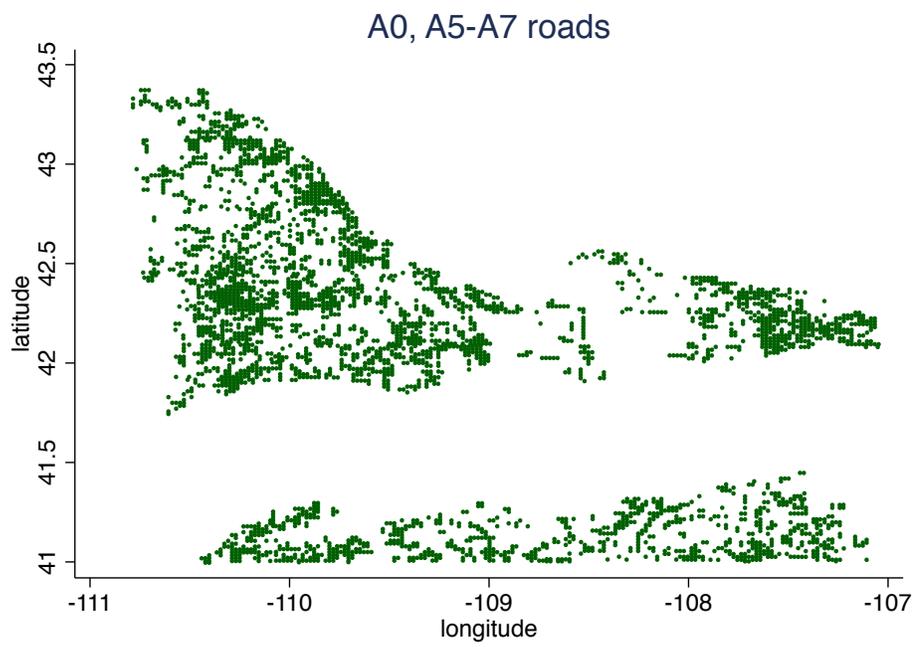
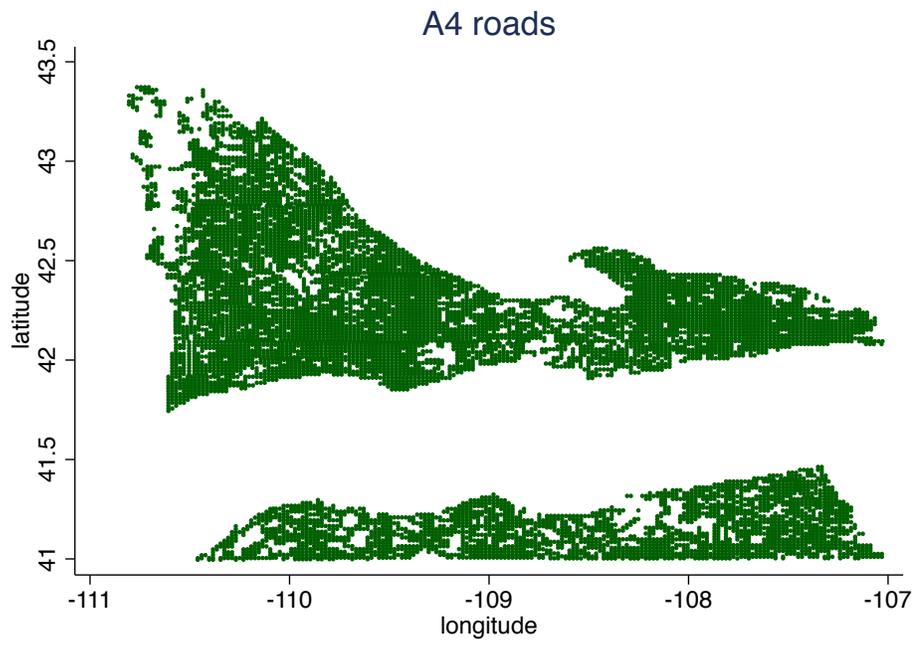


Figure A9: Maps of section centroids with A4 roads (top panel) and A0, A5, A6, and/or A7 roads (bottom panel).

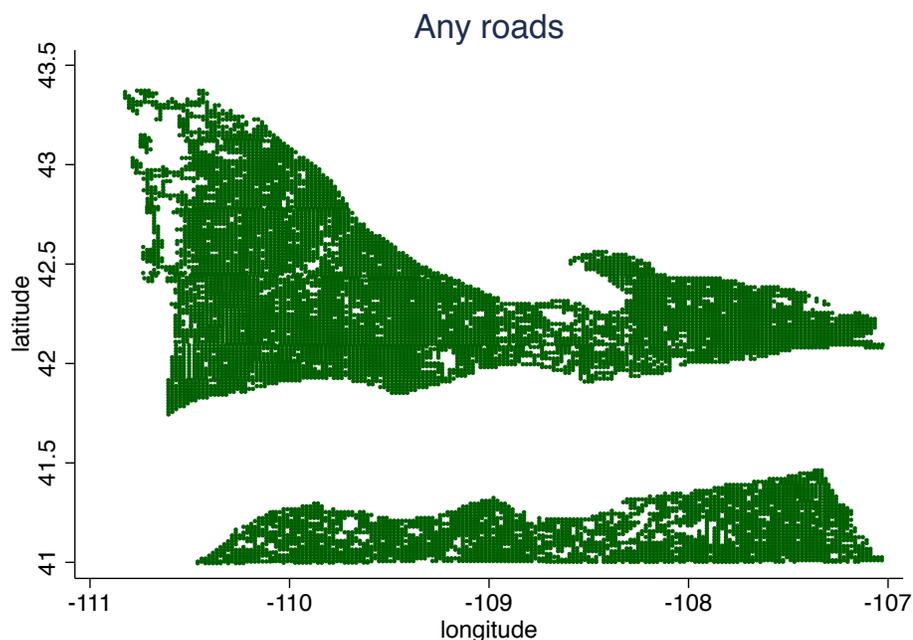


Figure A10: Maps of section centroids with any road presence.

so are A0, A5, A6, and A7 roads (lower panel of Figure [A9](#)). Overall, the vast majority (over 80%) of sections in the sample have at least some road, as mapped in Figure [A10](#).

**Roads and the Instrument:** Next I examine how road presence as of 2010 is correlated with the instruments—16/36 sections and proximity to 16/36 sections. For each of these categories of roads, I run regression [3](#) where the dependent variable is an indicator for the presence of that road type. Regression results are in Table [A7](#).

I find that overall road presence is correlated with the instruments (column 5 of Table [A7](#)). In particular, I find that sections 3 miles from state land (the excluded group) are the most likely type of sections to have a road. When I examine each type of road individually, I find that the differences seem to be driven by A3 roads (county roads): Sections three miles from state land are more likely to have A3 roads than any other type of section. In contrast, for the other categories of roads (A2, A4, and A0/A5-A7 roads), I do not find evidence that road presence is correlated with the instrument.

Similar results hold when I examine road distance. In Table [A8](#), I do comparable regressions where instead of using an indicator variable of road

presence as the dependent variable, I instead use the MacKinnon-Magee transform of distance, where the dependent variable is  $\log(\text{meters} + \sqrt{\text{meters}^2 + 1})$ . There I find similar results. Overall, sections 3 miles from state land has longer length of total road than other types of sections (column 5). Again, I find that the effect seems to be driven by the A3 roads (column 2).

**Regression Results With Road Controls:** I find that the various drilling and production regressions still show the spillover patterns even after controlling for roads. Table [A9](#) is a table of wildcat regressions. The first column does not control for roads and is identical to column 3 of Table [6](#). The second through sixth columns control for various indicators of road types. I find that while the coefficients on roads are statistically significant, the coefficients on the section location instruments (e.g., 16/36,  $\approx 1$  mile away, etc.) remain almost completely unchanged. Importantly, this holds for column 3 where I control for A3 roads (county roads) and column 6 where I control for all roads.

Similar results hold in Table [A10](#) where I examine overall drilling probabilities. Column 1 is identical to column 7 Table [6](#). The remaining columns control for different types of road presence. Again, I find that regression estimates are very similar when controlling for various roads, including A3 roads (column 3) and all roads (column 6).

**Local spatial correlation of drilling and A3 roads:** Finally, I show evidence that there does not seem to be any local spatial correlation between drilling and A3 roads. While I do find that sections 3 miles from state land are both more likely to have drilling and to have A3 roads, a higher likelihood of having an A3 road on a section  $\approx 3$  miles away from state land relative to a neighboring section  $\approx \sqrt{5}$  miles away is not correlated with higher likelihood of drilling on a section  $\approx 3$  miles away from state land relative to a neighboring section  $\approx \sqrt{5}$  miles away.

To do this analysis, I construct a data set where an observation is a pair of sections, where one section is  $\approx 3$  miles away from the 16/36 section and where the other is a neighboring section that is  $\approx \sqrt{5}$  miles from the closest 16/36 section.<sup>32</sup> For this pair of sections, I construct an indicator variables for whether the  $\approx 3$  mile section has drilling, whether the  $\approx \sqrt{5}$  mile section has drilling, whether the  $\approx 3$  mile section has an A3 road, and whether the  $\approx \sqrt{5}$  mile section has an A3 road.

To examine the correlation in spatial patterns of drilling and A3 roads,

---

<sup>32</sup>More precisely, the data set consists of all 13/33 sections and their adjoining sections. Because of the regular numbering pattern, 13/33 sections are about 3 miles from 16/36 sections and their neighboring sections are about  $\sqrt{5}$  miles from 16/36 sections.

	(1)	(2)	(3)	(4)	(5)
	A2	A3	A4	A0,A5-A7	any roads
is 16/36	-0.009 (0.015)	-0.058*** (0.016)	-0.022 (0.025)	-0.008 (0.021)	-0.023 (0.022)
$\approx 1$ mile away	-0.015 (0.011)	-0.054*** (0.016)	-0.021 (0.022)	-0.006 (0.013)	-0.026 (0.017)
$\approx \sqrt{2}$ miles away	-0.012 (0.010)	-0.050*** (0.011)	-0.023 (0.020)	-0.009 (0.015)	-0.030** (0.014)
$\approx 2$ miles away	-0.011 (0.008)	-0.046*** (0.016)	-0.025 (0.017)	-0.011 (0.013)	-0.034*** (0.011)
$\approx \sqrt{5}$ miles away	-0.009 (0.008)	-0.046*** (0.010)	-0.011 (0.022)	-0.004 (0.015)	-0.028* (0.015)
constant	0.077*** (0.017)	0.243*** (0.028)	0.765*** (0.045)	0.287*** (0.025)	0.867*** (0.036)
R squared	0.000	0.001	0.000	0.000	0.000
p value joint significance	0.519	0.000	0.181	0.935	0.001
p value non-16/36	0.669	0.000	0.110	0.876	0.000
Observations	12549	12549	12549	12549	12549

Table A7: Regressions of whether there is a road of a given type on indicators for various section types. Category A2 refers to non-divided highways which are typically state highways. A3 refers to secondary and connecting roads, and in Wyoming are often local county roads. A4 refers to local, neighborhood, and rural roads. A0, A5, A6, and A7 refer respectively to unclassified roads, vehicular trails (e.g., four-wheel drive), roads with special characteristics, and roads not used in public vehicular travel (like bike paths and private roads). The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

	(1)	(2)	(3)	(4)	(5)
	A2	A3	A4	A0,A5-A7	any roads
is 16/36	-0.063 (0.118)	-0.409*** (0.116)	-0.178 (0.201)	-0.052 (0.141)	-0.227 (0.195)
$\approx 1$ mile away	-0.116 (0.082)	-0.389*** (0.108)	-0.179 (0.170)	-0.039 (0.081)	-0.253* (0.141)
$\approx \sqrt{2}$ miles away	-0.091 (0.077)	-0.371*** (0.080)	-0.174 (0.166)	-0.037 (0.094)	-0.259** (0.120)
$\approx 2$ miles away	-0.075 (0.065)	-0.325*** (0.109)	-0.221 (0.140)	-0.060 (0.093)	-0.314*** (0.101)
$\approx \sqrt{5}$ miles away	-0.065 (0.065)	-0.336*** (0.066)	-0.092 (0.173)	0.006 (0.103)	-0.235* (0.128)
constant	0.585*** (0.128)	1.843*** (0.204)	6.223*** (0.387)	2.013*** (0.174)	7.277*** (0.327)
R squared	0.000	0.001	0.000	0.000	0.000
p value joint significance	0.052	0.000	0.572	0.889	0.006
p value non-16/36	0.548	0.000	0.437	0.796	0.003
Observations	12549	12549	12549	12549	12549

Table A8: Regressions of the McKinnon-Magee transform of road mileage on indicators for various section types. Dependent variable is  $\log(\text{meters} + \sqrt{\text{meters}^2 + 1})$ . Category A2 refers to non-divided highways which are typically state highways. A3 refers to secondary and connecting roads, and in Wyoming are often local county roads. A4 refers to local, neighborhood, and rural roads. A0, A5, A6, and A7 refer respectively to unclassified roads, vehicular trails (e.g., four-wheel drive), roads with special characteristics, and roads not used in public vehicular travel (like bike paths and private roads). The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

Dependent variable: Any wildcat by 2010						
	(1)	(2)	(3)	(4)	(5)	(6)
is 16/36	0.029 (0.013)	0.029 (0.013)	0.030 (0.013)	0.030 (0.013)	0.029 (0.013)	0.030 (0.014)
$\approx 1$ mile away	-0.004 (0.008)	-0.004 (0.008)	-0.003 (0.008)	-0.003 (0.008)	-0.004 (0.008)	-0.003 (0.008)
$\approx \sqrt{2}$ miles away	-0.008 (0.007)	-0.008 (0.007)	-0.007 (0.007)	-0.007 (0.007)	-0.008 (0.007)	-0.007 (0.007)
$\approx 2$ miles away	-0.021 (0.007)	-0.021 (0.007)	-0.021 (0.007)	-0.021 (0.007)	-0.021 (0.007)	-0.020 (0.007)
$\approx \sqrt{5}$ miles away	-0.010 (0.009)	-0.009 (0.009)	-0.009 (0.008)	-0.009 (0.008)	-0.009 (0.009)	-0.008 (0.009)
A2		0.032 (0.015)				
A3			0.020 (0.006)			
A4				0.034 (0.006)		
A0,A5-A7					0.024 (0.008)	
Any road						0.047 (0.008)
physical chars.	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.108	0.108	0.108	0.110	0.109	0.110
p value joint test	0.007	0.007	0.009	0.008	0.006	0.010
p value non-16/36	0.007	0.006	0.009	0.007	0.005	0.011
Observations	12549	12549	12549	12549	12549	12549

Table A9: Regressions of the probability of a wildcat (exploratory) well drilled on a section by 2010. The variables A2, A3, A4, “A0,A5-A7”, and “Any road” are indicator variables for various types of roads. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. Other physical characteristics that are controlled for are township fixed effects, mean elevation, elevation range, whether elevation data was missing, presence of wetlands, and presence of irrigated crops. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

Dependent variable: Any drilling by 2010						
	(1)	(2)	(3)	(4)	(5)	(6)
is 16/36	0.015 (0.017)	0.015 (0.017)	0.017 (0.017)	0.016 (0.018)	0.015 (0.017)	0.016 (0.017)
$\approx 1$ mile away	-0.031 (0.009)	-0.031 (0.009)	-0.029 (0.009)	-0.030 (0.009)	-0.031 (0.009)	-0.030 (0.009)
$\approx \sqrt{2}$ miles away	-0.028 (0.010)	-0.028 (0.010)	-0.026 (0.009)	-0.027 (0.010)	-0.028 (0.010)	-0.026 (0.010)
$\approx 2$ miles away	-0.031 (0.008)	-0.030 (0.008)	-0.029 (0.008)	-0.030 (0.008)	-0.030 (0.008)	-0.029 (0.008)
$\approx \sqrt{5}$ miles away	-0.026 (0.009)	-0.026 (0.009)	-0.025 (0.009)	-0.025 (0.009)	-0.026 (0.009)	-0.024 (0.009)
A2		0.028 (0.019)				
A3			0.039 (0.015)			
A4				0.052 (0.014)		
A0,A5-A7					0.031 (0.013)	
Any road						0.065 (0.010)
physical chars.	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.351	0.352	0.352	0.353	0.352	0.354
p value joint test	0.000	0.000	0.000	0.000	0.000	0.000
p value non-16/36	0.000	0.000	0.001	0.001	0.000	0.000
Observations	12549	12549	12549	12549	12549	12549

Table A10: Regressions of the probability of any drilling on a section by 2010. The variables A2, A3, A4, “A0,A5-A7”, and “Any road” are indicator variables for various types of roads. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. Other physical characteristics that are controlled for are township fixed effects, mean elevation, elevation range, whether elevation data was missing, presence of wetlands, and presence of irrigated crops. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

I examine the correlation between two differences:

$$(A58) \quad \text{correlation}[1(\text{drilling in } \approx 3 \text{ mile section}) - 1(\text{drilling in } \approx \sqrt{5} \text{ mile section}), \\ 1(\text{A3 road in } \approx 3 \text{ mile section}) - 1(\text{A3 road in } \approx \sqrt{5} \text{ mile section})]$$

I also examine this correlation using a regression framework:

$$(A59) \quad 1(\text{drilling in } \approx 3 \text{ mile section}) - 1(\text{drilling in } \approx \sqrt{5} \text{ mile section}) = \\ \alpha + \beta \cdot [1(\text{A3 road in } \approx 3 \text{ mile section}) - 1(\text{A3 road in } \approx \sqrt{5} \text{ mile section})] + \varepsilon$$

Both the correlation framework and the regression framework are a way to test whether higher A3 road presence on the  $\approx 3$  mile section relative to the  $\approx \sqrt{5}$  mile section is correlated with higher drilling presence on the  $\approx 3$  mile section relative to the  $\approx \sqrt{5}$  mile section. If the two are positive correlated, then that suggests causality in one direction or the other, or perhaps a third factor that affects both roads and drilling. However if the two are not correlated, then this suggests that there is no causality in either direction.

I find almost no correlation. The correlation between the two variables (the correlation given in Equation [A58](#) is positive but very tiny—less than 1%. In Table [A11](#) I display the results of the regression specification where I allow for spatial correlation of the error term. I find that the coefficient  $\beta$  is tiny and statistically insignificant.

This implies that the higher probability of A3 roads (county roads) on sections 3 miles from state land is not driving the higher likelihood of drilling on sections 3 miles from state land. Therefore this implies that higher drilling probabilities on sections 3 miles from state land is driven by spillovers of state land, and not by spatial patterns of roads.

### A.6.7 16/36 Sections That Are Not In State Ownership

In the main paper, I find that a 16/36 section is only predicted to have 80% of it owned by the state government. This suggests that there are potentially a significant number of 16/36 sections with no state ownership. This also suggests that in areas where few 16/36 sections were assigned to state ownership, that the instrument is not strong. Therefore in this section I explore this feature of the data. I show that most of the 16/36 sections with no state ownership are in a small fraction of the GGRB located in the northwest where there is very little oil and gas production. Next, I show that drilling and production

(1)	
difference in A3 roads	0.005 (0.031)
constant	0.033 (0.011)
R squared	0.000
Observations	2820

Table A11: Regression estimate of equation [A59](#). The dependent variable is the difference in drilling outcomes by 2010. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

results are similar if we exclude this northwest area.

**Where are the 16/36 sections that are not in state ownership?**

In Figure [A11](#), I graph the GPS coordinates of the centroids of all 16/36 sections in my sample, and circle the ones with no state mineral ownership. The graph shows that most of these are in the northwest part of the sample.

To compare this northwest area with the rest of the GGRB, I categorize a section as being part of this northwest GGRB if either (the latitude is greater than 42.3 and the longitude is less than -110.45) or (latitude is greater than longitude + 110.45 + 42.9). The graph of which sections fall into this area is shown in Figure [A12](#).

This northwest part of the GGRB is fairly mountainous and includes two mountain ranges—the Salt River Range and the Wyoming Range. Due to its mountainous nature, there is very little oil and gas in this area: Only 5% of the sections in this region have had drilling by 2010 and only 0.5% report production. In contrast, for the rest of the GGRB, 25% of sections have had drilling by 2010 and 11% report production.

**Excluding the northwest of the GGRB from the analysis:** If this northwest portion of the GGRB region is excluded from the analysis, then regression results are very similar. Table [A12](#) shows wildcat and ever-drill regression results using the same specification as Table [6](#), but excluding the northwest portion. Because the northwest GGRB has very little drilling at all, excluding the northwest of the GGRB gives slightly higher expected

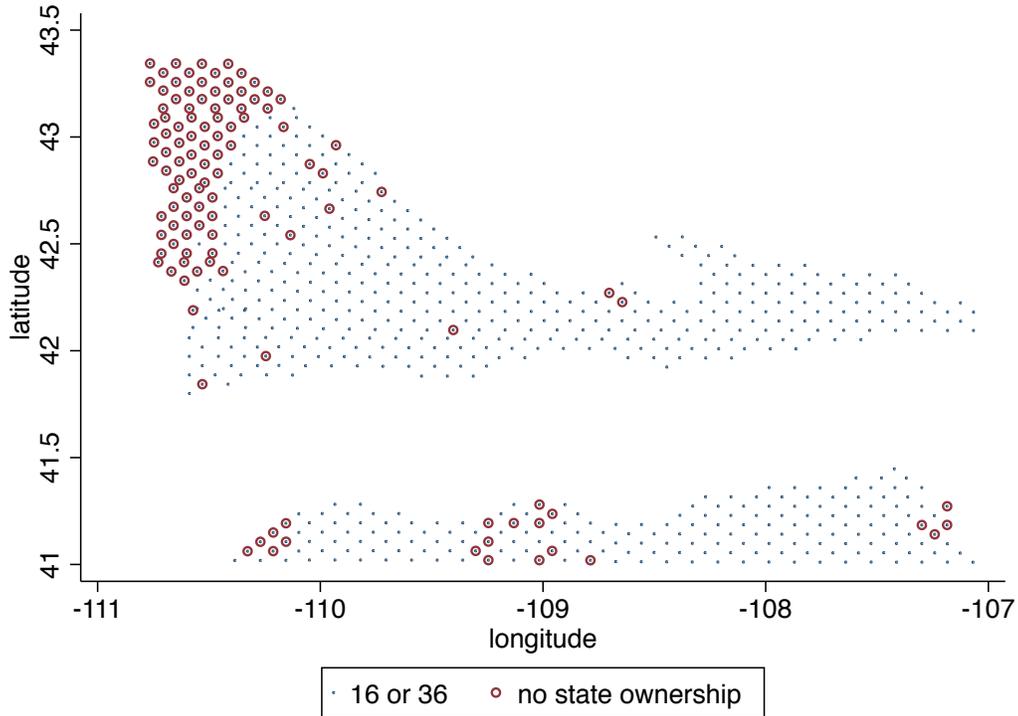


Figure A11: A map of the centroids of all 16/36 sections. Circles around the centroids indicate that the state owns none of the mineral rights within the section. The middle band is excluded from the sample because it is the area close to the transcontinental railroad, where the odd-numbered sections were transferred to the Union Pacific Railroad.

wildcat and ever-drill probabilities for all section types. For example, in the full sample column 1, expected probability of a wildcat well on a non-16/36 well is 9.8% in the full sample but 10.6% in the sample that excludes the northwest GGRB. However the overall pattern of coefficients remains the same: Drilling is most likely on 16/36 sections and sections 3 miles away. P value tests of joint significance also are similar.

For regressions of section-level production, Table [A13](#) uses the same specification as Table [7](#) but excludes the northwest part of the GGRB. Both point estimates as well as p values are similar. The one exception is the “oil” specification with no township fixed effects, where I find that the p values for each of the joint tests are statistically significant.

For regressions of the first 24 months of production, Table [A14](#) uses

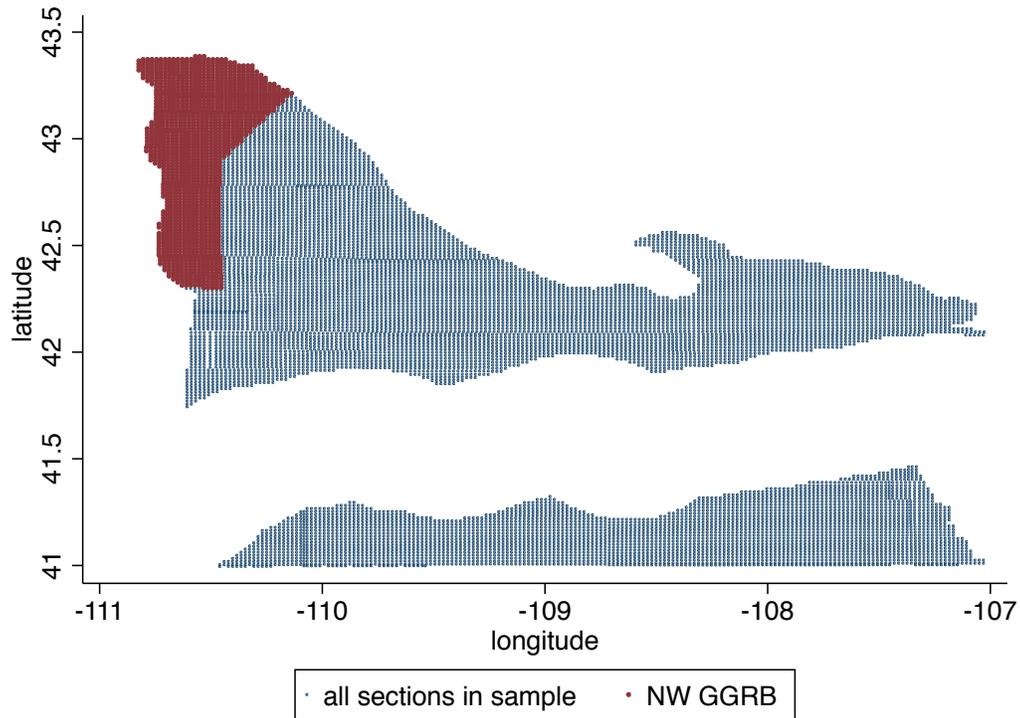


Figure A12: A map of the centroids of all sections in the sample. The centroids that are in red are those which I categorize as being in the northwest part of the GGRB. I exclude these sections in the robustness check regressions contained in Tables [A12](#), [A13](#), and [A14](#). The middle band is excluded from the sample because it is the area close to the transcontinental railroad, where the odd-numbered sections were transferred to the Union Pacific Railroad.

the same specification as [8](#) in the revised version of the paper except that it excludes the northwest part of the GGRB. For these, the coefficients and the p values are similar.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	wildcat	wildcat	wildcat	wildcat	drill	drill	drill	drill
is 16/36	0.045 (0.015)	0.035 (0.013)	0.033 (0.013)	0.028 (0.021)	0.047 (0.015)	0.019 (0.019)	0.017 (0.019)	0.016 (0.022)
$\approx 1$ mile away		-0.003 (0.009)	-0.003 (0.008)	-0.006 (0.008)		-0.029 (0.012)	-0.032 (0.011)	-0.024 (0.008)
$\approx \sqrt{2}$ miles away		-0.007 (0.008)	-0.006 (0.007)	-0.006 (0.010)		-0.027 (0.015)	-0.027 (0.011)	-0.023 (0.016)
$\approx 2$ miles away		-0.023 (0.009)	-0.022 (0.008)	-0.019 (0.013)		-0.030 (0.013)	-0.031 (0.009)	-0.012 (0.008)
$\approx \sqrt{5}$ miles away		-0.010 (0.011)	-0.010 (0.010)	-0.010 (0.015)		-0.031 (0.015)	-0.029 (0.010)	-0.026 (0.019)
constant	0.106 (0.012)	0.116 (0.013)		0.117 (0.012)	0.244 (0.042)	0.272 (0.049)		0.284 (0.045)
physical chars.	No	No	Yes	No	No	No	Yes	No
R squared	0.001	0.002	0.107	0.001	0.001	0.001	0.346	0.001
p value joint test		0.006	0.014	0.122		0.002	0.000	0.009
p value non-16/36		0.045	0.011	0.472		0.073	0.003	0.004
Observations	11085	11085	11085	6863	11085	11085	11085	6863

Table A12: Excludes Northwest part of GGRB. Regressions of the probability of wildcat (exploratory) well on a section by 2010 (columns 1-4) or that a section ever had a wildcat (exploratory) well by 2010 (columns 5-8). The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. Physical characteristics that are controlled for in columns 3 and 6 are township fixed effects, mean elevation, elevation range, whether elevation data was missing, presence of wetlands, and presence of irrigated crops. Columns 5 and 8 are limited to the sample where the land ownership pattern has remained particularly strong. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

	(1)	(2)	(3)	(4)	(5)	(6)
	BOE	BOE	Gas	Gas	Oil	Oil
is 16/36	-0.01 (0.17)	-0.04 (0.18)	0.04 (0.19)	0.01 (0.20)	0.00 (0.15)	-0.01 (0.14)
$\approx 1$ mile away	-0.27 (0.15)	-0.31 (0.14)	-0.28 (0.18)	-0.32 (0.17)	-0.21 (0.12)	-0.23 (0.12)
$\approx \sqrt{2}$ miles away	-0.21 (0.16)	-0.23 (0.15)	-0.22 (0.18)	-0.24 (0.18)	-0.15 (0.12)	-0.17 (0.12)
$\approx 2$ miles away	-0.18 (0.12)	-0.22 (0.13)	-0.17 (0.14)	-0.22 (0.15)	-0.12 (0.11)	-0.14 (0.11)
$\approx \sqrt{5}$ miles away	-0.27 (0.16)	-0.27 (0.16)	-0.28 (0.19)	-0.29 (0.18)	-0.17 (0.14)	-0.17 (0.13)
constant	1.75 (0.48)		1.92 (0.55)		1.22 (0.36)	
township FE	No	Yes	No	Yes	No	Yes
R squared	0.00	0.47	0.00	0.47	0.00	0.46
p value joint test	0.13	0.31	0.20	0.40	0.00	0.20
p value non-16/36	0.11	0.20	0.18	0.33	0.00	0.17
Observations	11085	11085	11085	11085	11085	11085

Table A13: Excludes Northwest part of GGRB. Regression table where dependent variable is MacKinnon-Magee transform of total 1978-2010 production for wells drilled between 1978 and 2010, aggregated to the section level. BOE refers to barrel-of-oil equivalent production. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

	(1)	(2)	(3)	(4)	(5)	(6)
	BOE	BOE	Gas	Gas	Oil	Oil
is 16/36	-0.48 (0.30)	-0.48 (0.15)	-0.61 (0.33)	-0.59 (0.19)	-0.23 (0.27)	-0.21 (0.03)
$\approx 1$ mile away	-0.18 (0.11)	-0.12 (0.09)	-0.21 (0.10)	-0.15 (0.05)	-0.09 (0.15)	-0.08 (0.07)
$\approx \sqrt{2}$ miles away	-0.17 (0.08)	-0.20 (0.08)	-0.22 (0.10)	-0.22 (0.04)	0.01 (0.13)	-0.15 (0.07)
$\approx 2$ miles away	-0.23 (0.15)	-0.13 (0.08)	-0.19 (0.14)	-0.09 (0.05)	-0.11 (0.18)	-0.08 (0.06)
$\approx \sqrt{5}$ miles away	-0.25 (0.07)	-0.20 (0.06)	-0.28 (0.11)	-0.18 (0.05)	-0.08 (0.12)	-0.13 (0.07)
constant	11.29 (0.49)		13.01 (0.52)		8.24 (0.52)	
township FE	No	Yes	No	Yes	No	Yes
R squared	0.00	0.49	0.00	0.52	0.00	0.49
p value joint test	0.25	0.00	0.10	0.00	0.00	0.15
p value non-16/36	0.00	0.00	0.00	0.00	0.34	0.15
Observations	7788	7788	7576	7576	7312	7312

Table A14: Excludes Northwest part of GGRB. Regressions of log well-level production for the first 24 months of production for wells drilled between 1978 and 2010. BOE refers to barrel-of-oil equivalent production. The first p value is the test that the coefficients for 16/36 and 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away are all equal to zero. The second p value is the test that the coefficients for 1,  $\sqrt{2}$ , 2, and  $\sqrt{5}$  miles away all equal to zero. I use Conley standard errors with a uniform weight and maximum correlation distance of 20 miles.

## A.7 Geological Evidence on Common Pools

Finally, I present some additional evidence on common pools in the GGRB. The GGRB is a relatively impermeable basin relative to spacing rules. Most oil and gas fields in the GGRB are “tight” formations, consisting of relatively impermeable rock, meaning that oil and gas cannot flow long distances underground (Nelson, Ewald, Santus, and Trainor, 2009; DeJarnett, Lim, and Calogero, 2002). In figure A13 I graph permeability from a number of wells using figures taken from the 1979 and 1992 editions of the Wyoming Oil and Gas Fields Symposium (Wyoming Geological Association, 1979, 1992). The distribution of permeability in Figure A13 shows that permeability for most wells ranges from 0.1 millidarcys to 100 millidarcys. This is in the low range of permeability for conventional oil and gas fields (see for example Bear (1972), page 136). This range of permeability leads to relatively small drainage areas relative to default spacing units. For example, the Hogsback and the Bunker Hill natural gas fields have estimated permeability at 94 and 26.5 millidarcys and have estimated drainage areas of 37.3 and 39.3 acres, respectively.<sup>33</sup> This is significantly smaller than the default natural gas spacing unit size of 640 acres. As these permeability measures are on the high end of the permeability distribution graphed in Figure A13, and as most fields in the GGRB are predominately natural gas, this suggests that common pools are unlikely to occur in the region.

---

<sup>33</sup>This evidence is taken from geological evidence submitted by firms to request smaller spacing units available on [wogcc.state.wy.us](http://wogcc.state.wy.us).

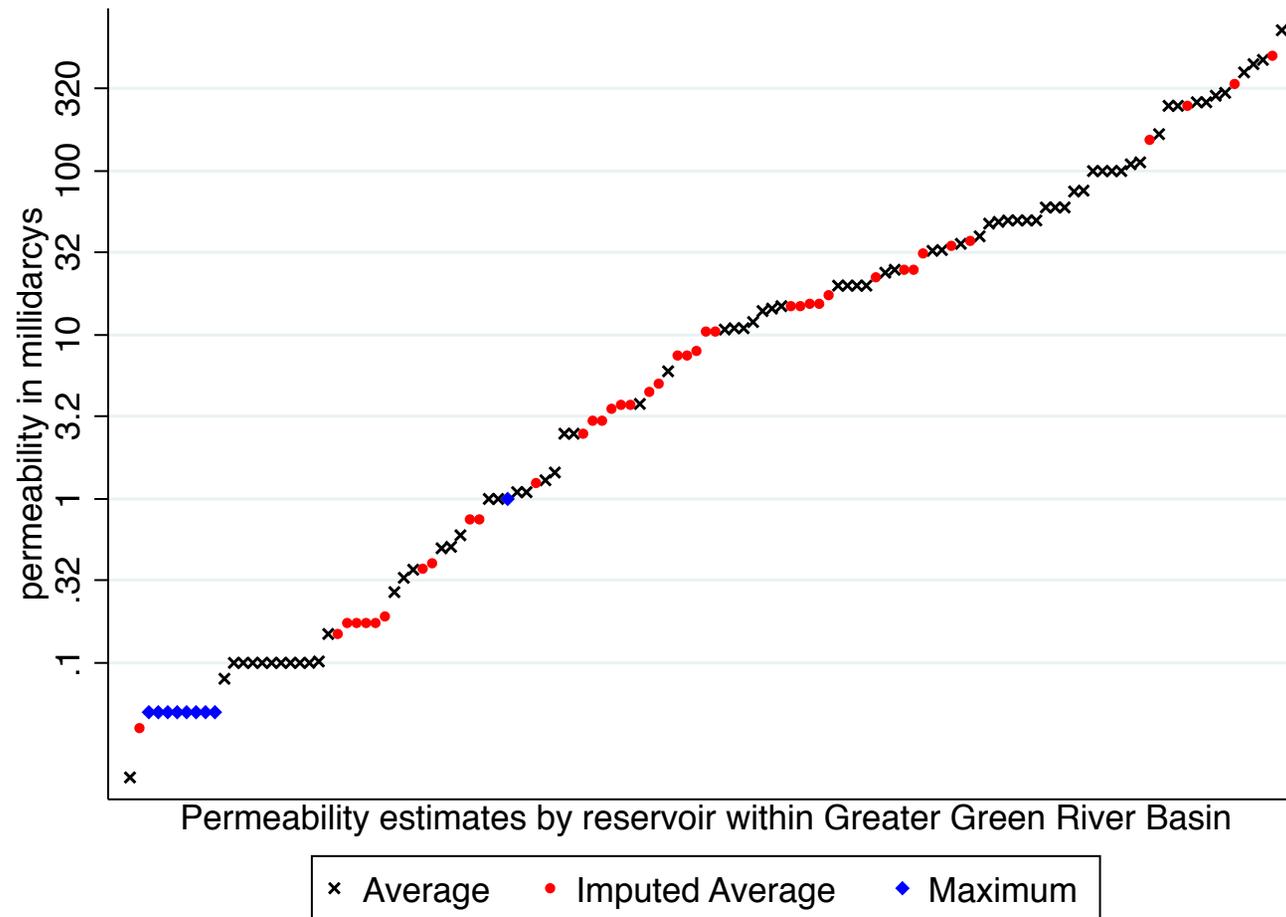


Figure A13: Geological estimates of permeability in reservoirs within fields in the Greater Green River Basin. Excludes reservoirs that lie predominately under the transcontinental railroad checkerboard region. This plots the average permeability if reported. If a maximum and minimum permeability is reported but not an average, this charts the mean of the maximum and minimum. In some cases only a maximum permeability is reported

## References

- ARPS, J. J. (1945): “Analysis of Decline Curves,” *Transactions of the American Institute of Mining Engineers*, 160, 228–247.
- BALTHROP, A. AND K. SCHNIER (2012): “A Regression Discontinuity Approach to Oil and Natural Gas Regulation,” Working paper.
- BEAR, J. (1972): *Dynamics of Fluids in Porous Media*, Dover.
- BIEWICK, L. R. (2011): “Geodatabase of Wyoming statewide oil and gas drilling activity to 2010,” U.S. Geological Survey Data Series 625.
- CONLEY, T. G. (1999): “GMM Estimation with Cross Sectional Dependence,” *Journal of Econometrics*, 92, 1–45.
- (2008): “Spatial Econometrics,” in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. Blume, Basingstoke, Hampshire ; New York: Palgrave Macmillan, 2 ed.
- COVERT, T. R. (2015): “Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale,” Working Paper.
- DEJARNETT, B. B., F. H. LIM, AND D. CALOGERO (2002): *Greater Green River Basin Production Improvement Project*, U.S. Department of Energy [www.netl.doe.gov/KMD/cds/Disk28/NGP7.PDF](http://www.netl.doe.gov/KMD/cds/Disk28/NGP7.PDF).
- EDWARDS, E. C., T. O’GRADY, AND D. JENKINS (2016): “The Effect of Land Ownership on Oil and Gas Production: A Natural Experiment,” Working paper.
- ENERGY INFORMATION AGENCY (2010): *Top 100 U.S. Oil & Gas Fields By 2009 Proven Reserves*, [http://www.eia.gov/oil\\_gas/rpd/topfields.pdf](http://www.eia.gov/oil_gas/rpd/topfields.pdf).
- FITZGERALD, T. (2010): “Evaluating Split Estates in Oil and Gas Leasing,” *Land Economics*, 86, 294–312.
- FITZGERALD, T. AND A. STOCKING (2014): “The Effects of Oil and Gas Fiscal Regimes on Exploration and Production Decisions,” Working Paper.
- FLANERY, S. O. AND R. J. MORGAN (2011): *Overview of Pooling and Unitization Affecting Appalachian Shale Development*, [www.steptoeh-johnson.com/sites/default/files/null/A\\_National\\_Survey\\_of\\_Statutory\\_Pooling\\_and\\_Unitization](http://www.steptoeh-johnson.com/sites/default/files/null/A_National_Survey_of_Statutory_Pooling_and_Unitization).

- GERKING, S. AND W. E. MORGAN (2007): “Effects of Environmental and Land Use Regulation in the Oil and Gas Industry Using the Wyoming Checkerboard as a Natural Experiment: Retraction,” *The American Economic Review*, 97, 1032.
- HENDRICKS, K. AND D. KOVENOCK (1989): “Asymmetric Information, Information Externalities, and Efficiency: The Case of Oil Exploration,” *RAND Journal of Economics*, 20, 164–182.
- HENDRICKS, K., R. H. PORTER, AND G. TAN (1993): “Optimal Selling Strategies for Oil and Gas Leases with an Informed Buyer,” *The American Economic Review: Papers and Proceedings*, 83, 234–239.
- HOLMES, T. J., B. SEO, AND M. H. SHAPIRO (2014): “Leasing, Learning, and Oil Field Development,” Working Paper.
- INTERSTATE OIL AND GAS COMPACT COMMISSION (2015): *Horizontal Well Development Pooling, Spacing, and Unitization: A Regulatory Toolbox for Key Policy, Regulatory, and Statutory Considerations*, <http://iogcc.myshopify.com/products/horizontal-well-development-pooling-spacing-and-unitization>.
- KELLOGG, R. (2014): “The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling,” *American Economic Review*, 104, 1698–1734.
- KRUPNICK, A. J., M. GOTTLIEB, AND R. J. KOPP (2014): *Attaining Sustainable Development of Oil and Gas in North America: Appendix: US Policy Briefs*, Resources for the Future. <http://www.rff.org/RFF/Documents/RFF-Rpt-NAmericanEnergy-US.pdf>.
- KUNCE, M., S. GERKING, AND W. MORGAN (2002): “Effects of Environmental and Land Use Regulation in the Oil and Gas Industry Using the Wyoming Checkerboard as an Experimental Design,” *The American Economic Review*, 92, 1588–1593.
- LARSON, T. A. (1990): *History of Wyoming*, University of Nebraska Press.
- LEVITT, C. J. (2011): “Learning through Oil and Gas Exploration,” Working Paper.
- LIBECAP, G. D. AND S. N. WIGGINS (1984): “Contractual Responses to the Common Pool: Prorating of Crude Oil Production,” *The American Economic Review*, 74, 87–98.
- (1985): “The Influence of Private Contractual Failure on Regulation: The Case of Oil Field Unitization,” *The Journal of Political Economy*, 93, 690–714.

- LIN, C.-Y. C. (2013): “Strategic Decision-Making with Information and Extraction Externalities: A Structural Model of the Multistage Investment Timing Game in Offshore Petroleum Production,” *The Review of Economics and Statistics*, 95, 1601–1621.
- LORIS, N. (2013): “Energy Production on Federal Lands: Handing Keys Over to the States.” <http://westernenergyalliance.org/wp-content/uploads/2012/05/Final-Combined-ES-Econ-Impacts-of-OG-Dev-on-Fed-Lands-in-the-West.pdf>.
- MACKINNON, J. G. AND L. MAGEE (1990): “Transforming the Dependent Variable in Regression Models,” *International Economic Review*, 31, 315–39.
- MCKIE, J. W. (1960): “Market Structure and Uncertainty in Oil and Gas Exploration,” *Quarterly Journal of Economics*, 74, 543–571.
- NELSON, P. H., S. M. EWALD, S. L. SANTUS, AND P. K. TRAINOR (2009): *Gas, Oil, and Water Production from Jonah, Pinedale, Greater Wamsutter, and Stagecoach Draw Fields in the Greater Green River Basin, Wyoming*, United States Geological Survey. Open-File Report 2009–1290.
- PENDERY, B. M. (2010): “BLM’s Retained Rights: How Requiring Environmental Protection Fulfills Oil and Gas Lease Obligations,” *Lewis and Clark Law School Journal of Environmental Law*, <http://elawreview.org/articles/volume-40/issue-40-2/>.
- SNOW, N. (2013): *House votes to restrict federal regulation of fracking*, Oil & Gas Journal <http://www.ogj.com/articles/2013/11/house-votes-to-restrict-federal-regulation-of-fracing.html>.
- (2014): *Witnesses attack, defend recent BLM policies at House hearing*, Oil & Gas Journal <http://www.ogj.com/articles/2014/02/witnesses-attack-defend-recent-blm-policies-at-house-hearing.html>.
- SOUDER, J. A. AND S. K. FAIRFAX (1996): *State Trust Lands: History, Management, and Sustainable Use*, University Press of Kansas.
- SWCA ENVIRONMENTAL CONSULTANTS (2012): *Economic Impacts of Oil and Gas Development on Public Lands in the West*, <http://westernenergyalliance.org/wp-content/uploads/2012/05/Final-Combined-ES-Econ-Impacts-of-OG-Dev-on-Fed-Lands-in-the-West.pdf>.
- TAO, T. AND V. H. VU (2006): *Additive Combinatorics*, Cambridge University Press.

U.S. DEPARTMENT OF THE INTERIOR, BUREAU OF LAND MANAGEMENT (2012): *Public Land Statistics 2011*, [http://www.blm.gov/public\\_land\\_statistics/pls11/pls2011.pdf](http://www.blm.gov/public_land_statistics/pls11/pls2011.pdf).

VISSING, A. (2016): “An Empirical Study of Market Power in Natural Gas Leasing,” Working Paper.

WYOMING GEOLOGICAL ASSOCIATION (1979): *Wyoming Oil and Gas Fields Symposium: Greater Green River Basin*, vol. 1, 2, Wyoming Geological Association.

——— (1992): *Wyoming Oil and Gas Fields Symposium: Greater Green River Basin and Overthrust Belt*, Wyoming Geological Association.