The Costs of Induced Seismicity: A Hedonic Anlalysis *

Neil E. Metz
Department of Economics
University of Central Oklahoma
nmetz@uco.edu

Travis Roach
Department of Economics
University of Central Oklahoma
troach2@uco.edu

Jordan A. Williams
Oklahoma Geological Survey
University of Oklahoma
jordan.a.williams@ou.edu

June 22, 2017

Abstract

New developments in drilling technology and the subsequent increase in hydraulic fracturing has brought unprecedented change to energy markets domestically and internationally. Unintended effects of this extraction technique have been felt, quite literally, due to induced seismicity from waste water injection. This research measures the costs of induced seismicity through changes in home prices using a hedonic price analysis within a differences-in-differences framework. We find the revealed cost to be between 3.15%-4.7% of home values, a $6,660 reduction at the average.

JEL Codes: H23, R31, R32, Q49
Keywords: Earthquakes, Hydraulic Fracturing, Oil and Gas Extraction

*This research was supported through funding from the NSF Oklahoma EPSCoR ROA+ program

1 Introduction and Background

Along with an unprecedented supply of new oil and natural gas in the United States, the process of hydraulic fracturing[1] has drastically increased the supply of produced water[2]. Much of this produced water is pumped back underground or "injected" as wastewater into class II injection wells used by the oil and gas industry, though some areas are recycling and reusing the water in various industrial practices. In Oklahoma, a major producer of oil and gas, waste water injection is a common practice. In 2009 over 849 million barrels of wastewater were injected underground. This amount has since grown dramatically with 1,538 million barrels injected in 2015 alone. The ramifications of such practices are being felt through increased induced seismic activity. In 2009 there were 20 earthquakes that registered as a magnitude 3.0 or greater. The amount of 3.0+ earthquakes grew consistently to 581 events in 2015 (Murray (2015)).

The Oklahoma experience is not anecdotal, and there is broad scientific consensus that swarms of induced earthquakes are correlated with injection (Ellsworth (2013); McGarr et al. (2015); Weingarten et al. (2015)). The increase in earthquakes related to waste water disposal has been established within the scientific literature for nearly 50 years. Healy et al. (1968) showed how high pressure injection caused earthquakes to occur in the Denver area. Following the Denver earthquakes, scientists were later able to run a controlled experiment and influence the amount of earthquakes by changing fluid pressure at four wells in Rangely, Colorado (Raleigh et al. (1976)). Other research has reported trends between naturally occurring and induced earthquakes. Studies have shown that the maximum magnitude of induced earthquakes may be smaller than what is seen with natural earthquakes, but they also suggest that induced earthquakes can trigger larger earthquakes on known or unknown faults (McGarr et al. (2015)). Additionally, induced earthquakes tend to occur in swarms (many happening in the same area in quick succession) and they tend to happen at shallower depths than natural earthquakes (Gomberg and Wolf (1999); McNamara et al. (2015)).

While induced seismicity has been an ongoing research subject geologically, the impacts through economic channels are less well-defined. Indeed, the recent, unanticipated nature of such earthquake activity lends itself to a 'natural experiment' setting where these unintended costs may be calculated. One mechanism to calculate the costs of induced seismicity is to measure changes in home prices. Given consumer theory, and assuming households are mobile, one would expect that homes in 'high risk areas' (those that have witnessed relatively

[1] Colloquially referred to as "fracking".

[2] excess saltwater and waste water produced during the drilling process

more earthquakes) would be priced lower than equivalent homes in lower-risk areas. For example, if a swarm of earthquakes occurred in north-east Oklahoma City, one would have to be 'compensated' through reduced home prices to bear the risk of experiencing future earthquake swarms. These equalizing differences may be recovered by estimating a hedonic price model in the tradition of Rosen (1974).

Hedonic pricing models are a common way to price externalities, and a number of studies have used the technique to study the 'fracking' boom in particular. Muehlenbachs et al. (2015) used differences in home water supply (e.g. well water or public utility) and distance to hydraulic fracturing sites to quantify externalities of shale development. These authors found there to be a positive effect for homes that are connected to the public utility water system due to royalties being capitalized into the home's price, but that the effect of increased royalties goes away if a home is reliant on well-water. Their conclusion is that the risk of ground water contamination from the hydraulic fracturing process, even if it may be misinformed, is capitalized into the home price as a compensating risk differential. Boslett et al. (2016) sought to quantify the costs and benefits of shale gas development through home prices by using state-to-state policy differences in the Marcellus shale region. These authors found that homes outside of New York, in which there is no fracking ban, witnessed a steeper increase in value because they were able to receive royalty payments.[3] Of particular interest to this study is a working paper by Cheung et al. (2016). These authors also seek to measure the effect of induced seismicity through home prices in Oklahoma. While the present paper comes to a similar conclusion quantitatively as Cheung et al. (2016), we differ in a number of meaningful ways. First, the present study encompasses a longer time-frame of sales surrounding the onset of seismic activity while the Cheung et al. (2016) study encompasses a larger area using home sales from across the state.[4] Our focus on a single county aids in identification of the causal effect on home prices as discussed below. Second, the present research uses a treatment effect model whereas Cheung et al. (2016) relies on indicator variables for various magnitude events, and in an alternate specification the cumulative count of earthquakes at various magnitudes. Regardless of the differences in modeling strategies, we both find that the onset of earthquakes in Oklahoma has had a negative effect on home prices.

[3]We note here that royalty payment differences are not relevant to the present study as they have for the former two studies because these studies cover an area in which there was little to no oil and gas activity historically. Oklahoma has a long history of oil and gas development, and, almost universally, land and mineral rights are severed.

[4]We also make use of different data sources for home price and characteristics information. Our data comes from state assessors whereas their data is through MLS records

In order to accurately estimate the revealed cost of induced seismicity, we estimate hedonic pricing models using single family home sales in Oklahoma county from 2000 to present.[5] The model uses a differences-in-differences treatment effect framework to isolate the effect of seismicity. This differencing method essentially compares the change in home prices in seismically active regions before and after the onset of the earthquake boom to the change in prices in other areas that were not seismically active while controlling for all relevant home features. This enables us to filter any effects from global or national factors (e.g. changes in the prime rate) and isolate the outcome due to seismic events. Further, by focusing on only Oklahoma county sales we are able to eliminate potential endogeneity from factors like changing employment mix across industries by differencing over time. For example, consider a large industry exposed to earthquakes in one city compared to a different industry located somewhere without earthquakes. Then, consider if one of these industries drastically reduced their labor force and this subsequently effected home demand. If this occurred while earthquakes took place, then the estimates could be biased in either direction. Cheung et al. (2016) account for this type of issue by including information on the distance to Tulsa or Oklahoma City, and by including employment in the extractive industries as regressors.

In addition to standard hedonic pricing models that account for house characteristics (e.g. number of bedrooms and square feet) we also estimate individual home fixed effects models. In these specifications we control for the exact same house being sold at different points in time. This method refines our estimate of the impact of seismicity on home prices by eliminating unseen, or at least undocumented, home attributes. Across specifications we find that homes are now priced differently in earthquake-prone areas.

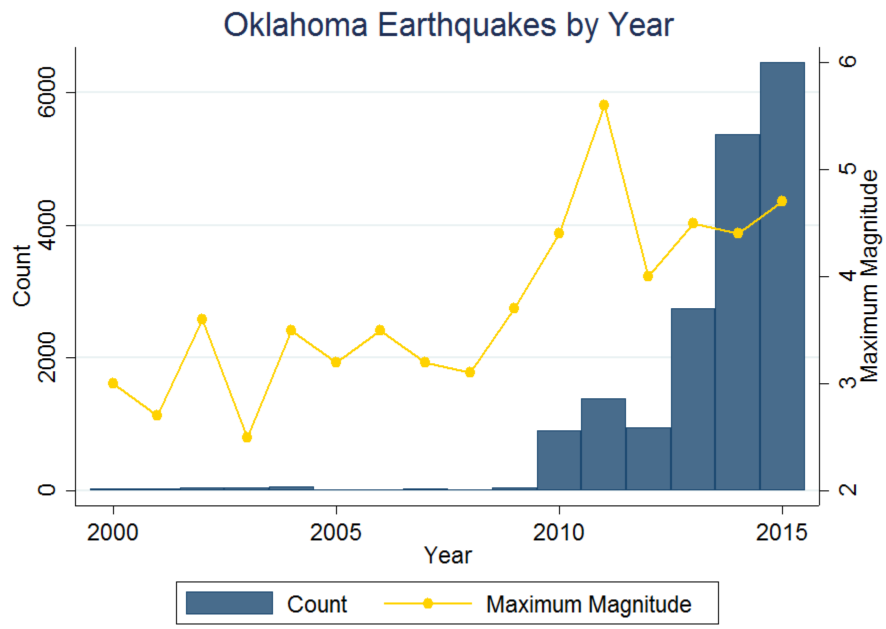
To preview results, the onset of earthquakes has reduced home prices in seismically active regions by 3.15% - 4.7%. This finding is robust to different definitions of a 'seismically active' region.[6] The results from this study confirm the results of Chueng, et al. (2017) while using a separate modeling strategy and more refined set of observations.

The balance of this paper continues as follows: Section 2 discusses the earthquake magnitude and location data as well as the home price and characteristics data; section 3 outlines the hedonic treatment effect model and the various robustness specifications we run; section 4 presents results; and section 5 concludes with remarks pertaining to the interpretation of our results.

[5]Oklahoma County covers 371 square miles of land, and contains the state capital, Oklahoma City. The real-estate makeup in Oklahoma county is very diverse with homes located in historic districts, suburbs, exurbs, etc.

[6]When a smaller circle radius from the earthquake epicenter is used, or a larger threshold for earthquake magnitude.

Figure 1



2 Data Description

2.1 Earthquake Location and Impact Data

Information on earthquake location and magnitude come from the Oklahoma Geological Survey. Figure 1 displays earthquake counts for a 1.0 magnitude or greater and the maximum magnitude for each year since 2000. It is clear that the number of earthquakes per year increased dramatically following 2010 with more than 6000 earthquakes occurring in 2015 alone. Using various radii surrounding earthquake epicenters[7], as well as various magnitude qualifications, we define seismically active regions to merge with our home price and characteristics data. The baseline specification defines a region, more specifically a zipcode, as seismically active if there are more than 50 magnitude 1.0+ earthquakes within a 10km distance of the zipcode's centroid. Evidence from United States Geological Survey's "Did You Feel It?" program shows that magnitude 1.0-2.0 earthquakes are certainly felt within this distance [8]. While lower magnitude events may not cause damage to a home or building, they may still influence the purchasing behavior of those that feel the earthquake or those that learn of the increased seismic activity from the news or their peers. Within our sample 41.3% of home sale observations are within a seismically active region. It is possible that

[7] Haversine formula distances.
[8] *Did You Feel It?* (n.d.)

the quantity of earthquakes is not the driving factor of home price differences, but rather large scale events. We account for this in separate robustness specifications where we define seismically active regions as those zipcode centroids that are within a 10 km distance to a 3.0 magnitude earthquake. We also reduce the distance that the earthquake could affect homes in zipcodes to 5km. Our results remain consistent across seismically active region specifications.

2.2 Home Price Data

Detailed records on single family home sales price(s), location, characteristics, and quality were purchased through the Oklahoma County Assessor. Within the raw assessors data we limit the analysis to homes which are sold multiple times from 2000 to June 2016 leaving us with 94,211 observations.[9] These data allow us to control for individual home characteristics through home characteristics and home fixed effects models. Summary statistics for all control variables are shown below in Table 1.

Table 1. Summary Statistics

	Avg	Std Dev	Min	Max
Home Price	139,272	101,583	30,000	995,000
Home sf	1,767	772.1	500	8,025
Land sf	15,304	34,572	2,000	863,359
Bedrooms	3.088	0.655	1	8
Bathrooms	1.911	0.705	0.750	6.500
Year Built	1970	22.85	1895	2015
Garage Dummy	0.892	0.311	0	1
Seismic Period Dummy	0.354	0.478	0	1
Seismic Region Dummy	0.413	0.492	0	1

Notes: N = 94,211

3 Empirical Specification

We quantify the cost of induced seismicity by measuring differences in home prices before and after earthquakes became prevalent, while controlling for relevant home attributes, with non-seismic regions in the same county serving as a treatment-free control group. The onset of induced seismicity was certainly not a "treatment" in the traditional sense, but the sudden increase in earthquake activity was discrete and unanticipated. In other words, areas of Oklahoma county did not reduce in price prior to earthquakes occurring because little to

[9] We also eliminate home sales with likely errors, such as price per sqft below \$15 and above \$500, and homes with less than 500 sqft. We also restrict the analysis to only include homes with a sale price between \$30k and \$1mil.

no seismic activity occurred prior to 2010, nor were earthquakes expected despite existing scientific evidence.

The treatment effect home price model is shown below in equation (1).

$$\begin{aligned} \ln Price_{it} &= \beta_0 + \beta_1 SeismicRegion_i + \beta_2 SeismicPeriod_t + \beta_3 Region * Period_{it} \\ &+ \sum_{j=1}^{10} \lambda_{0+j} Quality_{it} + \pi Attributes_{it} + \mu_i + \omega_t + \phi_t + \varepsilon_{it} \end{aligned} \quad (1)$$

Where $\ln Price_{it}$ indicates the natural log of the sale price for home i at time t; $SeismicRegion_i$ is a dichotomous indicator variable that is set to one if the home is in a seismically active region; $SeismicPeriod_t$ is a dichotomous indicator variable that is set to one after earthquakes became prevalent in the seismically active region.[10] The interaction term, $Region * Period_{it}$, is the difference-in-differences estimate which captures the effect of earthquake activity on home prices. The coefficient on this variable is able to explain home price differences while eliminating (subtracting away) factors that affect both seismically-active an non-active regions.

We also control for variation in home attributes with the vector of home-characteristics, $Attributes_{it}$, which includes information the square footage of the home, the square footage of the property, number of rooms and bathrooms, the year the home was built, and whether or not the home had a garage. Additionally, the Oklahoma County Assessor data provides a rating of the home quality at the time of sale which can range from 'fair' to 'excellent' with 11 distinct home quality designations. We control for these quality distinctions using separate binary variables for $Quality_{it}$. To capture variation that may be due to school locations, access to parks, etc., we include subdivision fixed effects, μ_i. Finally, we also include year-of-sale and month-of-sale fixed effects, ω_t and ϕ_t, respectively.[11]

We limit our sample to only include data on homes that were sold multiple times over the 15 year time period studied here (i.e. we remove any sales information on homes that were sold only once within 15 years). This allows us to test whether we fully control for individual home characteristics by running separate robustness checks. The last two columns of each table show the results of regressions that make very limiting assumptions. The goal of these

[10] We use January 1, 2010 as a start-date for earthquake activity. Separate models using two quarters before and after the beginning of 2010 yield results that are statistically similar.

[11] In a separate specification that is not included here we included year-month interaction effects to capture idiosyncratic variation that may bias our results. We conclude that our results are robust to including interacted month-year fixed effects variables.

robustness exercises is to show that there are not any individual homes that are biasing the average treatment effect we find in our baseline specification. First, we limit our sample to only those homes that have a sales observation before and after earthquakes began. By doing this we remove any sales record that either has (i) no exposure to earthquakes, or (ii) only has exposure to earthquakes. Second, we estimate a model that has individual home fixed effects (e.g. separate estimates for square footage, number of bedrooms, etc., are eliminated when they do not change over time with this home.) In this model we are able to find the effect of induced seismicity holding even unobserved home attributes constant. In all models we cluster standard errors by household.

4 Results

Table 2 presents our main results. In this table a home is considered to be in a seismically active region if there were more than 50 1.0+ magnitude earthquakes within 10 km of the zipcode centroid for a home. Column 1 shows that homes in seismically active regions decreased in price by appromiately 3.15%. This model does not include year-of-sale or month-of-sale fixed effects, though. Column 2 shows that affected homes decreased in price by 3.23% when these fixed effects are included. To put this in context, this means that the average home sold for $4,387 less in areas exposed to earthquakes than equivalent homes elsewhere.

Columns 3 and 4 make limiting restrictions by narrowing homes in the sample to only those sold both before and after the onset of earthquakes (column 3) and by using individual home fixed effects (column 4). From these results we find that homes prices were 4.15% and 4.71% lower, respectively. Using the fixed effects model this represents a $6,660 reduction in price at the average.

Table 3 presents estimates using an alternative definition of seismically active region. In columns 2, 3, and 4 a seismic region is defined as having a magnitude 3.0 earthquake or greater within 10 km of the zipcode centroid for a home. Column 1 further makes the limitation that there was a 3.0 or larger earthquake within 5 km of the zipcode cenroid for a home. Note, that we include time-of-sale fixed effects in all specifications.

Across specifications we find that home prices have fallen in seismically active areas. Column 2, which is analogous to column 2 of table 2, shows that there was a 3.09% decrease in price due to the onset of earthquakes. When we limit the radius to 5 km we find that homes decreased in price by 2.48%.

Columns 3 and 4 show the estimates when we only use homes sold before and after

Table 2. More than 50 EQs within 10 km

Dep Var: ln Home Price	(1)	(2)	(3)	(4)
Seismic Region	0.0038	0.0076	0.0127	-
	[0.0096]	[0.0091]	[0.0118]	
Seismic Period	0.1427***	-	-	-
	[0.0028]			
Region*Period	**-0.0315*** **	**-0.0323*** **	**-0.0415*** **	**-0.0471*** **
	[0.0038]	[0.0038]	[0.0038]	[0.0037]
ln Home SF	0.4723***	0.4692***	0.4819***	-
	[0.0087]	[0.0085]	[0.0111]	
ln Land SF	0.1008***	0.1029***	0.0956***	-
	[0.0058]	[0.0058]	[0.0072]	
Bedroom	-0.0044*	-0.0032	-0.0048	-
	[0.0027]	[0.0026]	[0.0032]	
Bathroom	0.0646***	0.0630***	0.0590***	-
	[0.0035]	[0.0034]	[0.0041]	
Year Built	0.0032***	0.0031***	0.0029***	-
	[0.0003]	[0.0003]	[0.0003]	
Garage Dummy	0.0187***	0.0243***	0.0296***	-
	[0.0048]	[0.0047]	[0.0063]	
Quality Dummies	Y	Y	Y	Y
Subdivision FE	Y	Y	Y	Y
Month of Sale FE	N	Y	Y	Y
Year of sale FE	N	Y	Y	Y
Individual Home FE	N	N	N	**Y**
Only obs before and after EQ era	N	N	**Y**	N
Observations	94,211	94,211	58,027	94,211
R-Squared	0.821	0.844	0.851	0.881

Notes: Clustered robust standard errors shown. Asterisks denote statistical significance at the traditional levels.

earthquakes began and while controlling for home fixed effects. In these specifications we find that home prices fell in seismically active regions by 4.22% and 3.38%, respectively.

5 Discussion and Conclusion

We identify the costs of induced seismicity by estimating hedonic pricing models using the sudden onset of earthquake activity in Oklahoma. We are able to measure the costs of earthquake activity by comparing the change in single-family home prices before and after earthquakes began with similar homes sold in areas that have not witnessed this change in seismicity while controlling for individual home characteristics.

We find that home prices fell in seismically active regions after the onset of earthquake activity in 2010. This finding is robust to how a seismically active region is defined, as well as when different radii from earthquake epicenters are used. Specifically, we find that homes fell by approximately 3.23%. This estimate is robust to model specifications that only include observations for homes sold both before and after the onset of earthquake activity, as well as model specifications that even eliminate unobserved home characteristics that are fixed over time. When individual home fixed effects are accounted for we find that home prices fell by as much as 4.71%. Given the average price for a home this represents as much as a $6,660 decrease. To put this in context, after removing data on single observation home sales we are left with 94,211 observations. If 20% of these homes are within a seismically active region, then a total of 18,842 homes would be negatively effected by earthquake activity. Using the average sales price this results in an external cost to society of over $125.49 million dollars. We must note, though, that this external cost necessarily represents a lower bound to the true external cost of induced seismicity in Oklahoma because we do not include extra costs that have been incurred to repair public infrastructure such as roads or schools, commercial real estate effects, or even homes outside of Oklahoma county.

The ramifications of induced seismicity are many, and they certainly include damage to building structures and infrastructure. We note, however, that the estimates in the paper do not represent the literal costs to homeowners through foundation repair, brickwork, or other problems. Instead, our estimates indicate the revealed preference of home buyers to pay more for homes in areas that have had less earthquakes (or, rather, less for those in areas that have had earthquakes). This price differential may be affected by a distaste for actual damages that have occurred to a home, a risk-aversion toward future earthquakes, or both. Thus, while the effect of earthquakes on home prices can be precisely estimated, the rationale behind these price differences is vague. Regardless, induced seismicity is a

Table 3. Magnitude 3.0+

Dep Var: ln Home Price	(1)	(2)	(3)	(4)
Seismic Region	0.0145*	-0.0184	-0.0062	-
	[0.0084]	[0.0232]	[0.0293]	
Seismic Period	-	-	-	-
Region*Period	**-0.0248*****	**-0.0309*****	**-0.0422*****	**-0.0338*****
	[0.0039]	[0.0038]	[0.0039]	[0.0038]
ln Home SF	0.4692***	0.4690***	0.4816***	-
	[0.0085]	[0.0085]	[0.0111]	
ln Land SF	0.1028***	0.1029***	0.0956***	-
	[0.0058]	[0.0058]	[0.0072]	
Bedroom	-0.0033	-0.0032	-0.0047	-
	[0.0026]	[0.0026]	[0.0032]	
Bathroom	0.0631***	0.0630***	0.0590***	-
	[0.0034]	[0.0034]	[0.0041]	
Year Built	0.0031***	0.0031***	0.0029***	-
	[0.0003]	[0.0003]	[0.0003]	
Garage Dummy	0.0244***	0.0244***	0.0296***	-
	[0.0047]	[0.0047]	[0.0063]	
Quality Dummies	Y	Y	Y	Y
Subdivision FE	Y	Y	Y	Y
Month of Sale FE	Y	Y	Y	Y
Year of sale FE	Y	Y	Y	Y
Individual Home FE	N	N	N	**Y**
Only obs before and after EQ era	N	N	**Y**	N
Observations	94,211	94,211	58,027	94,211
R-Squared	0.844	0.844	0.851	0.881

Notes: Spec (1) uses 5 km radius, all others at 10 km; Clustered robust standard errors shown in brackets; Asterisks denote statistical significance at the traditional levels.

clear negative externality of wastewater injection. Future market-oriented policies, such as a tax per gallon of waste water injected, would internalize the costs imposed on homeowners, would reduce waste water injection at the margin, and could even pay for future damages.

References

Boslett, A., Guilfoos, T. and Lang, C. (2016). Valuation of expectations: A hedonic study of shale gas development and New Yorks moratorium, *Journal of Environmental Economics and Management* **77**: 14–30.
URL: *http://www.sciencedirect.com/science/article/pii/S0095069615001023*

Cheung, R., Wetherell, D. and Whitaker, S. (2016). Earthquakes and House Prices: Evidence from Oklahoma.
URL: *https://www.clevelandfed.org:443/newsroom and events/publications/working papers/2016 working papers/wp 1631 earthquakes and house prices*

Did You Feel It? (n.d.).
URL: *https://earthquake.usgs.gov/data/dyfi/*

Ellsworth, W. L. (2013). Injection-Induced Earthquakes, *Science* **341**(6142): 1225942.
URL: *http://science.sciencemag.org/content/341/6142/1225942*

Gomberg, J. and Wolf, L. (1999). Possible cause for an improbable earthquake: The 1997 Mw 4.9 southern Alabama earthquake and hydrocarbon recovery, *Geology* **27**(4): 367–370.

Healy, J. H., Rubey, W. W., Griggs, D. T. and Raleigh, C. B. (1968). The Denver Earthquakes, *Science* **161**: 1301–1310.

McGarr, A., Bekins, B., Burkardt, N., Dewey, J., Earle, P., Ellsworth, W., Ge, S., Hickman, S., Holland, A., Majer, E., Rubinstein, J. and Sheehan, A. (2015). Coping with earthquakes induced by fluid injection, *Science* **347**(6224): 830–831.
URL: *http://science.sciencemag.org/content/347/6224/830*

McNamara, D. E., Hayes, G. P., Benz, H. M., Williams, R. A., McMahon, N. D., Aster, R. C., Holland, A., Sickbert, T., Herrmann, R., Briggs, R., Smoczyk, G., Bergman, E. and Earle, P. (2015). Reactivated faulting near Cushing, Oklahoma: Increased potential for a triggered earthquake in an area of United States strategic infrastructure, *Geophysical Research Letters* **42**(20): 2015GL064669.
URL: *http://onlinelibrary.wiley.com/doi/10.1002/2015GL064669/abstract*

Muehlenbachs, L., Spiller, E. and Timmins, C. (2015). The Housing Market Impacts of Shale Gas Development, *American Economic Review* **105**(12): 3633–3659.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20140079*

Murray, K. E. (2015). Class II Saltwater Disposal for 20092014 at the Annual-, State-, and County-Scales by Geologic Zones of Completion, Oklahoma, *Oklahoma Geol. Surv. Open File Rep., OF5-2015* .
URL: *http://ogs.ou.edu/docs/openfile/OF5-2015.pdf*

Raleigh, C. B., Healy, J. H. and Bredehoeft, J. D. (1976). An Experiment in Earthquake Control at Rangely, Colorado, *Science* **191**: 1230–1237.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy* **82**(1): 34–55.
URL: *http://www.journals.uchicago.edu/doi/10.1086/260169*

Weingarten, M., Ge, S., Godt, J. W., Bekins, B. A. and Rubinstein, J. L. (2015). High-rate injection is associated with the increase in US mid-continent seismicity, *Science* **348**(6241): 1336–1340.
URL: *http://science.sciencemag.org/content/348/6241/1336.short*