ASSA JMP Talk

Mark Agerton

2.5 minute intro

- I'm an energy economist with interests in dynamic investment problems and applied econometrics.
- My job market paper is about how oil & gas companies learn where to **concentrate drilling** in unconventional shale formations. We associate this kind of drilling with fraccing.
- **Since 2007**, US gas production from shale has increased more than tenfold. A lot of this because wells drilled today produce more than they used to.
- To date, the literature has focused on the idea that these **productivity improvements** are the result of firms learning *how* to drill better. These papers ignore the possibility that firms may have also been learning about *where* to drill, and increasingly concentrated drilling on locations with more productive geology.
- Empirically, learning *how* to drill and *where* to drill both result in increasing well productivity in the short run. However, in the long run, if we've learned *where* to drill and consequently deplete the best spots first, we'll have to drill locations that produce less. This could lead to over-estimates of long-run supply.
- To assess the potential for this, I'm studying (1) what firms know about how productive a location will be before they drill it, and (2) how what firms learn helps them concentrate drilling in more productive locations.
- My main **empirical challenge** is that we can't see what firms believe. I'm going to identify and estimate firms' information by incorporating learning into a dynamic discrete choice model of firms' investment decisions.
- To estimate the model, I assemble a **rich geospatial dataset** on Louisiana's Haynesville shale during the 2003–2015 period. The dataset draws on several public and private sources of data and includes each mineral lease, well, and natural gas production stream in the Haynesville.
- I find that companies' **initial signals** before drilling are not very informative. However, they **learn a lot about geological quality** from the initial wells they drill. With this information, they can concentrate subsequent drilling on better locations.
- My results imply (1) that learning *where* to drill is likely to have contributed to our productivity increases and (2) that it may lead to earlier-than-anticipated depletion in the long-run.

More detailed description

Model

- Let me tell you about the model.

  My unit of observation is a one square-mile parcel of land called a "section." The regulator exogenously defines these and allows only one firm to make decisions in each one. Each section can hold a finite number of wells, and the quantity produced by each one depends on that section's geological quality.

  By *where* to drill, what I mean is a firm's choice to drill wells in a section based on what it believes about the geology there.

  I assume that learning takes a simple form. Before the first well is drilled in a section, firms only have a noisy signal about quality. To learn actual quality, they have to drill an initial well. Once they do, firms can then use the actual quality of the section to decide whether to drill more wells. Initial signals and actual quality are latent variables.

  I'm interested in two things: how much variation in a section's actual quality explains variation in production volumes, and the precision of firms' initial signals about that quality.

- The signal and actual quality affect three outcomes: royalty rates, drilling decisions, and production volumes.

- First, **royalty rates**.

  I start observing firms once they have signed a mineral lease contract that gives them the right to drill and produce. Each lease specifies a royalty rate, which is the share of production revenues firms have to pay mineral owners.

  I assume that royalty rates are a function of firms' willingness to pay, including their initial signal, and landowners' willingness to accept drilling, which is based on the alternative uses of the land. I make an identifying assumption that landowners' willingness to accept drilling doesn't affect the profitability of drilling except through royalty rates.

- Second, **firms' drilling decisions**.

  This is the heart of the model in which I specify and solve firms' dynamic optimization problem of when to drill each well. Dynamics are important because firms' decisions today affect their choice set and information tomorrow.

  Every period, the firm has to decide on the discrete number of wells to drill. All the wells are the same, ex-ante, so the choice is just how many to drill. The static payoff from drilling is the expected revenue minus a fixed, per-well cost and a random cost-shock.

  The firm's first well is a little different than its later wells because the firm only has a noisy signal about the quality at this point, and it knows that the first well will reveal the actual quality of the section.

  The firm has rational expectations with respect to an exogenous, Markov price process, its current information about geology, and the number of drilling opportunities left.

- Finally, I see the **monthly production** from each well.

  Production depends on the actual quality of each section, an unpredictable well-productivity shock, and natural decline.

Identification

- Here's some intuition behind how I identify learning.
- Suppose I'm a firm who gets a high signal. This means (1) I'm willing to pay more for royalty rates conditional on the landowners willingness to accept, and (2) I drill my first well quickly, conditional on prices.
- Now suppose I learn that the actual quality of the section is high. That means (1) I drill more additional wells more quickly, and (2) that the wells produce a lot.
- The correlation between signal and actual quality is going to be identified by 2 things. First, we can compare the values of signal and true quality. Second, firms have less incentive to change the timing of initial drilling behavior when signals are less informative.

Data

- Let me tell you about my data on Louisiana's Haynesville shale now. There are two new things I'm doing.

- First, I'm linking the terms of individual, private mineral leases with the individual drilling decisions they affect and the resulting production outcomes. There are some papers that look at mineral leases, and others that look at drilling and production, but no paper actually combines the datasets and incorporates the way mineral leases shape firms' investment problems into estimation.

- I'm able to do this by exploiting the way the regulator requires all firms and leases in each section to participate in each well.

- To assemble my dataset, I take spatial data on sections and wells from Louisiana's regulator.

- I use proprietary data on mineral leases that I obtained from a company called Drillinginfo. Drillinginfo visits courthouses where mineral lease contracts are stored and records the key characteristics and location of each one. I also use their data on well-level production.

- I use Census Block Group characteristics and satellite-derived landcover data to form the variables that affect landowners' willingness to accept.

- The second data-innovation in my paper is that I include an exogenous, geology based measure of resource quantity for each section as a control variable in all three parts of the model. This captures some of the coarse geology information available to firms. To my knowledge, no other empirical energy economics uses a geology variable with such high spatial resolution.

- Finally, I take natural gas prices to be a weighted average of the forward curve.

Estimation

- Each section contributes to my likelihood in 3 ways: royalty rates, drilling decisions, and production outcomes.
- Following the Rust (1987) nested fixed point approach, I solve firms' dynamic program for each likelihood iteration in order to evaluate the probability of each observed choice.
- Since I don't see the signal or actual quality, I integrate them out to recover their joint distribution.

Findings

- I find that about half of the residual variance in production can be explained by the actual quality of each section. However, there is a surprisingly low correlation between initial signal and actual quality—around 0.16.
- These estimates have implications for the dynamics of drilling and productivity over the lifecycle of a shale play.
    - First, firms can't target initial drilling very precisely, so initial wells will be less affected by the selection issue I talked about.
    - Second, the informational value of initial wells is very high since firms can resolve half of the uncertainty about production. This speeds up initial drilling, contributing to a boom.
    - Third, in aggregate, firms' information about quality over space will improve. So, drilling will get concentrated in the best undrilled sections. Average production will initially go up because of increasing selection.
    - Fourth, just as increasing selection led to increasing productivity in the short run, in the long run, it will imply falling productivity as firms have to drill the worse locations left.
- So, the selection and depletion effects I mentioned are real concerns. We need to account for these, or we risk making overly optimistic long-run supply forecasts.
- It's great that papers like Covert (2015) and Montgomery and O'Sullivan (2017) control for high degrees of spatial variability. However, what they can't speak to is how precisely firms can target geology. Firm behavior is what affects supply, which is the issue my paper addresses.

Literature review

- In summary, what I'm doing in this paper is estimating a model of firms' drilling decisions that allows me to identify what they learn about the inherent quality of the underlying geology as they drill different locations. That learning determines the degree to which selection on geology contributes to the dynamics of shale productivity.
- This differs from the existing literature because of (1) the question I ask, and (2) the higher demands on data and estimation.
- The literature to date has focused on firms learning about *how* to drill. Covert (2015) and Fitzgerald (2015) are two particularly relevant papers in this regard. Both estimate relationships between well inputs and hydrocarbon outputs. To do this, they have to ignore the possibility that firms might be learning *where* to drill, not just *how* to drill. By ignoring this feature I show that firms are learning *where* to drill, and so long-run forecasts based on a *how* to drill concept will be biased upward.

- Answering my question is best done with a dynamic discrete choice framework, which has been used with success in Kellogg (2014) and Muehlenbachs (2015). By adding learning, I'm significantly increasing the demands on the model, computation, and data.
- In particular, to identify learning, I need to link drilling decisions with their respective mineral leases and production outcomes both in the data and the estimation procedure. This is new in the set of shale investment papers.

Potential questions

- How big could the effect be?

  This is clearly a very important question, and I'm currently working on some simulations to address this. However, I don't have a number I'm comfortable sharing yet.

- Why not *how* vs *where*—that seems to be the important question?

  Yes, to be very clear, what I'm not doing is decomposing productivity into learning about how vs where to drill. Both Covert (2015) and Montogmery and O'Sullivan (2017) do use sophisticated spatial econometrics with more detailed well inputs to get at this how/where tradeoff. However, they can't say anything about how geology affects firms' decisions. Without that, they can't think about long-run forecasts that incorporate depletion.

- What about spatial correlation?

  There are two concerns here.

  The first is that actual quality is correlated across space. This is true, but I include a geology-based spatial variable that will soak up a lot of spatial variation.

  The second concern is informational externalities that mean firms learn from their neighbors. That's going to make firms delay initial drilling that reveals information, but not later drilling. That limits the effect of information externalities to the first well. Furthermore, mineral leases usually have expiration dates that force the firms to use-it-or-lose-it. That will also limit strategic incentives to delay. Finally, we know from Covert (2015) that firms don't do a great job of learning from what competitors do.

  Information externalities is definitely an interesting one I'd like to look at more. A dataset like mine is the right one to do this.

- What about correlation between geological quality and the way a well is drilled, e.g., boosting production in especially productive locations by using more sand and water?

  What you're addressing is the *intensive* margin—how long do I make a well, how much sand and water, etc? That's going to be much more related to the *how* to drill question I ignore. Firms are going to learn about *where* to drill by making decisions on the *extensive* margin.

  Ignoring *how* to drill is going to mean I overestimate learning about *where* to drill. So, you'll need to take my estimates as an upper bound for the role of geology, just as the *how* to drill papers are a lower bound.

- Why this particular form of learning?

Empirically, what we see is that firms tend to drill wells the first well and then stop drilling for a while. If they do drill more wells, those are usually drilled one right after the other. This pattern suggests that firms don't update their beliefs much once they learn from the first well.

- What are the IID action-specific shocks
  - Weather, rig availability, etc