How to Read an Academic Article

With a discussion of Muehlenbachs, Spiller, and Timmins (2015)

Prof. Richard Sweeney

ECON3391.01, Boston College

Strategies for reading an academic article

- Have you done this in other classes?
- What aspects do you find easiest / hardest?

Discussion of Muehlenbachs, Spiller, and Timmins (1999)

"The Housing Market Impacts of Shale Gas Development"

What is the research question?

What is the research question?

- What is the "impact" of shale gas development

How to Read an Academic Article

Prof. Richard Sweeney

Intro
Research Question
Background
What this paper does
Why does this matter?
Data
Empirical Strategy
Details
Results
Wrap up

What is the research question?

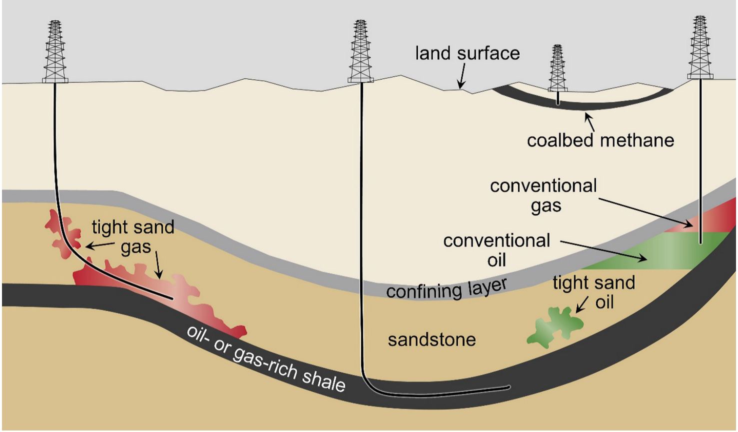
- What is the "impact" of shale gas development
- More specifically: Impact of X on house prices:
  - Threat of water contamination (-)
  - "visibility" (-)
  - Local economy (+)

What is hydraulic fracturing?

- Technique for increasing oil and gas yields
- "Fracking" involves injecting high pressure fluids to break up rock formations
- The fluid generally consists of
  - water
  - chemicals
  - proppant (sand/ synthetic beads that hold open fractures)
- Hydraulic fracturing has been used for decades.

Recent boom due to the combination with horizontal drilling

- Allowed for economical extraction from coalbeds, shale and tight formations

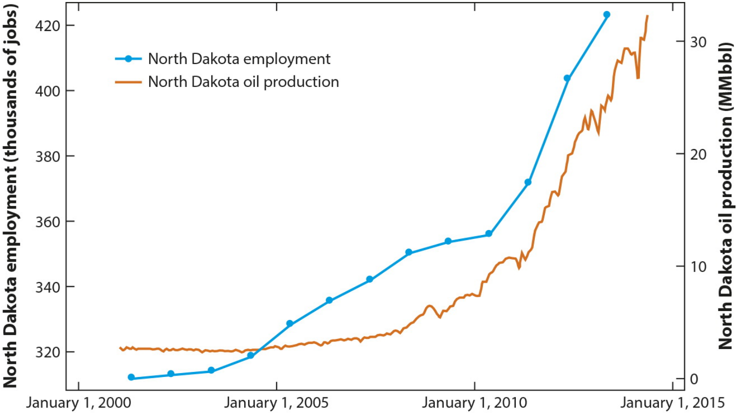


Source: EPA (2015)

What have you heard about fracking?

Many benefits of shale gas boom

- Shale accounted for over 40% of US gas production in 2014
- US pays considerably less for gas than the rest of the world
- Positive externality: offset coal
- Shale has also revitalized many rural US communities



Mason CF, et al. 2015.
Annu. Rev. Resour. Econ. 7:269–289

How to Read an Academic Article

Prof. Richard Sweeney

Intro

Research Question

Background

What this paper does

Why does this matter?

Data

Empirical Strategy

Details

Results

Wrap up

Hydraulic fracturing also involves many negative externalities

Hydraulic fracturing also involves many negative externalities

What risks are most salient to you?

Drilling process is destructive

- Have to bring in large trucks, heavy equipment
- Often have to build gathering pipelines
- Can fragment landscape and have long lasting impacts on habitat

Water contamination

- Most media coverage has focused on the water impacts of fracking
  - see, for example, The New York Times “Drilling Down" series
- Fracking uses significant amounts of water and generates large amounts of waste water
  - this wastewater needs to be disposed of and treated properly
- Groundwater depletion
- Groundwater contamination
  - either leaking well casings or seepage from surface storage pits
  - wellbores often traverse drinking-water aquifers

This paper attempts to disentangle these effects using house price data

Goal is to get an estimate of people's true willingess to pay (in $) using revealed preference

ie they want to make a statement like: "The average person in Pennsylvania would be willing to pay / accept $X to allow / remove a fracked well near their home."

Using house prices to infer preferences for environmental amenities

- How much do people value a good? For normal goods the answer is easy: Look at the price (estimate demand)
- Environmental externalities are inefficiently supplied because there is no price. That also makes them difficult to value.
- MST use hedonic property method to solve this problem.
- Access to safe reliable drinking water is an important home attribute
- If people know the impact of fracking on water quality, changes in home prices should reflect their willingness to pay for clean water
- Extra challenge in this setting: Economic activity from drilling could also increase local property values
    - Landowners receive up to thousands of dollars in (unobserved) bonus payments and 12.5-21% royalty payments per unit of gas extracted

How does the research question relate to the existing literature?

- Science studies on how much these measures are changing, drinking water tests, etc.
- Some great reporting on some affected people/ communities, but little large scale evidence
- Some surveys and opinion polls about how people feel
- Also unclear how to translate those studies into $'s
- A couple econometric papers have also looked at house prices

How does the research question relate to the existing literature?

- Science studies on how much these measures are changing, drinking water tests, etc.
- Some great reporting on some affected people/ communities, but little large scale evidence
  - *Anecdotes are important for generating questions, but often suggest testable implications if widespread*
- Some surveys and opinion polls about how people feel
  - *Economists generally prefer revealed preference.*
- Also unclear how to translate those studies into $'s
  - *Why is it important to get these in dollars?*
- A couple econometric papers have also looked at house prices
  - *"A major obstacle to accurately estimating the impact of shale gas development on surrounding homes is the presence of correlated unobservables"*

Why is this question worth asking?

Why is this question worth asking?

Many places have banned fracking outright (CA,MD,NY)

President Trump: "The shale energy revolution will unleash massive wealth for America.... I think probably no other business has been [more] affected by regulation than [fracking]"

President Biden placed a moratorium on federal drilling. Considerable pressure on his left to go further.

Potentially effected population

- Between 2000 and 2013, approximately 9.4 million people lived within one mile of a hydraulically fractured well
- Approximately 6,800 sources of drinking water for public water systems were located within one mile of at least one hydraulically fractured well during the same period.
- These drinking water sources served more than 8.6 million people year-round in 2013.

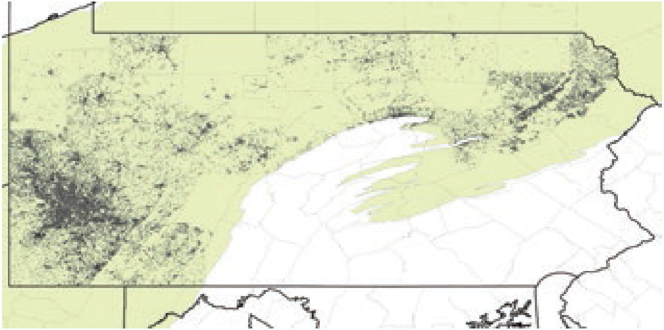
If you had an ideal data set for answering the research question, what would it look like?

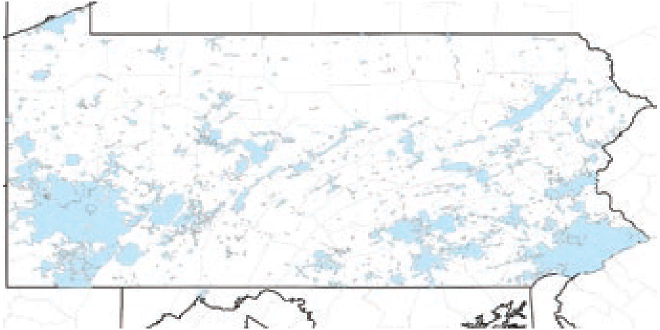
And how would you use it?

What is the data used in the paper? How was it collected?

What is the data used in the paper? How was it collected?

- Transaction records of all PA properties sold 1995-2012 (Corelogic)
- Drilling locations and dates from PADEP
  - data contain 6,260 wellbores which MST group into 3,167 well pads
- Also observed quantity produced from each well
- Use GIS Viewshed tool to predict how many wells are within eyesight of each property





What are the key variables?

Every paper will have a summary of the key variables

Table 1—Summary Statistics by Sample

	Full sample		Boundary subsample	
	Mean	SD	Mean	SD
Transaction price (k 2012 dollars)	134	(98.4)	120	(92.1)
Age of house	55.7	(32.1)	61.3	(34.9)
Total living area (1,000 sq ft)	1.59	(0.67)	1.54	(0.634)
No. bathrooms	1.82	(0.852)	1.68	(0.799)
No. bedrooms	2.96	(0.933)	2.91	(0.984)
Lot size (acres)	0.578	(3.9)	0.53	(4.5)
Distance to nearest MSA (km)	22.3	(12.4)	26.4	(13.4)
Groundwater dependent	0.0771	(0.267)	0.0563	(0.231)
Distance to closest well pad (km)	11.7	(5.35)	11.2	(5.5)
Pads in 1 km	0.00329	(0.081)	0.00596	(0.113)
Pads in 1.5 km	0.00855	(0.164)	0.015	(0.226)
Pads in 2 km	0.0178	(0.289)	0.0314	(0.401)
Pads in 20 km	4.73	(18.1)	5.11	(21)
Pads in view in 1 km	0.000474	(0.024)	0.000844	(0.0325)
Pads in view in 1.5 km	0.00113	(0.0425)	0.0022	(0.0599)
Pads in view in 2 km	0.00189	(0.0671)	0.00368	(0.0955)
Producing pads in 1 km	0.00263	(0.0736)	0.0049	(0.104)
Producing pads in 1.5 km	0.00694	(0.152)	0.0127	(0.214)
Producing pads in 2 km	0.0147	(0.274)	0.0273	(0.388)
Observations	229,946		66,327	

Notes: Samples are the same as those used in our main estimation (i.e., only include properties that were sold more than once during the sample period). The boundary subsample includes only properties in the narrow band on either side of the border of the public water service area.

What is a unit of observation in this paper?

What is a unit of observation in this paper?

Home - sale.

What is the "empirical strategy"?

This is another way of asking: why do the authors believe their estimates will uncover the causal effect of fracking on home values?

What is the simplest regression you could write down to answer the research question? (What is the Y and the main X's?)'

What is the "empirical strategy"?

This is another way of asking: why do the authors believe their estimates will uncover the causal effect of fracking on home values?

What is the simplest regression you could write down to answer the research question? (What is the Y and the main X's?)'

Why is that problematic in this case?

Empirical challenge: wells are not located randomly

The willingness to allow drilling may be correlated with other attributes that affect housing values

- what direction do you think this bias will go?

Empirical challenge: wells are not located randomly

The willingness to allow drilling may be correlated with other attributes that affect housing values

- what direction do you think this bias will go?

What is their strategy for dealing with this?

Empirical challenge: wells are not located randomly

The willingness to allow drilling may be correlated with other attributes that affect housing values

- what direction do you think this bias will go?

What is their strategy for dealing with this?

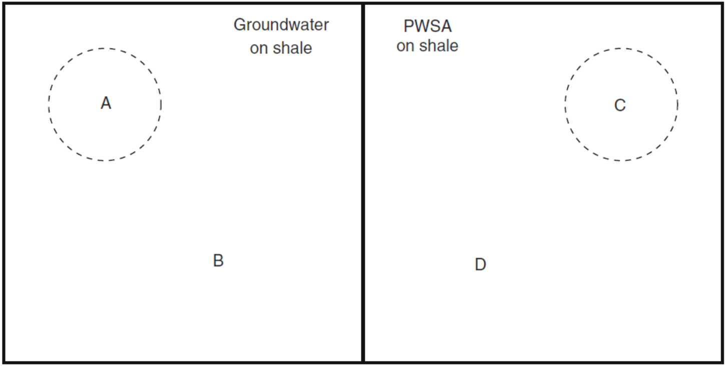
- MST attempt to overcome this by looking at how the same property's price changes over time as drilling expands around it
- This is known as **difference in differences** (more next week)

Impact categories

- Adjacency effects - costs and benefits of being near a well independent of water impacts
  - costs: noise, air pollution, visual disruptions, etc
  - benefits: lease and royalty payments
- Groundwater contamination risk (GWCR)
  - some properties rely on groundwater, others use publicly treated water
- Vicinity effects - general costs and benefits in a wider (e.g. 20 km) area
  - costs: traffic, accidents, etc
  - benefits: increased employment, spending, public finances, etc

MST identification strategy

Figure: Types of Areas Examined



- circles represent adjacency effect buffers
- rectangles distinguish areas that rely on groundwater for drinking

Empirical strategy

- Comparing change in property value before and after fracking generates different price changes by property type:

$$\begin{aligned}
\Delta P_A &= \Delta Adjacency + \Delta GWCR + \Delta Vicinity_{GW} + \Delta Macro \\
\Delta P_B &= \Delta Vicinity_{GW} + \Delta Macro \\
\Delta P_C &= \Delta Adjacency + \Delta Vicinity_{PWSA} + \Delta Macro \\
\Delta P_D &= \Delta Vicinity_{PWSA} + \Delta Macro,
\end{aligned}$$

- Groundwater effect identified using triple difference-in-differences estimator:

$$\Delta GWCR_{DDD} = [\Delta P_A - \Delta P_B] - [\Delta P_C - \Delta P_D]$$

What concerns might you have about this identification strategy?

What concerns might you have about this identification strategy?

- why do some houses lease?

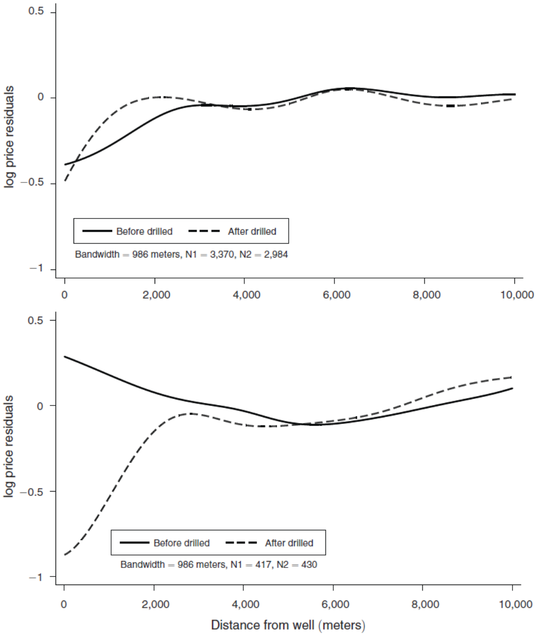
What are the main results of the paper?

Visualizing the results

Good empirical papers will illustrate their empirical strategy and results graphically before diving into the econometrics.

In this case, simple plot the residuals of house price within an area, in a year, as a function of distance to a well.

Gradient for PWSA (top) vs groundwater areas



MST Results: Groundwater impacts

TABLE 2—LOG SALE PRICE ON WELL PADS

	$K \leq 1$ km		$K \leq 1.5$ km		$K \leq 2$ km	
	Full (1)	Boundary (2)	Full (3)	Boundary (4)	Full (5)	Boundary (6)
Panel A. County-year fixed effects						
Pads in K km	0.028	0.026	0.029**	0.034*	0.016**	0.018*
	(0.025)	(0.035)	(0.014)	(0.02)	(6.9e-03)	(0.01)
(Pads in K km) × GW	−0.062	−0.165**	−0.042*	−0.099***	−0.023	−0.013
	(0.046)	(0.072)	(0.025)	(0.036)	(0.02)	(0.052)
Pads in 20 km	−7.8e-04***	−8.1e-04	−8.3e-04***	−9.3e-04*	−8.4e-04***	−9.4e-04*
	(3.0e-04)	(5.3e-04)	(3.0e-04)	(5.5e-04)	(3.0e-04)	(5.6e-04)
(Pads in 20 km) × GW	6.6e-04	2.0e-03***	7.0e-04	2.0e-03***	7.1e-04	1.7e-03**
	(4.7e-04)	(7.0e-04)	(4.9e-04)	(6.8e-04)	(5.2e-04)	(6.8e-04)
Property effects	Yes	Yes	Yes	Yes	Yes	Yes
County-year effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	229,946	66,327	229,946	66,327	229,946	66,327
p-value ($\alpha_3 + \alpha_4 = 0$)	0.414	0.051	0.544	0.090	0.740	0.919
Avg. pads in K km	0.003	0.006	0.009	0.015	0.018	0.031
Avg. pads in 20 km	4.725	5.108	4.725	5.108	4.725	5.108

TABLE 3—ADJACENCY EFFECTS

	$K = 1$ km (1) ln(price)	$K = 1.5$ km (2) ln(price)	$K = 2$ km (3) ln(price)
Panel A. log sale price on well pads in view			
Visible well pads in K km	1.1e-03	−0.019	0.019
	(0.072)	(0.058)	(0.035)
Not-visible well pads in K km	0.03	0.036***	0.015**
	(0.028)	(0.013)	(6.5e-03)
Pads in 20 km	−6.0e-04*	−6.4e-04*	−6.5e-04*
	(3.3e-04)	(3.3e-04)	(3.3e-04)
Panel B. log sale price on productive wells			
Unproductive pads in K km	−0.052	−0.043	−0.054*
	(0.077)	(0.035)	(0.03)
Producing pads in K km	0.044**	0.038***	0.02***
	(0.02)	(0.013)	(5.8e-03)
Pads in 20 km	−6.0e-04*	−6.4e-04*	−6.3e-04*
	(3.3e-04)	(3.3e-04)	(3.3e-04)
Panel C. log sale price on timing of wellbores			
Old bores (drilled $>$ 365 days) in K km	0.021	0.023**	0.011**
	(0.018)	(9.8e-03)	(4.4e-03)
New bores (drilled $\leq$ 365 days) in K km	−4.4e-03	−9.7e-03	−3.3e-04
	(0.029)	(0.013)	(8.0e-03)
Old undrilled permits ($>$ 365 days) in K km	0.055**	0.022	0.011
	(0.025)	(0.014)	(0.012)
New undrilled permits ($\leq$ 365 days) in K km	0.04*	7.2e-03	7.2e-03
	(0.023)	(0.014)	(7.9e-03)
Pads in 20 km	−6.0e-04*	−6.2e-04*	−6.3e-04*
	(3.3e-04)	(3.3e-04)	(3.3e-04)
Property effects	Yes	Yes	Yes
County-year effects	Yes	Yes	Yes
Quarter effects	Yes	Yes	Yes
Observations	212,207	212,207	212,207

Results Summary

- Estimate risk of groundwater contamination negatively affects house values within 1-1.5 km of a fracked well in PA
    - Note this impact measures the *perceived* impact
- Find that households that rely on piped water actually benefited from being near wells
    - results appear to be driven by royalty payments
    - positive finding is explained by wells that were drilled over a year prior to the sale (after drilling costs)
    - Only find these positive effects for wells that are not visible from the property
- Average annual loss for groundwater dependent homes within 1.5 km of a well is $30,167
- This is larger than the average annual gain for piped water properties within 1.5 km of a well of $4,802

What are the economic/ policy implications of these results?

What did you learn from this paper?

Intro
Research Question
Background
What this paper does
Why does this matter?
Data
Empirical Strategy
Details
Results
Wrap up

What questions does this paper leave unanswered? How might you answer them?

Announcements