Online Appendix for
The Local Economic and Welfare Consequences of Hydraulic Fracturing

By Alexander W. Bartik, Janet Currie,
Michael Greenstone, and Christopher R. Knittel*

Exploiting geological variation and timing in the initiation of hydraulic fracturing, we find that fracing leads to sharp increases in oil and gas recovery and improvements in a wide set of economic indicators. There is also evidence of deterioration in local amenities, which may include increases in crime, noise, traffic and declines in health. Using a Rosen-Roback-style spatial equilibrium model to infer the net welfare impacts, we estimate that willingness-to-pay (WTP) for allowing fracing equals about $2,500 per household annually (4.9% of household income), although WTP is heterogeneous, ranging from more than $10,000 to roughly zero across ten shale regions.
JEL: D61, R23, J61, Q51
Keywords: Labor Mobility, Cost Benefit, Resource Booms, Urban Economics

* Bartik: Assistant Professor of Economics, University of Illinois at Urbana-Champaign, abartik@illinois.edu. Currie: Henry Putnam Professor of Economics and Public Affairs and Director of the Center for Health and Well-Being, Princeton University, and NBER, jcurrie@princeton.edu. Greenstone: Milton Friedman Professor in Economics and the College and Director of the Energy Policy Institute at Chicago, University of Chicago, and NBER, mgreenst@uchicago.edu. Knittel: George P. Shultz Professor of Applied Economics, Sloan School of Management, and Director of the Center for Energy and Environmental Policy Research, MIT, and NBER, knittel@mit.edu. This paper has benefited from conversations with John Deutch, Lucas Davis, Manasi Deshpande, Joseph Doyle, Esther Duflo, and Edson Severnini, and seminar participants at the ASSA 2014 Meetings, the Center for Energy and Environmental Policy Research Spring Workshop, the Energy Institute @Haas Summer Energy Camp, MIT, Tufts University, University of Chicago, and the NBER Summer Institute. We thank Mateo Arbelaez, Brian Goggin, Faraz Hayat, Colin McDonnell, Ishan Nath, Lucy Page, Jonathan Petkun, Harshil Sahai, and Daniel Stuart for excellent research assistance. Greenstone gratefully acknowledges financial support from MIT's Provost Office. Bartik gratefully acknowledges funding from the MIT Center for Energy and Environmental Policy Research (CEEPR). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

In this appendix, we go into more detail on our model, empirical strategy, and setting. We also present some supplementary results. The appendix starts with Section I, where we provide a more detailed description of our model of the local welfare impacts of fracing. We then discuss the history and technology of fracing in more detail in Section II, before discussing some of the literature on fracing's potential impacts in Section III. Section IV describes our research design in greater detail and also discusses estimates of the effect of fracing using a propensity score design and compares these estimates to the estimates we present in the main text. We then present robustness checks and extensions of our main strategy in Section V. These results include estimates using alternative measures of exposure to fracing and IV estimates of the effect of hydrocarbon production on economic outcomes. Finally, Section VI includes a detailed discussion of our data sources and how the data were processed.

I. Model Appendix

We present a model closely based on Hornbeck and Moretti (2015) that allows us to develop expressions for the unobserved change in amenities and local welfare as a results of the advent of fracking in terms of the observed changes in rents, income, and population. We assume that household i in location j at time t obtains utility from the consumption of a numéraire good sold on a global market, C_{ijt} (with price normalized to 1), housing, H_{ijt}, location amenities, A_{jt}, and idiosyncratic place-based preferences and moving costs, μ_{ijt}. Assuming Cobb-Douglas utility yields:

$$u_{ijt} = C_{ijt}^{1-\beta} H_{ijt}^{\beta} A_{jt}^{\alpha} \mu_{ijt}^{s}, \tag{1}$$

where β is the share of income households spend on housing, the exponent s measures the size of moving costs or variance of idiosyncratic preferences,[1] and α measures the utility of amenities.

Each consumer in location j at time t earns wage and salary income, w_{jt}, and pays r_{jt} in rent.[2,3] Solving for the consumer's problem and taking logs yields the indirect utility function:

$$v_{ijt} = \ln w_{jt} - \beta \ln r_{jt} + \alpha \ln A_{jt} + s \times \epsilon_{ijt}, \tag{2}$$

where $\epsilon_{ijt} = \ln \mu_{ijt}$.

[1] In the canonical Roback model, these idiosyncratic preferences and moving costs do not exist, i.e. $s = 0$.

[2] We abstract away from differences in housing rents and housing prices. In the simplest model with competitive housing markets, the housing price will equal $\frac{1}{1-\rho}\bar{r}$, where ρ is the discount rate and $\bar{r}$ is the rental price. Therefore a permanent and immediate change in $\bar{r}$ will shift rents and house prices by the same percentage. We also assume that non-labor market income, such as interest and dividend income from lease payments, does not depend on individual location decisions and we abstract away from income effects of non-labor market income on the share of income spent on housing.

[3] Writing wags as w_{jt} makes the simplifying assumption that wages don't depend on worker characteristics, experience, or location decisions.

A key feature of the model is that housing supply is elastic, where inverse housing supply (i.e. X_{jt}) is given by:

$$\ln r_{jt} = \gamma_j + \kappa_j \ln X_{jt}. \tag{3}$$

For intuition on how prices allocate individuals across locations, consider the case where there are only two locations, a and b. Assuming that μ_{ijt} are independently drawn every time period so that future shocks do not affect current decisions, the household's problem simplifies to choosing the location that maximizes current period utility. Consequently, a household chooses to live in location a in period t if and only if $u_{iat} - u_{ibt} > 0$. Defining $\tilde{x} = x_a - x_b$ and using our expression for indirect utility in Equation 2, we can write the household's decision rule as:

$$\widetilde{\ln w_t} - \beta\widetilde{\ln r_t} + \alpha\widetilde{\ln A_t} + s \times \widetilde{\epsilon}_{it} > 0.$$

If $\frac{\mu_{ibt}}{\mu_{iat}} \sim U[0,2]$, then $s \times \widetilde{\epsilon}_{it}$ is distributed exponentially with the shape parameter equal to $\frac{1}{s}$, and we can express the share of households who choose to live in location a in time t as:

$$\frac{N_{at}}{N} = \exp\left[\frac{\widetilde{\ln w_t} - \beta\widetilde{\ln r_t} + \alpha\widetilde{\ln A_t} - s\ln 2}{s}\right].$$

Taking logs yields a linear expression for the log share of households who choose to live in location a:

$$\ln\frac{N_{at}}{N} = \frac{\widetilde{\ln w_t} - \beta\widetilde{\ln r_t} + \alpha\widetilde{\ln A_t} - s\ln 2}{s}. \tag{4}$$

Now, consider the case where an additional assumption holds, namely that location a is "small," relative to location b, such that $\kappa_b \approx 0$. In this case, the difference in log rents simplifies to $\widetilde{\ln r_t} = \xi + \kappa_a \ln N_{at}$, where $\xi = \gamma_a - \gamma_b$ is time-invariant. Imposing this assumption, as well as the assumption that each household consumes one housing unit,[4] combining Equations 3 and 4, and re-arranging yields the following expression for

[4]In light of this assumption that each household consumes one-unit of housing, the flexible H_{ijt} in Equation 1 can be interpreted as spending on all locally produced services, such as housing quality or personal services, whose price is affected by the price of housing. Relaxing the assumption that every household consume one unit of housing, and instead solving for the housing market equilibrium when $X_{jt} = N_{jt} \times H_{ijt}$ is total number of housing units consumed in location j, results in the slightly altered version of Equation 6: $\Delta \ln r_{at} = \frac{\kappa_a}{s+(\beta+s)\kappa_a}\bigg((1+s)\Delta\ln w_{at} + \alpha\Delta\ln A_{at}\bigg)$. The main qualitative difference when we allow households to consume more than one unit of housing, the amount of housing consumed rises when incomes rise. Consequently, even when moving costs are infinite, so no household changes locations, increases in local incomes will still cause an increase in local rents. This contrasts with the result when households only consume one-unit of housing, in which case, when moving costs are infinite, changes in income

the difference in rents between location a and location b:

$$\widetilde{\ln r_t} = \frac{\kappa_a(\widetilde{\ln w_t} + \alpha\widetilde{\ln A_t}) + s\kappa_a(\ln N - \ln 2) + s\xi}{s + \beta\kappa_a}.$$

Now, suppose that location a is situated over a promising shale deposit and chooses to allow fracing. Those activities increase demand for labor but also may increase the negative amenities discussed in the previous section. The introduction of fracing within a jurisdiction has an ambiguous effect on land values, while it is likely to lead to higher wages. Land values are affected by two forces. First, land prices will increase as oil and gas firms compete for access to land to drill, and the local labor force will presumably expand, increasing demand for housing. Second, any increase in negative amenities will reduce households' demand for housing and land. Which of these forces dominates is ultimately an empirical question.

In particular, suppose the introduction of fracing increases household income by some amount Δw_{at}[5] and changes amenities by some value ΔA_{at}. The change in relative rents is then given by:

$$\Delta\widetilde{\ln r_t} = \frac{\kappa_a(\Delta\widetilde{\ln w_t} + \alpha\Delta\widetilde{\ln A_t})}{s + \beta\kappa_a}. \tag{5}$$

If we further suppose that fracing does not change amenities or wages in location b, then equilibrium rents will not change in location b and we can then write the change in rents in location a purely in terms of changes in household incomes and WTP for amenities in location a:

$$\begin{aligned}\Delta \ln r_{at} &= \frac{\kappa_a}{s + \beta\kappa_a}(\Delta \ln w_{at} + \alpha\Delta \ln A_{at}),\\ &= \phi(\Delta \ln w_{at} + \alpha\Delta \ln A_{at}),\end{aligned} \tag{6}$$

where $\phi \equiv \frac{\kappa_a}{s+\beta\kappa_a}$.

The parameter ϕ captures how percentage changes in household income or willingness-to-pay for amenities translate into *percentage changes* in rents. ϕ varies between a minimum of 0 and $\frac{1}{\beta}$. In the canonical Roback model where there are zero moving costs and ho-

do not translate into changes in rents.

[5]In the case where the production function in place a is given by $F_{at} = \alpha_{at}N_{at}$. Suppose α_{at} increases by $\Delta\alpha_{at}$. If firms operate in competitive markets, then wages are $w_{at} = \alpha_{at}$, and because the production function is CRS there are no profits and $\Delta \ln w_{at} = \Delta w_{at} = \Delta \ln \alpha_{at}$, i.e. the percentage change in productivity. In the more complicated case when there are decreasing returns to scale (DRS) in production or firms do not pay workers their marginal product, because of search frictions, for example, then wages may rise less than the change in productivity. In addition, in practice firms make profits, of which local households receive a share through lease payments, and consequently the change in household income will include both the change in wages and lease payments. However, these lease payments will not depend on an individual's post-fracing choice of residence or work, and so would not enter the workers' location decision. However, these lease payments do impact the welfare and, consequently, we include them separately in our welfare calculations.

mogeneous tastes for locations such that $s = 0$, $\phi = \frac{1}{\beta}$ because changes in amenities and productivities are fully capitalized into rents, leaving real wages unchanged. A large value of ϕ implies that increases in household income or amenities will cause large increases in rents; this is the case when moving costs and idiosyncratic place-based preferences are small (i.e. s is small) or housing supply is very inelastic (i.e. κ_a is large). Conversely, ϕ will be low when moving costs and idiosyncratic place-based preferences are high, or housing supply is very elastic.

This model allows for calculations that are of tremendous practical value for inferring the local welfare consequences of fracing. In the subsequent empirical analysis, we will estimate the effect of fracing on housing prices[6], household income, and population $\widehat{\Delta \ln r_t}$, $\widehat{\Delta \ln y_t}$, and $\widehat{\Delta \ln N_t}$ respectively. Using these estimates, and values of s and β calibrated from Albouy (2008), Diamond (2016), and Suarez Serrato and Zidar (2016), it is possible to derive an implementable expression for the willingness-to-pay for the change in amenities in location a. Specifically, we can differentiate Equation 4 and re-arrange, yielding:[7]

$$\alpha \Delta \ln A_{at} = \quad s\Delta \ln N_{at} - \Delta \ln w_{at} + \beta \Delta \ln r_{at} \tag{7}$$

This is a remarkably useful expression because it provides an estimate of willingness-to-pay for the full set of amenity changes, even though a data set with the complete vector of amenities or information on willingness-to-pay for these amenities are unlikely to ever be available.

Additionally, it is possible to develop an expression for the change in welfare for all the people that either reside or own property in location a before the change in amenities and local productivity occurred.[8] To do so, we'll also need to incorporate the effects of fracing on dividend and interest income, which includes lease payment from oil and gas firms. Let $\widehat{\Delta y_a^{\text{owner}}}$ indicate the estimated change in interest and dividend income for home-owners. Now, let $\bar{W}_a$ be average baseline household wage and salary income, $\bar{Y_a}^{\text{owner}}$ be the average baseline interest and dividend income for homeowners and $\bar{R}_a$ be average baseline rent, then the welfare change in dollars for an individual renter is $\bar{W}_a(\widehat{\Delta \ln w_{at}} + \alpha \widehat{\Delta \ln A_{at}} - \beta \widehat{\Delta \ln r_{at}})$, and the welfare change for a landowner (who may or may not reside in location a) who owns one housing unit is $\bar{R}_a \times \widehat{\Delta \ln r_{at}} + \bar{Y}_a^{\text{owner}} \times \Delta \widehat{\ln y_{at}^{\text{owner}}}$. Thus, the expression for the total change in welfare for all individuals that either reside or own property in location a

[6] If fracing shifted rents in a place permanently, competitive housing markets would imply that the percentage change in rents and housing prices should be the same. However, the shift in rents may not be permanent because owning a home can entail homeowners to lease payments that renters do not receive, and renter and owner-occupied homes may not be perfect substitutes; for these reasons, the percentage change in rents and owner-occupied homes are likely to differ.

[7] We thank an anonymous referee for suggesting we use this expression, along with calibrated values of s and β, to estimate the change in amenities, which simplified our previous approach.

[8] This calculation ignores the change in welfare for in-migrants, as well as any profits received by oil and gas firms in excess of lease payments to local residents. It also assumes that the the average change in household income is attained by original residents, and is not due to high earnings by inmigrants. Finally, the expression omits profits of landowners who develop new housing units or rent previously vacant housing units.

before the change in amenities is:

$$\begin{aligned}
\Delta \widehat{V_{at}} \approx \quad & N_{at}\bar{W}_a \times \left(\widehat{\Delta \ln w_{at}} + \alpha \widehat{\Delta \ln A_{at}} - \beta \widehat{\Delta \ln r_{at}}\right) + N_{at} \times \left(\bar{R}_a \widehat{\Delta \ln r_{at}}\right) \\
& + N_{at}^{\text{owner}} \times \bar{Y}_a^{\text{owner}} \times \Delta \widehat{\ln y_{at}^{\text{owner}}} \\
= \quad & N_{at}\bar{W}_a \times \left(\widehat{\Delta \ln w_{at}} + \alpha \widehat{\Delta \ln A_{at}} - \beta \widehat{\Delta \ln r_{at}}\right) + N_{at} \times \left(\bar{W}_a \beta \widehat{\Delta \ln r_{at}}\right) \\
& + N_{at} \times \bar{Y}_a \times \widehat{\Delta \ln y_{at}} \\
= \quad & N_{at} \times \left(\bar{W}_a \widehat{\Delta \ln w_{at}} + \bar{Y}_a \times \widehat{\Delta \ln y_{at}} + \bar{W}_a \alpha \widehat{\Delta \ln A_{at}}\right) \qquad (8)
\end{aligned}$$

Therefore the total change in local welfare is equal to total population in place a, times the change in income per household (including both the change in wage and interest and dividend income per household) and the change in the WTP for amenities per household. The change in rents has dropped out, because renters' loss (gain) from the increase (decrease) in rents is exactly counterbalanced by the gain (loss) for property owners from the same increase (decrease) in rents.[9]

II. Fracing Background Appendix

In this section, we describe in more detail the history of fracing and provide more background information on the economic and technological features of fracing.

THE DEVELOPMENT OF A NEW DRILLING TECHNIQUE. — In this section, we briefly describe the history of the development of the technologies associated with hydraulic fracturing. In doing so, we draw on and summarize material from a number of more comprehensive histories of fracing. For more detailed information, the reader should consult the following sources. Montgomery and Smith (2010) outline the broad history of hydraulic fracturing. Gold (2014) and Zuckerman (2013) provide histories of the development and consequences of modern fracing in shale formations. Wang and Krupnick (2013) discuss the key factors leading to the development of modern fracing in shale formations. Trembath et al. (2012) discuss the role government policy played in the development of modern fracing techniques for shale formations.

Oil and gas producing firms have long used different methods of fracturing rock to stimulate production, starting in the early days of the US oil industry in Pennsylvania when drillers used torpedoes filled with black powder to increase well production. These methods became less important with the discovery of extremely productive fields in Texas and other parts of the American west in the late 19th and early 20th centuries (Gold (2014)).

[9]It is perhaps most straightforward to see this point in the case where all homes are owner occupied.

However, as production from these fields declined, oil and gas producing firms began exploring new techniques for improving well productivity in conventional reservoirs. In 1947, Stanolind Oil performed what is considered the first modern hydraulic fracturing job on the well Klepper No. 1 in the Hugoton Field in Grant County, Kansas, pumping napalm into the well under high pressure to fracture the rock (Gold (2014) and Montgomery and Smith (2010)). Two years later, in 1949, Halliburton completed the first two commercial wells to be hydraulically fractured (Montgomery and Smith (2010)). Fracing techniques and knowledge improved as fracing became a common practice.

These new fracing techniques were, for the most part, used to target conventional reservoir rocks such as limestone or sandstone. Geologists had long known that there were extensive hydrocarbon deposits contained in shale formations. However, shale wells did not produce enough hydrocarbons to be profitably exploited. Shale formations are much less permeable than conventional reservoir rocks, meaning that much of the oil or gas is unable to flow through the rock to the wellhead with conventional drilling. For decades, many in the oil and gas industry doubted whether shale formations would ever serve as anything more than a minor source of hydrocarbons.

In the early 1980s, George Mitchell, co-founder of Mitchell Energy, began experimenting with different fracing techniques in the Barnett Shale in Texas. Mitchell Energy had been producing from the Boonsville above the Barnett for decades, but production from those conventional wells had been declining, and Mitchell was searching for a new gas resource (Martineau (2007)). Mitchell experimented with different ways of fracing wells in the Barnett for many years with mixed success. However, in 1998 Mitchell began experimenting with fracs using much more water and less sand (rather than the gel fracs Mitchell Energy had been using previously). This experimentation paid off on June 11, 1998, when the S.H. Griffen Well No. 4 began producing gas at a much larger rate than previous Barnett wells (Gold (2014)). This new technique—called a "slick-water" frac because the frac fluid was much thinner than the gel-based frac fluids used previously—not only led to more productive wells, but also cost less than half of the specially designed gels (Martineau (2007)). Soon, Mitchell Energy was completing many Barnett wells using their water-based technique.

Initially, other firms were skeptical of reports regarding Mitchell's Barnett wells; the conventional wisdom was that wells drilled into shale formations could not be consistently great performers (Gold (2014)). However, as Mitchell drilled more wells, firms observed their surprising productivity levels. In late 2001, Devon Energy agreed to purchase Mitchell Energy for $3.1 billion (Zuckerman (2013)). Devon had experience drilling horizontal wells, and combined their horizontal wells with Mitchell's fracing techniques to great effect (Gold (2014)).[10] This combination of exploiting shale formations using horizontal wells with massive, slick-water frac completions is what is now often colloquially called fracing.

Once the success of these techniques in the Barnett became clear, firms began trying to use them in shale formations throughout the US. Indeed, Figure 1 shows that the share

[10] Mitchell Energy had experimented with horizontal wells in the 1980s and 1990s, but was never able to produce enough from them to justify their much higher cost (Wang and Krupnick (2013)).

of hydrocarbons produced from horizontal wells over shale formations has grown from less than 1 percent of US energy production to around 25 percent since fracing's discovery.

The application of Mitchell's innovation to other shale deposits has not always been straightforward. Although the broad approach is similar, the types and number of stages, along with the appropriate chemicals and proppants to use, vary significantly between different shale formations. Furthermore, surrounding formations significantly affect the appropriate fracing techniques, so detailed knowledge of the local geology is important. Additionally, because oil molecules are larger than gas molecules, some industry members initially thought that fracing techniques would not work as well for producing oil (Zuckerman (2013)). Finally, certain shale formations are more amenable to modern fracing techniques than others. Consequently, modern fracing techniques have slowly spread across the different shale formations in the United States.[11] The paper's empirical approach rests on the spatial and temporal variation in the diffusion of fracing across the United States and is described in more detail below.

III. Local Impacts of Hydraulic Fracturing Activity Appendix

In this section, we qualitatively describe some of the potential negative local impacts, many of which cannot be directly measured.

The most frequently discussed environmental concern has probably been the potential for water contamination by chemicals involved in the fracing process, hydrocarbons, such as methane, or"formation water" - water contained within shale or other rock formations that is produced during the drilling process - which may include potentially hazardous salts, minerals, and other materials (Environmental Protection Agency, Office of Research and Development (2015)). There are a number of potential sources of water contamination. First, some of the fractures produced during the fracing process could allow hydrocarbons or drilling fluids to leak up into the water table. The conventional wisdom is that this is unlikely to be a pervasive problem, because fraced wells are drilled extremely deeply—often more than a mile below the water table.[12] Second, improper casing of wells could also allow hydrocarbons or drilling fluids to leak into the groundwater. A series of studies by Duke University researchers has found evidence that, at least in some cases, methane has leaked into groundwater (Osborn et al. (2011)). Third, a mixture of frac fluid and formation water is produced during the drilling process. This wastewater - sometimes called "flowback" - must be recycled or disposed of and there are concerns that during this disposal process the water could contaminate drinking supplies or local bodies of water (Ground Water Protection Council and ALL Consulting (2009), Environmental Protection Agency, Office of Research and Development (2015)). Wastewater could contaminate groundwater or local water supply if there are accidents or spills while the wastewater is being stored or

[11] Indeed, there is even significant heterogeneity within a shale basin in the amenability to modern fracing techniques and in the optimal techniques. For example, Range Resources, the first firm to successfully use modern fracing techniques in the Marcellus, wrote in a report that "...the Marcellus is not created equal" Zagorski, Wrightstone and Bowman (2012).

[12] `http://pubs.usgs.gov/sir/2008/5059/section4.html`.

transported prior to injection.[13] When wastewater is not disposed of in injection wells, it is often sent to local water treatment facilities where it is processed and then released into local water supplies or bodies of water, potentially posing environmental or health risks if the treatment does not successfully remove hazardous materials from the water (Environmental Protection Agency, Office of Research and Development (2015)).[14]

In addition to water contamination, there is evidence that activities associated with fracing can lead to earthquakes, including in some regions that traditionally have been seismically stable (Connelly, Barer and Skorobogatov (2015)). The consensus within the geophysics community is that fracing itself rarely, if ever, causes large earthquakes, but that the disposal of frac fluids and other wastewater[15] into deep injection wells can cause earthquakes.[16] USGS researchers confirmed the possibility of injection well-induced earthquakes in experiments in Rangley, CO, in 1969 (Ellsworth (2013)). More recently, evidence from both case studies of particular areas (Frohlich (2012), Kim (2013)) and national studies (Weingarten et al. (2015)) suggests that modern injection wells are causing earthquakes, and may explain the rise in the frequency of earthquakes above magnitude 3.0 in the Central and Eastern US in recent years (Weingarten et al. (2015)). However, the risk of injection well-induced earthquakes depends on local seismic characteristics. For example, recent research suggests that the use of disposal wells in the Bakken has not caused a large increase in seismic activity (Frohlich, Walter and Gale (2015)). Additionally, it is currently unknown whether injection wells will cause earthquakes large enough to cause significant damage.[17] Because injection wells rather than production wells are the fracing activity associated with earthquakes, fracing-related earthquakes will not always occur in the same areas as actual drilling. Indeed, the injection wells that caused earthquakes near Youngstown, OH were predominantly used to dispose of wastewater from Marcellus shale wells drilled in Pennsylvania (Kim (2013)).

Fracing activities can also lead to elevated levels of local air and noise pollution, and increases in visual dis-amenities. The typical production well requires the delivery of between 550 and 1,400 truckloads of water.[18] Modern drilling pads can be quite large, the typical size being 2.5 acres, causing an eyesore and creating significant amounts of noise during the drilling and completion process (National Energy Technology Labratory (2013)). Further, drilling and associated activities at the well sites often rely on diesel

[13] For example, wastewater is sometimes stored in open impoundment ponds prior to disposal, and there are concerns that these impoundment ponds could leak (Phillips (2014)). Additionally, as in the case of wells drilled for production, disposal wells that were improperly sealed could leak waste into groundwater

[14] Ground Water Protection Council and ALL Consulting (2009) discuss commonly used wastewater disposal methods in different shale plays.

[15] It's important to note that conventional drilling also uses injection wells to dispose of drilling fluid and wastewater. Indeed, much of the wastewater disposed of in Oklahoma, the site of many recent earthquakes, comes from wastewater from conventional wells (Rubinstein and Mahani (2015)).

[16] See Ellsworth (2013) and Rubinstein and Mahani (2015) for discussions of the science of and evidence for injection well-induced earthquakes.

[17] There is active research on how injection wells affect the chances of very large earthquakes (Ellsworth et al. (2015)).

[18] Fracturing a well requires between 2 and 5 million gallons of water (http://www.usgs.gov/faq/categories/10132/3824%20), while a tank truck holds 3,600 gallons of water (http://www.truckinginfo.com/article/story/2012/11/trucking-fracking-water.aspx).

generators which can emit high levels of air pollution. In addition, some wells, primarily oil ones, flare natural gas giving an area an industrial feel that many judge to be unpleasant for residential purposes.

There are also substantial concerns about quality of life issues resulting from fracing activity. Anecdotally, fracing brings in large numbers of young men with weak connections to the local community. There have been concerns that this has led to elevated crime rates and inhospitable social environments for some groups, especially young women.[19]

Local manufacturing industries may also be hurt by increases in local prices caused by fracing. If there are agglomeration economies or other market failures, these higher prices could cause a "Dutch Disease," whereby the growth of the natural resource sector hurts the long-term prospects of the region. However, in the case of fracing, local natural gas and electricity prices often decline, which can mitigate some of these effects. Fetzer (2014) finds evidence consistent with the idea that reduced electricity prices may have reduced "Dutch Disease" in areas affected by fracing. Relatedly, fracing activity is responsive to oil and gas prices. Indeed, declines in natural gas prices (Krauss (2009)) and, more recently, oil prices (Takersley (2015)) have significantly reduced fracing activities in some shale plays. This sensitivity may lead to frequent local boom and bust cycles in areas with fracing activity that may affect the local economy and quality of life.[20]

Finally, there is some evidence that the national and global impacts of fracing influence local communities' preferences about allowing fracing locally. The increased supply of natural gas due to fracing has helped to reduce local air pollution and carbon emissions due to a shift in electricity production from coal power plants to relatively cleaner natural gas turbines. Others have argued that over the long-run fracing will reduce usage of renewable energy sources, raising local air pollution and carbon emissions. These impacts on local and global emissions have shaped opinions about fracing. The geopolitical consequence of fracing may also shape opinions about the advisability of allowing it in local communities.

IV. Research Design

In this section, we present an expanded version of Section II.B and describe the variation exploited by our empirical strategy in greater detail.

A. *Cross-Sectional Variation in Prospectivity within Shale Plays*

Shale plays are not homogenous and there is significant variation in the potential productivity of different locations within a shale play. Geological features of the shale formation affect the total quantity and type of hydrocarbons contained within a shale formation, the amenability of the shale to fracing techniques, and the costs of drilling and completing the well. Among others, these features include the depth and thickness of the shale formation,

[19] See, for example, `http://www.npr.org/2013/01/14/169363299/five-years-into-fracking-boom-one-pa-town-at-a-turning-point`.

[20] For example, see recent articles about the recent slowdown in activity in the Bakken (Oldham (2015)).

as well as the thermal maturity, porosity, permeability, clay content, and total organic content of the local shale rock (Zagorski, Wrightstone and Bowman (2012), Budzik (2013)). The thickness, porosity, and total organic content of the shale determine the quantity of hydrocarbons that could have formed in the shale formation. Thermal maturity, which measures how much heat the shale has been exposed to over time, determines whether hydrocarbons have formed and, if so, what types. Finally, the permeability, clay content, presence of natural fractures and depth influence how well the formation will respond to fracing, as well as how expensive drilling and completion will be.[21,22]

The relationship between these factors and amenability to fracing is often non-monotonic. For example, deeper formations are under higher pressure, which can make them easier to frac. However, drilling deeper wells is also more expensive. Additionally, the type of surrounding rock layers and the presence of natural fractures also influence the effectiveness of fracing. For example, initially fracing in the Barnett was less effective in areas where the Barnett was immediately above the Ellenberger salt-water reservoir, because the frac fluid would dissipate into the water reservoir rather than creating the desired fractures in the shale formation (Martineau (2007)).[23] Furthermore, the relationship between different geological features and productivity may differ across different shale plays.

Rystad Energy is an oil and gas consulting firm that provides research, consulting services, and data to clients worldwide. We purchased Rystad's NASMaps product that includes GIS shapefiles of Rystad's Prospectivity estimates for each North American shale play (Rystad Energy (2014)). Figure 2 maps the Rystad Prospectivity estimates for major US shale plays. The "prospectivity"values are estimates of the potential productivity of different portions of shale plays based on a non-linear function of the different geological inputs, including formation depth, thickness, thermal maturity, porosity, and other information, along with Rystad's knowledge and expertise on the impact of geology on productivity in different shale plays. The geological variables included and the functional forms used to transform them into prospectivity scores differ for each shale, so scores cannot be compared across shale formations.

We aggregated the Rystad prospectivity measure to the county level by computing the maximum and mean Rystad score within each county. We then divide counties, within a shale play, into Rystad score quartiles. Our preferred measure of fracing exposure is based on the maximum prospectivity score within each county. This decision is motivated by the observation that the quality of a county's best resources may more strongly impact hydrocarbon production than the average quality. We also explore the sensitivity of the results to alternative measures of fracing exposure. Figure 3 shows a map of the county assignments.

[21]Depth is also correlated with thermal maturity, because deeper formations have usually experienced higher levels of pressure and heat.

[22]See Budzik (2013) for a general discussion of the role played by different geological characteristics in determining the effectiveness of fracking. Zagorski, Wrightstone and Bowman (2012) describes the geological features of the Marcellus and their role in drilling productivity, Covert (2015) includes a discussion of the importance of different geological factors in the Bakken. See McCarthy et al. (2011). for an introduction to the science of hydrocarbon formation and a helpful discussion of thermal maturity.

[23]Firms eventually found that this obstacle could be surmounted by drilling horizontal wells.

Appendix Figures 1 through 3 illustrate how the Rystad prospectivity scores are used to structure the research design for the Bakken play, which is part of the Willston Basin. Figure 1 depicts the Williston Basin in green and the Bakken play in blue. Appendix Figure 2 adds the Rystad prospectivity data for the Bakken Play. Darker red indicates a higher prospectivity score.

Appendix Figure 3 reveals how the detailed shape files are summarized to develop our measure of fracing opportunities. Counties with land area that is in the top quartile of the Rystad prospectivity measure for a given shale play are coded in bright red. Counties without such land area that intersect the Bakken shale play are colored light blue. The identification strategy is based on comparing the red counties to the blue counties, within a play. Finally, counties that only intersect the Williston Basin are shown in green.

B. Temporal and Cross-sectional Variation in the Discovery of Successful Fracing Techniques

While geological features of the shale deposits provide cross-sectional variation, the paper's research design also exploits temporal variation in the initiation of fracing across shale plays. This time variation comes both from heterogeneity in the shale formations' geology and potential for oil and gas recovery that led to differences in the time elapsed before drilling and exploration firms devised successful fracing techniques in each play, as well as local and national economic factors influencing oil and gas development. We determined the first date that the fracing potential of each of the 14 shale plays in the US became public knowledge. When possible, these dates correspond to investor calls and production announcements when firms first began drilling operations involving fracing in an area or released information on their wells' productivity.

Table 1 summarizes the temporal variation in the initiation of fracing across shale plays, as well as the distribution of top-quartile counties within each play. The Barnett was the first play where modern hydraulic fracturing in shale plays combined with horizontal wells found success. This success started becoming public in late 2000 and early 2001. Fracing was initiated in 10 of the 14 plays by the end of 2009. In total, there are 95 top-quartile counties and 310 counties outside of the top quartile in these 14 plays.

As an example of how we determined the first date when the fracing potential of a play became public, below we outline the history of the Marcellus play. See Silver (2011) and Carter et al. (2011) for a more complete history of fracing in the Marcellus. The history below draws predominantly from these two sources. Range Resources, an independent oil and gas producer, had acquired leases in Washington County and other counties in the area of southwest Pennsylvania near Pittsburgh. They drilled a number of wells targeting non-Marcellus formations. However, in the early 2000s, Range's vice president of technology, Bill Zagorski, learned of the phenomenal production Mitchell Energy was achieving using fracing in the Barnett shale. Subsequently, Zagorski convinced his colleagues to try fracing the Marcellus. In October 2004, Range re-completed the Renz No. 1 well with a Barnett-style slickwater frac. Range then tried combining fracing with drilling horizontal wells. Their efforts paid off when they completed the well Gulla No. 9, which produced gas at impressive rates (Silver (2011)). Range announced their success fracing horizontal wells in the Marcellus in a press release on December 9, 2007 (Range Resources (2007)). Around the same time, academic geologists Terry Engelder and Gary Lash estimated that the Marcellus contained much more natural gas than had been previously thought. Their findings were publicly announced by a Penn State press release on January 17, 2008 (Engelder and Lash (2008)). Combined, the Range Resources announcement and findings of Engelder and Lash helped spur increased interest and development of the Marcellus. Appendix Section VI.J provides brief outlines of the history of fracing and our first frac date assignment for each of the shale plays in our sample.

There are a couple concerns raised by the use of play-specific dates for when the potential of fracing in the area became public knowledge. First, although there is an element of serendipity in when exactly the potential of different shale plays was discovered, these discoveries were the result of experimentation by oil and gas producers within the shale play. Consequently, if economic factors influence which areas oil and gas producers experiment in, then the timing of the development and publicization of different shale plays may be related to local economic trends. Second, although the development of successful local fraced wells is a large shock to information about the potential of a given shale play, previous developments of other shale plays likely affected beliefs about the potential of a given shale play as well. For example, the development of the Barnett signaled that profitable fracing of shale formations was possible, while the development of the Bakken proved that not only gas, but also oil could be profitably be produced from shale formations using fracing. Each of these events likely caused some increase in expectations that other shale plays could be successfully developed. This second concern is especially important for our investigation of the effect of fracing on housing prices, which, in a competitive housing market, would rapidly incorporate each additional piece of information as it became public.

We will address these concerns in two ways. First, we will carefully explore the possibility of pre-trends in hydrocarbon production and local labor market outcomes. Second, we will estimate the effect of fracing on housing prices by comparing the growth in housing prices between 2000 and 2010.[24] As Table 1 demonstrated, the date where fracing was initiated in all shale plays studied in this paper falls between these two years, so it is reasonable to presume that local housing markets have absorbed much of the available information about the prospects for fracing in their respective communities in each of these years.[25]

C. Propensity Score Matching

We explored an alternative identification strategy based on propensity-score matching that allows us to estimate the impact of fracing on non-top quartile counties and leverages control group outside of the play. Specifically we generate control groups for each shale play using the following procedure. First, we restrict the donor pool to counties that are not in a shale play and are neither neighbors nor neighbors-of-neighbors of top-quartile

[24] In practice, because the long-form of the decennial Census was replaced by the ACS, we use the pooled 2009-2013 ACS rather than data just from 2010

[25] Clearly, there is uncertainty about the total amount of fracing activity that is actually realized.

counties. Second, we follow Imbens and Rubin (2015) and select covariates and second-order terms to include in the propensity score using an iterative procedure[26], resulting in a list of 17 first-order and second-order terms.[27] Third, using these covariates selected by the Imbens (2015) procedure, we estimate the propensity score for being in a shale play. Fourth, we match each county to the five counties with the closest propensity score, with replacement.[28] Fifth, we construct a control sample for each shale play using the unique matched counties for that play.[29] This procedure provides a control group for the entire shale play, allowing us to estimate both the effect of fracing on non-top quartile counties and on top-quartile counties purged of potential spillovers.[30] When estimating the effect of fracing on top quartile or bottom-three quartile counties we use the same set of pscore-matched counties as the control group for each shale play. Consequently, except in specifications involving covariates or weights, the estimated effect on top-quartile counties using the p-score matching procedure will be the sum of the estimated effect on counties in the bottom three quartiles and the estimated effect of fracing on top quartile counties relative to counties in the bottom three quartiles.

Propensity Score Matching Results. — Appendix Tables 14 through 17 report estimates of Equation 8 using the propensity-score matching strategy described above. Although the results in Table 2 suggest that the pscore-matching procedure does not form the basis of a

[26]This procedure works as follows. You start with a large set of potential covariates for including in the propensity score. Starting with a model with just an intercept, in each stage you iteratively add variables one at a time to the logit propensity score model, estimate the model, and note the resulting improvement in the likelihood ratio relative to not including the given variable. After doing this for all potential covariates, you keep the covariate that resulted in the largest increase in the likelihood ratio. Then you start again with the remaining covariates, continuing until the improvement in the likelihood ratio of the estimated propensity score falls below 2.71. Then you follow the same procedure for all of the second-order terms involving the first-order terms chosen in the first step of the procedure. In this second-step, we use a stricter threshold that second-order terms are added until the improvement in the likelihood ratio drops below 10

[27]These covariates are log median home values in 1990,change in log-median home values between 1990 and 2000, change in log salary and wage income between 1990 and 2000, share black in 1990, change in agricultural/mining employment share from 1980 to 1990, log in-migration in 1993, share employed in agriculture/mining in 1990, share under 18 in 1990, change in total employment between 1990 and 2000, change in median home values between 1980 and 1990, change in share older than 65 between 1980 and 1990, total value of natural gas production in 1992, value of petroleum production in 1992, share employed in manufacturing in 1990, share college educated in 1990, share employed in mining in 1990 squared, change in log median home values between 1990 and 2000 interacted with share employed in mining in 1990, log median home values in 1990 squared, log median home values in 1990 to oil and gas production in 1992

[28]We restrict the donor pool to counties whose propensity score is at least .001.

[29]We only include each control county once for each shale play, although a control county can be in the sample more than once as a control for a different shale play.

[30]A final subtlety is that this procedure changes the number of counties in the control sample for each shale play, which implicitly changes in the estimand. To see this, note that it can be shown that, when shale play fixed effects are included, our long-difference specs can be written as a weighted average of the effect of fracking in each shale play, i.e. $\hat{\beta} = \sum_{p=1}^{P} \phi_p \hat{\delta}^p$, where δ^p is the effect of fracing in shale play p and $\phi_p = \frac{\mu_p \bar{Z}_p(1-\bar{Z}_p)}{\sum_{p=1}^{P} \mu_p(\bar{Z}_p)(1-\bar{Z}_p)}$ is the sample-share weighted variance of the top-quartile indicator within shale play p. Using the p-score matched sample as the control group changes ϕ_p, implicitly changing the estimand. Consequently, to keep the results comparable we reweight observations when using the propensity score matching procedure by $\frac{\phi_p}{\phi_p^{pscore}}$ where ϕ_p is the value when the control group is non-top quartile counties within shale play p and ϕ_p^{pscore} is the value of ϕ when the control group is the pscore matched sample.

credible empirical design, we report the estimates and discuss them below for completeness.

Starting with Appendix Table 14, which resports the propensity-score matching estimates of the effect of fracing on hydrocarbon production, we see that the matching estimator yields similar results to our main specifications for top-quartile counties. Three years after initiation of fracing hydrocarbon production is estimated to be $460 million higher. We also find a modest increase in hydrocarbon production in the other three quartiles although the effect is nearly 1/10 the size in the top quartile.

Turning to fracing's economic impacts, Appendix Table 15 provides evidence that our main estimates may be lower bounds on the impact of fracing on employment and income. Column (1) suggests that employment in top-quartile counties increases by 10.4 percent, roughly twice the estimates in Table 4. We find similar results for income and migration; the estimates in column (1) tend to be nearly twice those in Table 5. Column (2) provides a potential explanation for this difference; there may be economic spillovers or economies of scope that increase employment in nearby counties. The estimates suggest that employment in non-top quartile counties increases by over 5 percent. We also find statistically significant increases in income and migration. Given the comparison of levels and pre-trends, we acknowledge that the matching identification strategy is not as clean as our geology-based strategy, but these results suggest we may want to view the geology-based estimates as lower bounds.

In, Appendix Table 16, the matching estimator in column (1) again suggests that our main estimates in Table 5 may be best viewed as a lower bound on the change in employment and income in top-quartile counties as these estimates are significantly larger than those in Table 5. The results with respect to the employment-to-population ration, unemployment, and population are more mixed. We find statistically significant increase in income and reductions in unemployment for non-top quartile counties in column (2).

Turning to the corresponding results from the pscore-matching estimator in Appendix Figure 11, we see that as above the pscore estimates suggest that we may be understating the effects of fracing by using nearby counties as the control group. This understatement appears particularly large for construction, leisure and hospitality, financial activities, and professional and business services, while the estimates for the impacts on oil and gas employment are fairly similar. This pattern makes sense given that oil and gas employment may need to take place not far from actual drilling activities, while non-tradeables built by construction or supplied by leisure and hospitality firms can be consumed in nearby counties as well.

The estimates of the effect of fracing on crime in Table 6 were imprecise. The matching estimates suggest one potential reason for this imprecision: the possible presence of spillovers to nearby counties. In Appendix Table 17 we find a positive and statistically significant increase in the total and property crime rate for non-top quartile counties and a large, but imprecisely estimated, increase in the crime rate for top-quartile counties. The estimated impact of fracing on violent crime is imprecisely estimated.

The matching estimator for top quartile counties reported in Table 18 in general tell a similar story to our main estimates in Table 7 for total expenditures and revenues, although there are some differences when we look at specific components. For example, we find larger capital outlays in column (1), compared to Table 7, driven by an increase in capital outlays in non-top quartile counties (column (2)). The increase in public safety expenditures is also modestly larger in column (1), again driven by an increase in expenditures in quartiles 1 through 3. On the revenue side, we find a smaller increase in "other tax revenue" and charges and sales tax revenue in column (1) and a larger increase in property tax revenues compared to Table 7.

Finally, turning to housing price estimates using the matching estimator in Table Appendix 19, we find little evidence of increases in housing prices using the matching estimator, if anything, there is evidence that housing prices fell in non-top quartile counties. This may not be too surprising given the comparison of levels and trends in Table 2. We find statistically significant pre-trends in housing prices and housing units in top-quartile counties compared to their matched counterparts, and statistically significantly different housing price and units levels for the non-top quartile counties. As such, we may not expect to find reliable estimates from this identification strategy.

V. Robustness Checks and Extensions

In this section, we perform several robustness checks and extensions of the main analysis to help us better understand our results and probe their sensitivity to alternative assumptions. Specifically, we perform two main analyses. First, we investigate how our estimates of the effect of fracing on amenities and welfare change using alternative measures of changes in prices or quantities or alternative parameter assumptions regarding idiosyncratic location preferences or the share of income spent on housing. Second, we present Two-Stage Least Squares (2SLS) where we use our Rystad prospectivity measures as an instrument for oil and gas production, producing estimates of the effects of a dollar of oil and gas produced on local economic outcomes.

A. Welfare and Amenity Calculations Under Alternative Assumptions

In Table 10, we reported empirical estimates of the annual WTP for the change in amenities and annual WTP for allowing fracing using Equation 11 and Equation 6 and the estimates of the effect of fracing on local housing costs, population, wage and salary income, and income from rents and dividends in Tables 5 and 8 under particular assumptions about the the variance in idiosyncratic location preferences and the share of income spent on housing that we took from the literature. We also reported standard errors of these welfare estimates taking into account both sampling variation and uncertainty in the assumed parameters.

These standard errors, however, do not tell us how the results would change under particular alternative assumptions or how they would change under alternative measures of changes in housing costs, wage and slaary income, population, or income from rents and dividents. In Appendix Table 20, we fill this gap and explore how our welfare calculations change when we use alternative measures of changes in prices or quantities or different parameter values. Panel A, Column (1) and (2) duplicates our base estimates of the amenity and welfare impacts of fracing in dollars per household per year. Recall that for these calculations, we use the change in local housing costs, which me measure using the change in local housing rents (2.9 percent), the change in local population, which we measure using the change in the total local population (2.7%), the change in wages and salaries, which we measure using the change in wage and salary income per capita (7.5 percent), and the change in rent and dividend income, which we measure using the change in average income from rents and dividends multiplied by one over the share of homeowners (13.1 percent). Combining these measures with an assumed standard deviation of idiosyncratic location preferences (s) of .30 and share of income spent on housing (β) of .65, we find that amenities decline by $-\$2,225$ per household per year, but that the increase in local earnings is large enough to offset this amenity decline, resulting in an overall increase in welfare per household per year of $\$1,716$.

In Columns (3)-(8) we investigate how these estimates change using alternative ways of measuring the change in housing costs (columns (3) and (4), population (columns (5) and (6), and wages and salaries respectively (columns (7) and (8)). Columns (3) and (4) report estimates using the change in self-reported values of owner-occupied housing units rather than rents. Although rents directly correspond to the theoretical object in the model, changes in the value of owner occupied housing units may be better measured as the majority of households own their home.[31] The larger increase in the value of owner-occupied housing compared to rental prices (5.7% compared to 2.9%) reduces the amenity decline by about a third (-$1,405 versus -$2,225) and increases the welfare rise by about 50% ($2,537 versus $1,716). Columns (5) and (6) use alternative measures of the change in population. Labor market migration is most relevant to those at prime working ages of about 18 to 55. Those younger than 18 will usually be in school, while many of those older than 55 are less likely to be willing to pay the costs of migrating or retired. Consequently, the change in the 18-55 population may more accurately reflect how people respond to the combined change in labor market opportunities or amenities. Unsurprisingly, the estimated change in the 18-55 population (5.3%) is larger than the change in the overall population (2.7%). This results in a smaller estimated decline in amenities (-$1,855), resulting in a larger estimated local welfare effect ($2,086). The estimated decline in amenities is smaller because a smaller change in amenities is required to rationalize the larger population response to the same change in real wages.

In Columns (7) and (8), we use an alternative measure of the change in wage and salary income. The measure used in our base models, wage and salary income per capita, includes both changes in hourly wages, but also changes in hours worked per year. The theoretical model presented in Section I abstracts from these issues - workers inelastically supply labor and their is no involuntary unemployment. However, in other models, hourly wages or wages for full-time, full-year workers may be the correct measure rather than wages

[31] Although it should be noted that it's not totally clear that changes in the value of owner occupied housing are better measured. Self-reported changes in the value of owner-occupied homes may include the value of royalty payments, leading to double counting when we include these income sources. Furthermore, homeowners may be overly optimistic about how much their homes have appreciated in value since fracing began.

and salary income per capita.[32] In Columns (7) and (8), we present welfare estimates where the change in wage and salary income is measured based on the change in wage and salary income for full-time, full-year workers (4.1%), which is somewhat smaller than the change in wage and salary income per capita (7.5%). This smaller estimated change in real wages requires a smaller decline in amenities (-$560) to rationalize the observed change in population, but also yields smaller welfare gains ($1,529).

In Panel B, we then show how the estimates change under an alternative empirical specification to those used in Tables 5 and 8 where we also allow for differential pre-trends in top-quartile compared to counties outside the top quartile. Allowing for pre-trends leaves the results qualitatively unchanged: we estimate economically large declines in local amenities and increases in local welfare in all cases. Quantitatively, the effects of allowing for pre-trends depends on the measure of changes in prices or quantities. Using our base measures of changes in quantities in Columns (1) and (2), allowing for pre-trends increases the absolute value of the decline in amenities, reducing the positive welfare effects. A similar effect happens in Columns (5) and (6) when we measure the change in population using the change in the 18-55 population and in Columns (7) and (8), when we measure the change in earnings using the change in earnings for full-time, full-year workers. Conversely, in Columns (3) and (4), when we measure the change in housing costs using the change in self-reported owner-occupied home-values, we find that allowing for pre-trends reduces the estimated decline in amenities, increasing the estimated welfare improvement.

Finally, in Panel C, we then show how the estimates change under alternative assumptions about s and β. Row B1 reports results assuming a lower standard deviation in idiosyncratic location preferences ($s = .1$, compared to .30 in our base model). With a lower standard deviation of idiosyncratic preferences, a larger population response would be expected for a given change in real wages, so the estimated decline in amenities is correspondingly larger (-$2,466) and the corresponding change in welfare is consequently smaller ($1,475). In Row B2, we report results assuming a higher standard deviation of idiosyncratic preferences ($s = .5$), which generates the opposite effect from Row B1, a smaller estimated decline in amenities (-$1,980) and larger increase in welfare ($1,961). Finally, Row B3, reports results using a smaller share of income spent on housing (β). Our base model uses $\beta = .65$, which includes the effect of housing prices on other local prices, and allows for wage and salary income to be only a portion of tot al income. A lower β increases the estimated change in real wages, which then requires a larger change in amenities (-$2,403) to rationalize the small observed change in population, resulting in lower overall welfare increases ($1,538).

Overall, these results are quite consistent, with every specification suggesting a decline in amenities of at least -$350 per household per year and an increase in local welfare of at least $1,300 per household per year. This consistency shows that our conclusions are not driven by subjective measurement or parameter choices, but instead reflect broad empirical

[32]Note that this is not true of all models that allow for workers to work different numbers of hours. For example, in models of involuntary unemployment, a tightening labor market may lead to more hours or weeks worked per year. Consequently, in such models our preferred specifications including per-capita income unconditional on hours worked may be appropriate.

patterns that yield similar conclusions under quite different modeling assumptions.

B. Effects of Fracing per Dollars of Hydrocarbons Produced per Capita: Two-Stage Least Squares Results

Instrumental variables estimates of the effects of hydrocarbon production on local economic outcomes are shown in Table 21. The estimates are from long-differenced regressions of the change in county aggregate economic outcomes between 2000 and 2009/2013 on measures of oil and gas production per household, using the balanced sample of counties. The OLS estimates in column (1) suggest that an additional $10,000 worth of production per household would be associated with small increases in home values, rental income, wages and salaries, rental income. The IV estimates are much larger, suggesting that the OLS is biased downwards by omitted variables that are correlated with both larger hydrocarbon output, and lower values for these economic indicators. The IV estimates suggest a $10,000 increase in oil and gas production per household causes a 2.4 percent increase in home values, a 1.3 percent increase in rental values, a 3.1 percent increase in wage income, and a 3.8 percent increase in rent and dividend income.

VI. Data Appendix

In this section, we describe the data sources and construction in more detail than in the main text.

A. Geological Information

Shapefiles of the locations of shale plays and basins come from the Energy Information Agency (EIA) May 9, 2011 map of "Shale Gas and Oil Plays in the Lower 48 States" (Energy Information Agency (2011)). The play and basin shapefiles were the spatially joined using GIS software with shapefiles of US counties from the Census TigerLine files (US Census Bureau (2010)), allowing us to determine which shale plays and basins intersect each county. For each county, we also compute the fraction of the area of the county covered by each shale play that the county intersects. We then compute indicators for intersecting the given shale play at all, having 50 percent of the county covered by the shale play, having 75 percent of the county covered by the shale play, and having 99 percent of the county covered by the shale play.

Prospectivity estimates come from the NASMaps product purchased from Rystad Energy, which we discussed in more detail above. Just as we did with the basin and play shapefiles, we merge the prospectivity estimates with county shapefiles. For each county, we compute the maximum and average.[33] Prospectivity values for each shale play the county intersects. We then compute indicators for being in the top quartile and top octile

[33] If the county is not completely covered by the prospectivity shapefile, we compute the average prospectivity value coding the missing portions of the county as having a "0" prospectivity score.

of the prospectivity score for each shale play. Frequently, the extent of the prospectivity valuation shapefiles for a given shale play does not exactly correspond with the boundaries of the shale play itself. In these cases, we compute the quartiles and octiles of the prospectivity distribution based only on the counties that overlap both the prospectivity shapefile and the EIA shale play shapefile.[34,35]

B. Hydrocarbon Production

Oil and gas production data for 1992 through 2011 come from data purchased from Drilling Info, Inc (2012) and used with permission.These data were purchased from HPDI before they because Drilling Info, Inc. These data contain data on oil and gas production by well for most wells, with the data starting in different years for different states. 1992 is the first year in which all states that intersect shale plays in our sample have non-missing data.[36,37] Wells are identified by a unique American Petroleum Institute (API) number and include the latitude and longitude of the well, as well as information on the county and state. County information is sometimes missing or inaccurate. Consequently, we instead determine the county ourselves by spatially joining the well latitudes and longitudes with the US county shapefile discussed above using GIS software.[38] We then aggregate oil and gas production to the county year level, creating a data set with total oil and gas production in each year for all US counties.

The oil and gas data come in number of barrels for oil and thousands of cubic feet for gas. We convert both of these units to quantities that are comparable. First, We compute the value of oil and gas production in each year using the EIA price time-series. The EIA oil price is the average Cushing, OK WTI spot price while the gas price is the average National Citygate prices. We also compute the energy content of the produced oil and gas in terms of BTUs and Joules.[39] Finally, using EIA data on the heat rates of different types of power plants, we also convert oil and gas production into GWh.[40]

In an effort to confirm the accuracy of our oil and gas production data, we aggregated the data to the state-year level and compared our figures with EIA data for total production by state. We found that our state year aggregates matched the EIA data very closely for state-years in our sample.

[34]There are also cases where the extent of the prospectivity shapefile is smaller than the extent of the shale play shapefile. In these cases we compute the prospectivity distribution based only on counties within the shale play with non-missing prospectivity information

[35]Because of our data use agreement, we are only able to include county-level aggregates of the Rystad energy prospectivity information in our replication files. Fortunately, these data are sufficient to replicate all of our results.

[36]Data for most states begins well before 1992. However, Maryland data only begins in 1992 and Pennsylvania only mandated production reporting starting in 1991, so Pennsylvania data is unrepresentative before 1991.

[37]Note that because of our data use agreement, the replication files posted along with this paper only include county-year level aggregates.

[38]We drop wells where the latitude and longitude differs across years and are greater than 1 km from one another. We also drop well-year combinations where there are multiple, conflicting production reports in the same year.

[39]Using the fact that 1 tcf of natural gas roughly equals $1.03 * 10^6$ BTU, 1 barrels of oil equals $5.6 * 10^6$ BTU, and that 1 BTU roughly equals 1061.48 Joules

[40]The EIA average heat rate value in 2011 was 10829 BTUs per KWh for petroleum power plants and 8152 BTUs per KWh for natural gas power plants.

C. Housing

Housing price, quantity, and characteristics data crime from three datasets produced by the US Census bureau: the American Community Survey (ACS), the Decennial Census, and the New Residential Construction series.

1990 and 2000 housing variables come from the Decennial Census. For later periods, we use the pooled 2009-2013 American Community Survey, which provides data for all counties (individual ACS years supress information on a number of variables for many counties). Consequently, for median and mean[41] home values for owner-occupied units, median and mean gross rent for renter-occupied units, mobile home variables, and data on housing characteristics we use data from the 1990 and 2000 Decennial Census as well as the pooled 2009-2013 ACS. We use gross rent, which includes utility costs regardless of whether utilities are included in the contract rent or not, because this will reduce measured rent differences resulting from differing practices regarding the inclusion of utilities in contract rent or not. The housing characteristics we use in the analysis are: the share of housing units with the following characteristics 0 bedrooms, 1 bedroom, 2 bedrooms, 3 bedrooms, 4 bedrooms, 5 or more bedrooms, full indoor plumbing, a complete kitchen, mobile, gas utility heating, taking heating, electricity heating, fuel heating, coal heating, wood heating, solar heating, other heating, no heating, built in the last year, built between 2 and ten years ago, built between ten and twenty years ago, built between twenty and thirty years ago, built between thirty and forty years ago, built between forty and fifty years ago, built more than 50 years ago. We compute these shares separately for renter-occupied and owner-occupied units.

The usage of Census and ACS housing data is slightly complicated by changes in the universe of owner and renter-occupied price data over time. In particular, prior to 2000 for owner-occupied values and 2005 for renter-occupied gross rents, the Census only collected data from specified housing units. Specified owner-occupied units include only single family houses on less than ten acres of land, which excludes "mobile homes, houses with a business or medical office, houses on 10 or more acres [...], and housing units in multi-unit buildings" (US Census Bureau (2013), page 35). The definition of specified renter owner-occupied is less restrictive, and only excludes "1-family houses on 10 or more acres" (US Census Bureau (2013), page 36). This change in the universe of housing variables causes two complications. First, for gross rents, we must use gross rents for specified renter-occupied units in 2000 and all renter-occupied units in 2009/13. Second, for owner occupied housing units, we can only test pre-trends in Table 2 for specified owner-occupied housing units, whereas our main housing results Table 8 uses all owner-occupied units. This means that, in theory, there could be pre-trends for the full universe of owner-occupied housing units even though there are no pre-trends for specified owner-occupied housing units.

We investigate the extent to which these two complications should be concerning in Appendix Table 10 using individual Public Use Microdata (Ruggles et al. (2015)), which

[41] Mean home values and rental prices are computed by dividing the aggregate value (aggregate rent) of owner-occupied (renter occupied) units and dividing by the total number of owner occupied (renter occupied) units

allows us to directly compare differences in levels and changes of owner-occupied values and renter-occupied gross rents for specified housing units vs. all housing units.[42] In Column (1) the table shows the percentage difference in the number of specified housing units compard to the full universe of housing units. Column (2) reports the percentage difference in the change in the given housing outcome between 2000 and 2007/11 computed using the universe of specified housing units in 2000 and the full universe in 2007/11 compared to using the full universe of housing units for both years.[43] Similarly, Column (3) reports the difference in the percentage change in the given housing outcome between 2000 and 2007/11 when computed using the universe of specified housing units for both years compared to using the full universe for both years.[44]

Appendix Table 10 provides a number of re-assurances. First, Column (1) shows that, as we might have expected given the small number of excluded units, there is only a small difference in the average number of units when using the specified versus the full universe. Consequently, the change from specified to the full universe between 2000 and 2009/13 is unlikely to meaningfully affect our results. This is borne out in Columns (2) in Panels A2 and A3, which shows that the average change in median or mean gross rent of renter occupied housing units differs by less than one-quarter of one percent when computed using specified units in the base year or the full universe in the base year. Turning to owner occupied units in Panel B, Column (1) shows that non-specified units make up a much large fraction of total housing units - more than twenty percent. Consequently, we may be more concerned that there may be different trends in home prices for specified units compared to the full housing universe, which would make our tests for differential pre-trends between top quartile and other housing units in Table 2 less meaningful. However, Column (3), Panels B2 and B3, assuage these concerns. Specifically, the average difference in trends between specified housing units and the full universe is less than two percent, suggesting that differential pre-trends between non-specified housing units are unlikely to influence our finding in Table 2 that top quartile counties have similar trends to other counties within shale plays.

We also use data on the number of housing permits from 1990 to 2013 from the Census Bureau's New Residential Construction data-series (US Census Bureau (2014)). We use the reported number of permits and aggregate the number of 1 family, 2 families, 3 or 4 families, and 5 or more families to a single measure of the number of permits. We do this for both the number of building permits and the number of unit permits.

[42]The smallest geographic unit reported in the public use microdata is the Public Use Microdata Area (PUMA), which are collections of counties. Consequently, this entire analysis is performed at the PUMA level.

[43]Starting with the 2012 ACS, the Census Bureau began using the 2010 PUMAs, which differ in some cases from the 2000 PUMAs used for the 2000 Census and 2005-2011 ACS'. Consequently, we must use the 2007/11 ACS' to maintain geographic comparability.

[44]Ideally, we would have reported the differences using changes between 1990 and 2000 in Column (3). However, because the PUMA boundaries changed between 1990 and 2000, the data do not have consistent geographic definitions.

D. Agricultural Land

We compute the number of acres of agricultural land and the fraction of land devoted to agriculture using data from the Census of Agriculture (National Agricultural Statistics Service (NASS)). These contain data on agriculture by county at 5 year intervals: 1997, 2002, 2007, and 2012. We compute the fraction of land devoted to agriculture by dividing the total agricultural land by the total land variable.

E. Income and Employment

Data on total employment, and personal income by type come from the Local Area Personal Income (LAPI) data, which is from the Regional Economic and Information Systems (REIS) data produced by the Bureau of Economic Analysis (BEA) (US Bureau of Economic Analysis (BEA)). All of the employment and income variables are measured at the county-by-year level. Specifically, we use the data from the series CA04 for personal income and CA25 for employment. For personal income we use the variable CA04-10, for total employment we use CA25-10, and for total population we use CA04-20. We also compute a total wage and salary income by place of residence variable, including supplemental income, which is the sum of CA04-50, CA04-42 (adjustment for place of residence), and CA04-61 (employer contributions for employee pension and insurance funds). We compute per-capita versions of these variables by dividing by population (CA04-20).

Data on wages by industry by place of work come from the Quarterly Census of Employment and Wages (QCEW) produced by Bureau of Labor Statistics, US Department of Labor (2014). We use the data from the county high-level QCEW NAICS-based files from 1990 to 2013. We compute wages per worker by dividing by total employment (within the given industry, when appropriate).

F. Migration

We derive migration counts and rates for 1990-2012 from the Internal Revenue Service (IRS) U.S. Population Migration Data (Internal Revenue Service (2015)), which is a part of the Statistics on Income (SOI). The IRS migration data comes in the form of all county-to-county migration flows. The data are computed based on addresses listed by income tax-filers. We aggregate the data by summing all in-migration and out-migration for each county to create a total in-migration and out-migration for each county. We also create a county-level population measure by summing the number of people originally in the county (i.e. population for county i in year t is equal to the sum of non-migrants and migrants from county i in $t+1$). We can then compute the in and out migration rates as $\text{in(out)-migration rate}_{it} = \text{in(out)-migration}_{it}/\text{pop}_{i,t-1}$. We compute net-migration as $\text{in-migration}_{it} - \text{out-migration}_{it}$.

G. County Demographics

We draw county demographic and economic characteristic data from the Decennial Census and the ACS. Data for 1990 and 2000 come from the Decennial Census. For the later time period, we draw from pooled the 2009-2013 ACS. Consequently, we use the 1990 and 2000 decennial censuses and the pooled 2009-2013 ACS for percent of the county that is urban, employment-to-population ratio, percent of people with less than a high school diploma, fraction with a college degree or more, percent employed in manufacturing, and percent employed in mining.

The census sex by age variables come in one-year age groups. We aggregate these to broader age groups. Specifically, the "prime" age population is anyone 18 to 64.

H. Crime

Crime data come from the FBIs Uniform Crime Reporting (UCR) program (Federal Bureau of Investigation (2015)). Individual law enforcement agencies (e.g., City of Cambridge Police, MIT Police, etc...) report "index crimes" to the FBI, including murder, rape, aggravated assault, robbery, burglary, larceny, and motor-vehicle theft, as well as the population that the law enforcement agency covers.[45] Reporting is non-mandatory (although some federal grants are conditioned on reporting UCR data, so there is an incentive to report), and consequently not all agencies report crimes in all years or months. Consequently, we must do a fair amount of cleaning to generate consistent, annual data and then aggregate these data to the county level. We proceed in first steps. First, we clean the annual data for each law enforcement agency. Second, we define a consistent sample of policy agencies that consistently report crime data and then aggregate these data to the county level. We now discuss each of these steps in turn.

If an agency reports between 3 months and 12 months of crime, we scale all of their reported crimes up by the proportion of months reported.[46] If an agency reports less than 3 months of crime, we code all of their crime types as missing for the year. In practice, most agencies either report more than 11 months or 0 months of crime. We also recode agencies' crime to missing if they report 0 for total crime or an individual crime type in a given year, but positive crimes for the given crime type in other years and, based on their population and reported crimes in other years, they would be expected to have at least 20 crimes of the given crime type.[47] We follow a similar rule for population, and recode an agencies population and crime data as missing if the agency reports zero population in the

[45] The UCR program data also include simple assault. However, simple assault is not an index crime, and therefore reporting is much more inconsistent than for other crimes. Consequently, we do not use simple assault in any or our cleaning procedures or analysis.

[46] For example, if an agency reports 4 months of crime, we scale up all of their crime variables by 3 to generate an estimate for crime over the full year. This amounts to assuming that the crime rate is the same in reported months and non-reported months.

[47] We estimate expected crime for a given crime type in agencies that report 0 crime for the given crime type using the average reported crime for the crime type for the given agency in years when it reports positive crime for the given crime type. We follow this procedure because we want to allow for agencies that, given their small population or low crime rates, may actually have zero reported crimes, while also not including zeroes that are clearly misreported.

given year but positive population in other years and the average population reported in positive years is more than twenty. To avoid throwing out data from agencies that report an index crime (population) in all years except for one or two, we interpolate each index crime (population) for an agency in year t if the agency reports the given index crime in year $t+1$ and $t-1$ and the agency is missing the given index crime (population) data in no more than three years of crime data from 1990 to 2013.

Finally, some agencies are not contained within one county, but rather overlap multiple counties. Typically, the vast majority of the population of the agency is in one of these counties, but in some cases the population covered by the agency is more evenly split. Unfortunately, we only know the identity of the county in which the agency covers the largest population. However, we do know the population the agency covers for the three counties in which the agency covers the largest population. Consequently, to reflect the fact that for these multi-county agencies the reported crime did not all occur within the county with the largest covered population, we downweighted reported crime proportionally to the share that the county with the largest covered population had of the total population in the three counties with the largest covered population for the given agency. This procedure results in an agency-year dataset of agencies that consistently report crime information.

To avoid within-county sample composition changes over time from influencing our results, we then define a consistent sample for each county of agencies that report crimes in most years. We define the consistent sample as agencies for which we have either a reported or an interpolated crime value for all[48] index crimes in every year from 1992[49] to 2013.[50] To ensure that the consistent sample agencies are representative of the county as a whole, we only include counties in our sample if the consistent sample agencies represent at least 20 percent of total county crimes between 2011 and 2013. We assign agencies to counties using the crosswalk from United States Department of Justice, Office of Justice Programs, Bureau of Justice Statistics (2012). A few counties do not have any agencies that report crimes for at least three months in every year, and consequently our sample size is smaller for crime than our other outcome variables, containing 56 Rystad top-quartile counties and 340 total counties, compared to 65 top-quartile counties and 405 total counties in the full sample. Following the FBI, we sometimes group crimes into the categories of violent crimes and property crimes. Violent crimes include murder, rape, aggravated assault, and robbery, while property crimes include burglary, larceny, and motor-vehicle theft.

I. *Inflation*

All dollar denominated outcomes are inflation adjusted to constant 2010 USD using the Consumer Price Index (CPI) produced by the BLS. We use the annual averages of the All

[48]Note that there are in fact a number of cases where agencies consistently report some index crimes, but not others. Our procedure excludes these agencies from our sample.

[49]We start in 1992 because in the Bakken shale play a number of agencies don't start reporting until 1991 or 1992

[50]If an agency is covered by another agency in some year - i.e. agency b reports crimes for agency a, we include agency a in the consistent sample if in every year either i) agency a has non-missing crime data, or ii) agency b has non-missing crime data and is covering agency a in the given year.

Urban Consumers (CPI-U) price index for all items.

J. First-Frac Dates

As discussed above, we determined different "first-frac dates" from the history of the oil and gas industry in each shale play. We undertook this exercise for all shale gas plays that were listed as being actively developed in the National Energy Technology Laboratory's Modern Shale Gas Development in the United States: An Update (National Energy Technology Labratory (2013)), as well as for the major shale oil plays.[51] We determined first-frac dates using two different definitions of the start of fracing. The first definition we determined is the first date that a well using modern hydraulic fracturing techniques was completed and showed promising production in the given shale formation. The second definition we determined is when the first major publicity occurred regarding the success and potential of modern fracing technologies in the play (this often came from press releases, annual reports, or investor calls from firms regarding the first particularly successful fraced wells). In many plays there are a number of years between these two dates. In Table 1 below, we list the first-frac dates using both definitions for all of the shale plays in the U.S. that are being actively developed. In our analysis, we use the first publicity date, because we think this more closely corresponds to when information about fracing would be reflected in housing prices and when fracing related activities may increase.

Choosing the first-completion and first-publicity dates for fracing can be quite subjective, and there is substantial uncertainty regarding the correct dates. For some shale plays, such as the Marcellus, both the first modern-well date and the first-publicity date are relatively straightforward and agreed upon. However, for other plays, such as the Bakken, the correct dates are less straightforward. The first difficulty is that the date when the potential of a play became public knowledge can be difficult to determine. Operators are often quite secretive in releasing promising results, because they want to lease land before other firms know about the area's potential. Consequently, operators some time release only bits of information—mentioning that a horizontal well has been drilled into a certain shale formation—without disclosing how positive the results have been. Additionally, although for some cases it is quite clear what the first "successful" modern, fraced well is (such as the Eagle Ford), in others there are a series of moderately successful wells that firms continually improve upon, making it more difficult to pinpoint then the first fraced well that was particularly productive was completeld. Finally, it is often the case that people doubt whether success in a particular area can be generalized to other areas in a formation. For example, few thought that the success of Arco Energy and Continental Energy fracing wells in the Elm Coullee field in Montana in the early 2000s could be replicated in other parts of the Bakken (Zuckerman (2013)). Consequently, even after positive results in one location it is not always the case that these results lead to widespread changes in people's perceptions of the potential of a given shale play. Often it took success in several different

[51] Including the Bakken, Eagle Ford, Niobrara DJ-Basin and Niobrara Powder-River Basin. In practice, because all of these plays also produce some natural gas, they were also included in the National Energy Technology Labratory (2013) document as well.

locations in the shale play or success in parts of the shale formation previously not thought strong candidates for production.

Some first-frac dates can be very late in the year. For example, the first publicity date for the Marcellus shale play is December 9, 2007. Because hydrocarbon production, economic and other data correspond to the entire year, it would be incorrect to code 2007 as a year with fracing for the Marcellus. To account for this issue, for the analysis we use the year of the first-frac date if the month is June or before, and the year of the first-frac date plus one if the month is July or later.

Below, we discuss our first-frac assignments and sources for each shale play in more detail. We highlight for which plays we are confident in our dates, and for which plays there is more uncertainty.

Bakken. — The Bakken shale is part of the Williston basin and lies under much of North Dakota, as well as parts of Montana and Saskatchewan, Canada. The Bakken is actually composed of three layers, the upper and lower shale layers, and the middle Bakken layer which is composed of more permeable, non-shale layers.[52]

Olesen (2010) and Nordeng (2010) provide more detailed histories of oil production from the Bakken, while Gold (2014) and Zuckerman (2013) include extensive discussions of the more recent history of hydraulic fracturing in the Bakken. Here we summarize the points from these sources that are most relevant to determining a first-frac date, primarily drawing on these four sources. The Bakken was first exploited in 1953. Later, developers found success in the Bicentennial and Elkhorn fields, first with vertical wells and later, after 1987, with horizontal wells. However, producers had mixed success, and new drilling eventually tapered out (Nordeng (2010) and Olesen (2010)). Around 2000, Lyco energy started developing the Elm Coullee field in eastern Montana. Using horizontal wells drilled into the middle Bakken and single-stage frac-jobs, Lyco Energy found significant success (Nordeng (2010) and Williams (2004*a*)).[53] The success of the Elm Coullee field, however, did not attract much attention and here were doubts whether the success at the Elm Coullee field would apply generally to the Bakken (Zuckerman (2013) pp. 204, 250).

People started to take notice when EOG Resources began releasing results from wells drilled near Parshall, in Mountrail County, North Dakota. EOG complete the 1-36 and 2-36 Parshall wells, drilled in May and September 2006, and both wells were good producers (Nordeng (2010)). EGO was initially circumspect regarding the success of their Parshall wells—the first public mention of their Bakken success seems to be February 1, 2007, when they reported that they had completed five horizontal wells in the Bakken and mention especially high production from the Warberg 1-25H well (EOG Resources (2007) and Zuckerman (2013)). However, it remained unclear whether the potential of the Bakken

[52] The Three-Forks formation lies under the Bakken and also responds well to horizontal drilling and fracing. Consequently, many firms in the area also target the Three-Forks formation. Because of this, the Bakken play is sometimes called the "Bakken-Three-Forks play."

[53] Lyco's first horizontal, fraced well was completed on May 26, 2000 (Williams (2004*a*)) and Montana Board of Oil and Gas Conservation (2015)).

extended throughout the entire play, or just in specific areas such as the Elm Coullee and the Parshall fields. Brigham Energy would provide this confirmation in late 2008, when it found success outside of the Parshall field at the Olsen 10-15 No. 1H well (Brigham Exploration (2009), Gold (2014), and Crooks (2015)).[54] In the meantime, the United States Geological Survey (USGS) had released an updated assessment of the technically recoverable reserves in the Bakken that increased the estimated recoverable reserves by about twenty-five times (United States Geological Survey (2008)). Coupled with this announcement and EOGs stellar results in the Parshall field, Brigham's success rapidly led many firms to enter the Bakken. We use Lyco's first horizontal well in the Elm Coullee, which was completed on May 26, 2000 (Montana Board of Oil and Gas Conservation (2015))), as the first completion date and EOGs description of their success in the Parshall field on February 1, 2007 as the first publicity date. However, an argument could be made that the appropriate date is actually September 7, 2008 (Gold (2014)), when Brigham completed the Olsen 10-15 No. 1H or April 2008, when the USGS released their new assessment of the technically recoverable reserves in the Bakken.

BARNETT. — The Barnett shale lies in the Forth Worth Basin in Northeast Texas. In the early 1980's, George Mitchell, co-founder of Mitchell Energy, began experimenting with different fracing techniques in the Barnett Shale in Texas. Mitchell Energy had been producing from the Boonsville above the Barnett for decades, but production from those conventional wells had been declining, and Mitchell was searching for a new gas resource (Martineau (2007)). Mitchell experimented for many years with mixed success. However, in 1998 Mitchell began experimenting with fracs using much more water and less sand (rather than the gel fracs Mitchell Energy had been using previously). This experimentation paid off on June 11, 1998, when the S.H. Griffen Well No. 4 began producing gas at a much larger rate than previous Barnett wells (Gold (2014)).

Initially, other firms were skeptical that Mitchell's Barnett wells were producing so well (the conventional wisdom was that wells drilled into shale formations could not consistently produce enough oil or gas to justify their costs (Gold (2014)). However, as Mitchell drilled more wells firms observed their surprising productivity levels. On August 14, 2001 Devon Energy agreed to purchase Mitchell Energy for $3.1 billion (Sidel and Cummins (2001)). Devon had experience drilling horizontal wells, and combined their horizontal wells with Mitchell's fracing techniques to great affect (Gold (2014)).[55] Choosing the first-frac date for the Barnett is difficult, because although Mitchell Energy began having success using slickwater fracs in mid 1998, many industry observers were extremely skeptical that it was possible to profitably produce for shale formations. It was only after Mitchell Energy showed increasingly impressive production that the industry started to take notice. Additionally, it was only when Devon bought Mitchell Energy and combined Mitchell's

[54] Brigham's innovation was to dramatically increase the number of frac stages compared to previous Bakken wells, using 20 stages in the case of the Olsen 10-15 No. 1H well (Stell (2009)).

[55] Mitchell Energy had experimented with horizontal wells in the 1980s and 1990s, but was never able to produce enough from them to justify their much higher cost (Wang and Krupnick (2013)).

slickwater frac techniques with horizontal wells that all of the key elements for producing from the Barnett were put together.[56] It is easy to pinpoint the date of the first successful completion of a fraced well in the Barnett as June 11, 1998. However, the first date the potential of the Barnett was widely known is more difficult because information about Mitchell's success only slowly trickled out and because the industry was originally so skeptical of producing from shale formations in general. Furthermore, the full potential of the Barnett only began to be realized when Devon started drilling horizontal wells. Consequently, one could also argue that the appropriate first publicity date is when Devon started having success with horizontal wells in late 2002. We chose the date of July 19, 2000, when an article in Oil and Gas Investor covered Mitchell's success using slickwater fracing in the Barnett (Fletcher (2000)), but recognize that arguments can be made supporting somewhat earlier or later dates.

Eagle Ford. — The Eagle Ford is a large shale formation that stretches from Northern to Southern Texas. Although the Eagle Ford was known to be a source rock for the Austin Chalk formation, which was successfully exploited over many decades using a variety of techniques, the Eagle Ford itself did not attract much attention until late 2008.[57] On October 21, 2008, Petrohawk Energy announced impressive natural gas production from its wells in La Salle County (Vaughn (2012)). Meanwhile, Pioneer Energy and Rosetta Energy had also started drilling wells in the Eagle Ford. The Petrohawk announcement combined with the success of other firms spurred a flurry of development of the Eagle Ford (Williams (2009)). Later, firms discovered that areas of the Eagle Ford farther to the northwest had substantial potential for producing oil. This discovery spurred even more investment in the play, and the Eagle Ford rapidly became one of the most important shale plays in the United States (Oil and Gas Investor (2013)). We use the completion date for Petrohawk's first Eagle Ford well, August 28, 2008 (Railroad Commission of Texas (2015)), as our the date of the first completion that was very successful of a modern, fraced well and we define the first publicity date as the October 21, 2008 Petrohawk announcement.

Fayetteville. — The Fayetteville shale formation is the eastern portion of the Arkoma basin in north-central Arkansas. In 2002, Southwestern Energy "noticed that completions in the overlying Wedington Sandstone sometimes produced greater volumes of gas than

[56] Using horizontal wells both improved production of wells throughout the Barnett, it also allowed firms to produce from areas of the Barnett that overlay the Ellenberger formation, which had previously been difficult.

[57] Although the Eagle Ford attracted little public attention before 2008 and the Petrohawk well completed August 28, 2008 is often cited as the first Eagle Ford well, there was actually development in the Eagle Ford for several years before 2008. Lewis Petroleum started drilling wells into the Eagle Ford in 2002 (Lewis Energy Group (2015)) and experienced some success years before Petrohawk's 2008 well (Toon (2011*a*)). Indeed, a case could be made for using the one of the early Lewis wells rather than the Petrohawk well as the first modern, fraced well that was successful in the Eagle Ford. However, ultimately we decided to use the Petrohawk well because even the CEO of Lewis Energy Group acknowledges that Petrohawk made technological advances, telling *Oil and Gas Investor* that " 'I give all the credit to Petrohawk for coming up with the proper frac technology, but they didn't discover the Eagle Ford. We drilled a lot of horizontal wells in the middle 2000s that no one knew about. We and EOG (Resources) were pushing this formation way before these other guys even thought about it." (Toon (2011*a*))

would be expected from its reservoir properties" (Williams (2004*b*)) and hypothesized that the Wedington sand wells were drawing gas from the surrounding Fayetteville shale formation (Williams (2004*b*)). After confirming that the Fayetteville had similarities with the Barnett shale, Southwestern began purchasing mineral rights in several areas overlying the Fayetteville (Taylor (2015)). They drilled several initial wells that showed promising production and announced their progress in a press release on August 17, 2004 (Southwestern Energy Company (2004)). In 2005, Chesapeake Energy followed suit and entered the Fayetteville and a number of other firms followed subsequently (Toal (2007)). We use the date that Southwestern's subsidiary, SEECO, completed their first well in the Fayetteville, May 13, 2004 (State of Arkansas Oil and Gas Commission (2015)) as date of the first modern, fraced well and the date of Southwestern Energy's press release, August 17, 2004, as the first publicity date.

Haynesville - Bossier. — The Haynesville - Bossier shale formation is part of the TX-LA-MS Salt basin, and is located within parts of Arkansas, Louisiana, and Texas. Although overlying formations had experienced development beforehand, the Haynesville - Bossier shale formations themselves had received little attention until 2006, when a few firms, including Chesapeake Energy and Cubic energy, began exploring whether modern hydraulic fracturing techniques would be effective in the Hayensville - Bossier shale (Durham (2008)). On November 27, 2007 Cubic Energy released a press release saying that the analysis of data from wells drilled into the Haynesville showing that the Hayensville - Bossier shared similar characteristics with the Barnett shale (Cubic Energy Inc (2007)). A few months later, on March 24, 2008, Chesapeake energy released results of initial wells drilled into the Hayensville - Bossier which showed promising production and announced plans for purchasing additional leases in the region (Chesapeake Energy Corporation (2008)). After Chesapeake and Cubic's announcements, interest in the Haynesville rapidly increased, and soon a number of companies were competing to buy leases and drill wells in the Hayensville (Durham (2008)). We use the date of the first well drilled into the Haynesville, drilled by Chesapeake and completed on March 2, 2007 (Department of Natural Resources: State of Louisiana (2015)) as the first modern well date, and Chesapeake's March 24, 2008 press release discussing its initial results as the first publicity date.

Marcellus. — The Marcellus is a large shale formation in the Appalachian Basin that underlies much of Pennsylvania, West Virginia, and New York, as well as parts of Ohio, Kentucky, Maryland, and Virginia. Pennsylvania is where the US oil and gas industries were born and has a long history of development, however the Marcellus itself had not undergone extensive development prior to the 2000s. Below, we outline how this changed in the 2000s with the advent of hydraulic fracturing. See Silver (2011) and Carter et al. (2011) for a more complete history of fracing in the Marcellus. The history below draws largely from these two sources.

Range Resources, an independent oil and gas producer, had acquired leases in Washington County and other counties in the area of southwest Pennsylvania near Pittsburgh.

They drilled a number of wells targeting non-Marcellus formations. However, in the early 2000s, Range's vice president of technology, Bill Zagorski, learned of the phenomenal production Mitchell Energy was achieving using fracing in the Barnett shale. Subsequently, Zagorski convinced his colleagues to try fracing the Marcellus. In October 2004, Range re-completed the Renz No. 1 well with a Barnett-style slickwater frac. Range then tried combining fracing with drilling horizontal wells. Their efforts paid off when they completed the well Gulla No. 9, which produced gas at impressive rates (Silver (2011)). Range announced their success fracing horizontal wells in the Marcellus in a press release on December 9, 2007 (Range Resources (2007)). Around the same time, academic geologists Terry Engelder and Gary Lash estimated that the Marcellus contained much more natural gas than had been previously thought. Their findings were publicly announced by a Penn State press release on January 17, 2008 (Engelder and Lash (2008)). Combined, the Range Resources announcement and findings of Engelder and Lash helped spur increased interest and development of the Marcellus. We use the re-completion of the Renz No. 1 well on October 20, 2004 (Harper and Kostelnik (NA)) as the date for the first well fraced using modern techniques and the December 9, 2007 Range Resources press release as the first publicity date for the Marcellus.

Niobrara: Denver-Julesburg. — The Niobrara Denver-Julesburg-Basin play is the portion of the Niobrara formation in the Denver-Julesburg Basin (often called the D-J Basin). It predominantly lies in Northeast Colorado, but also includes parts of Western Nebraska, and Southern Wyoming. Although the broader region and the D-J basin itself had a long-history of oil and gas production (Colorado Geological Survey (2011)), development with modern fracing techniques combined with horizontal wells didn't pick up speed until the late 2009, when EOG and Noble Energy began drilling exploratory wells. On April 7, 2010, EOG Resources released promising initial oil production results from its Jake 2-01H well (EOG Resources (2010)). Not long after the Jake well results, Noble Energy reported strong initial oil production from its Gemini well (Williams (2010*c*) and Colorado Oil and Gas Conservation Commission Department of Natural Resources (2015)). Combined, these two announcements spurred a significant amount of interest and coverage of the potential of the Niobrara D-J basin, with some even speculating that it may have similar potential to the Bakken.[58] We use the completion of the Jake 2-01H well, September 5, 2009 (Colorado Oil and Gas Conservation Commission Department of Natural Resources (2015)), as the completion date for the first successful fraced well in the Niobrara and the April 7, 2010 EOG releasing discussing the first publicity date.

Niobrara: Greater Green River Basin. — The Greater Green River Basin is a large basin that lies in Wyoming, Northwest Colorado, and Northern Utah. It contains several

[58] However, as time passed, firms found that the Niobrara D-J basin was much more heterogenous than the Bakken, with firms finding great success in some areas and little success in others. Consequently, although investment and production continued to rise, they did so less quickly than had been originally anticipated after the announcement of the Jake and Gemini wells (Klann (2012) and Sheehan (2013)).

sub-basins, including the Green River, Hoback, Washakie, Great Divide, and Sand Wash. The Washakie and Sand Wash basins, which lie predominantly in southern Wyoming and northern Colorado respectively, overly a part of the Niobrara shale (Finn and Johnson (2005)). Unlike other parts of the Niobrara, the Greater Green River Basin Niobrara did not immediately receive attention with the succeess of the Jake well in the Niobrara D-J basin in late 2009. Instead, the Greater Green River portion of the Niobrara would wait until 2011, when a number of firms started programs exploring the potential of the Niobrara in the Greater Green River Basin. On December 21, 2011, Entek Energy released an update on its exploratory drilling program in the Niobrara - Greater Green River, discusisng its Niobrara test wells, as well as activities and results released by other operators in the area (Entek Energy Limited (2011)). A month later, Oil and Gas Journal published an article discussing a similar release by one of Entek's Partners, Emerald Oil and Gas NL (Emerald Oil and Gas NL (2012)). We use December 21, 2011, when Entek Energy released its update on its Niobrara exploratory program in the Greater Green River. It was unclear when the first successful modern, fraced well had been drilled, so instead we just use the year 2011, when a number of companies seemed to have been experimenting with completing wells into the Niobrara in the Greater Green River Basin.

Niobrara: Powder River. — The Niobrara Powder River is the portion of the Niobrara formation lying in the Powder River basin, in Northeastern Wyoming and Southern Montana. It is difficult to date exactly when the potential of the Niobrara Powder River Basin first became public knowledge. Like the Niobrara Denver-Julesburg, there is a long history of conventional drilling in the area. We assigned the Niobrara Powder-River Basin the same first completion and publicity dates as the Niobrara: Denver-Julesburg, because both plays target the same formation, and they are often discussed in conjunction by industry publications and firms (Phish, McDermott and Waterous (2010)). For example, in 2011 Chesapeake energy issues a press release touting their joint venture with CNOOC Limited in the Niobrara in both the Powder River and Denver-Julesburg basins (Chesapeake Energy Corporation (2011*b*)).

Permian Basin Plays. — The Permian basin is a large basin in Western Texas and Eastern New Mexico that contains numerous overlapping formations that produce oil and gas that has been an important source of oil and gas production for many decades. Recently, production from the Permian had been declining, but the advent of massive slickwater fracing and horizontal drilling has reinvigorated oil and gas production from the Permian basin, both in traditional Permian targets and new targets such as the Wolfcamp shale (Dutton et al. (2005) and Gold (2014)).[59,60]

[59]Oil and Gas Investor quotes Tim Leach, CEO of Conco Resources Inc, saying "What we discovered in the Permian was that many of the zones that were bypassed historically were great producers using the new technology'" Anonymous (2011).

[60]Sutton (2015) has written a series of articles about the geology, production history, and current production activity in the Permian basin.

The Permian is made up of two main sub-basins: the Delaware sub-basin in the west and the Midland sub-basin in the east. The Wolfcamp formation underlies most of both the Delaware and Midland sub-basins, while the Avalon and Bone-spring formations lie in the Deleware sub-basin and the Spraberry and Cline formations lie in the Midland sub-basin. The overlapping nature of the many permian formations makes it very difficult to separately determine first-frac dates for different formations within the Permian, both because improvements in producing from one play makes producing from overlapping plays more attractive and because success using horizontal wells and hydraulic fracturing for one formation in the basin may raise expectations regarding the prospects using these technologies for other formations. Further complicating matters, many operators have found success using vertical wells drilled to produce through multiple formations, leading to "Wolfberry" (Wolfcamp and Spaberry), "Wolfbone" (Wolfcamp and Bone spring), and Avalon-Bone Spring plays. Finally, Rystad Energy provides prospectivity for the Permian Basin as a whole rather than for specific plays within the Permian Basin, making it impossible to implement our empirical strategy separately by shale play within the Permian Basin. Consequently, we use one first-frac date for the Permian Basin a whole, recognizing that this is a simplification and that the timing of the development of different plays using modern hydraulic fracturing techniques within the Permian Basin did differ somewhat in practice.

Williams (2006) and Williams (2008) provide a helpful history of the early development of the Permian basin using modern fracing techniques. Here we summarize some of the key events, which they discuss in more detail. In the Delaware Basin, modern fracing development began in 2003, when Perenco energy began drilling horizontal wells into the Wolfcamp, using acid completions. These wells were not tremendously successful. However, in 2004, EOG drilled horizontal wells and completed them using slickwater fracs, and immediately found success with the No. 1 Nile 22 state well. Several follow up wells confirmed that the combination of horizontal drilling with massive slickwater fracs would consistently payoff in the Wolfcamp (Williams (2006)). The first date this success became broad public appears to be May 9, 2005, when Parallel Petroleum - one of EOG's partners in the No. 1 Nile-22 state well - released an operations update discussed the promising results and plans to purchase more land in the area (Parallel Petroleum (2005)).[61]

Around the same time, in the eastern part of the Permian in the Midland sub-basin, firms also began finding success using fracing techniques to produce from both the Spraberry and Wolfcamp (the so-called Wolfberry play). In 1996 Arco Permian began experimenting with extending Spraberry wells into the Wolfcamp formation and completing the wells with frac jobs. A few years later, Henry Petroleum began trying out Wolfberry wells using slickwater frac jobs. It is not clear when the success of the Henry wells became widely public, but the news had certainly become public by November 2006, when St. Mary's made a substantial lease purchase in the area (Williams (2008)).

In the subsequent years, operators found that modern fracing and horizontal drilling techniques allowed them to dramatically increase production from a number of other for-

[61]It is not entirely clear to us the exact date that the No. 1 Nile-22 state well became public. The first official announcement we have found is the May 9, 2005 Parallel Petroleum announcement.

mations within the Permian basin, including the Avalon Shale, the Bone-Spring formation, the Cline shale, the Midland sub-basin portion of the Wolfcamp, and the Barnett Woodford shale (Reynolds (2013), Williams (2010*a*), Toon (2011*b*)). The histories of the use of fracing in the Wolfcamp in the Delaware basin and Wolfberry play in the Midland basin show that fracing techniques became prominent in both major areas of the Permian around the same: between 2003 and 2006. We use the first completion and first publicity dates for wells using modern fracing techniques in the entire Permian Basins. This means we define the first-completion date of a modern well in the Permian Basin as 1996, when Arco Petroleum started exploring the Wolfberry (Williams (2008)),[62] We assign May 9, 2005, when Parallel Petroleum released details of the success of the No. 1 Nile 1-22H, to be the first publicity date for the Permian Basin.

Utica. — The Utica is a shale formation in the Appalachian basin that extends through New York, Pennsylvania, Quebec, West Virginia, and Ohio. In most places, the Utica is a few thousand feet deeper than the Marcellus. In 2009, Chesapeake began purchasing leases in the Utica (Williams (2011)). Around the same time, on March 22, 2010, Range Resources completed one of the first modern wells in the US[63] part of the Utica (Harper (2011)).[64] A little more than a year later, on July 28, 2011, Chesapeake sparked increased interest in the Utica with a press release discussing promising results from vertical wells and comparing the Utica to the Eagle Ford because of the presence of gas, condensate, and oil windows (Chesapeake Energy Corporation (2011*b*)). Later in 2011, Chesapeake announced good results from its initial horizontal wells in the Utica, particularly the Buell well (Chesapeake Energy Corporation (2011*a*)). Around the same time, Rex Energy announced that it had also been receiving strong production from new Utica wells (Rex Energy (2011)). Combined, these announcements led to widespread interest in the Utica formation (Warlick (2012)). We use the completion of Range Resources No. 1 Lloyd well on March 22, 2010 as the completion date for the first modern well and Chesapeake's July 28, 2011 announcement as the first publicity date.

Woodford: Anadarko. — The Woodford Anadarko[65] lies in central and western Oklahoma, an area with an extensive history of conventional drilling. Devon Energy started ex-

[62] Note that it's not entirely clear which exact date to assign, because Arco was drilling many wells in Midland county at this time into the Spraberry and it's not entirely clear in which ones they first used modern fracing techniques, which ones or which ones also produced from the wolfcamp. For this reason we do not choose a particular well and instead simply use the year 1996 for the first completion date of a modern well.

[63] The portion of the Utica in Quebec was developed earlier, starting around 2004 (Marcil et al. (2012)).

[64] There were earlier wells completed in the Utica in both New York and Pennsylvania that did not produce (New York State Department of Environmental Conservation (2015) and Pennsylvania Department of Environmental Protection (2015)). The first Ohio Utica well wasn't permitted until September 2010 (Ohio Department of Natural Resources: Division of Oil and Gas Resources (2015)).

[65] We assign the three Woodford shale plays separate first-frac dates. There are several reasons for this choice. First, as detailed below, the Woodford plays are in three different basins, produce different types of hydrocarbons, with the Woodford Arkoma containing dry gas, the Woodford Anadarko producing dry gas and liquids, and the Woodford Ardmore producing oil and gas (Cardott (2013)). Second, these three parts of the Woodford had quite different histories and were developed at different times by different firms. Third, industrial publications generally treat the parts of the Woodford in separate basins as separate plays (For example, see Williams (2010*b*) or Cardott (2013)).

ploring the play in 2007, drilling its first horizontal, fraced well into the Woodford Anadarko on November 15, 2007. Subsequently, a number of other operators, including Cimarex Energy Company, also began completing horizontal, fraced wells in the Anadarko (Anonymous (2008)). By the end of 2008, both operators were realizing significant production in the play. The exact date that potential of the Woodford Anadarko became public is difficult to pin-down. Cimarex's 2007 annual report, released on February 28, 2008, mentions that they were drilling horizontal wells into the Woodford Anadarko (Cimarex Energy (2008)). A month later, the Oklahoma Geological Survey's annual "Drilling Highlights" newsletter mentions the initial Devon well, but does not include any information about its production (Boyd (2008)). In its 2008 annual report, Devon mentions the Woodford Anadarko play for the first time, saying that their horizontal drilling program had been successful and that they were expanding operations in the play (Devon Energy (2009)). We use the date of the completion of Devon's first horizontal Woodford Anadarko well, November 15, 2007 (Oklahoma Geological Survey (2015)), as the date of the completion first modern, fraced well in the play, and, February 28, 2008, when Cimarex first mentions the play in its annual report, as the first publicity date for the play, although one could also make a case for the release data of the 2008 Devon annual report, February 4, 2009.

WOODFORD: ARDMORE. — The Woodford Ardmore is the portion of the Woodford Shale lying in the Ardmore basin in south-east Oklahoma. The Woodford Ardmore shale differs from the other two Woodford plays in producing a significant amount of oil, in addition to natural gas. Firms started exploring the Woodford Ardmore in 2005, when Chesapeake Energy completed the first modern well in the formation on May 19, 2005 (Oklahoma Geological Survey (2015)) . On January 10, 2007, Bankers Petroleum announced results from their first vertical well into the Ardmore and plans to drill a horizontal well (Bankers Petroleum (2007*b*)). Later, on October 24, 2007, Bankers issued a press release detailing positive results from this first horizontal well, and plans to drill more horizontal wells (Bankers Petroleum (2007*a*)). Around the same time, Chesapeake energy also drilled a horizontal Ardmore well (Boyd (2008)). By mid 2008, the success of these wells had led a number of other firms to began developing wells in the Ardmore (Williams (2010*b*) and Anonymous (2008)). We use the date of the first modern well targeting the Woodford in the Ardmore basin, completed by Chesapeake Energy on May 19, 2005 (Oklahoma Geological Survey (2015)) as the first modern well date, and date of the release of Banker's Petroleum press release discussing their first vertical Ardmore well, on January 10, 2007, as the first publicity date for the Woodford Ardmore play.

WOODFORD: ARKOMA. — The Woodford Arkoma is the portion of the Woodford shale in the Arkoma basin in eastern Oklahoma.[66] The Woodford Arkoma is particularly attractive

Finally, Rystad Energy provides separate prospectivity information for each play, making it most straightforward to treat them as three separate plays given our empirical strategy.

[66]The Fayetteville shale, discussed above, is in the eastern part of the Arkoma basin in Arkansas

because it lies below the Caney Shale, providing firms with two shale formations that they can potentially produce from (Haines (2006)). In 2002, Newfield Exploration began purchasing leases in the Woordford Arkoma, in 2003 they drilled their first vertical well into the Woodford Arkoma shale, and in 2005 they drilled their first horizontal well into the Woodford Arkoma shale (Langford (2008)). The horizontal Woodford Arkoma well proved successful, and they released information about the results March 1, 2006 in their annual report and separate press releases (Newfield Exploration Company (2006) and Newfield Exploration (2006)). Around the same time, Devon energy had also began drilling horizontal wells into the Woodford, and soon a number of other operators had entered the play as well (Haines (2006)). We use August 15, 2004, the date of the first Newfield Woodford completion (Oklahoma Geological Survey (2015)) as the first completion date for a modern, fraced well in the shale formation, and March 1, 2006 - the date of the release of the 2005 Newfield annual report - as the first publicity date.

REFERENCES

Albouy, David. 2008. "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas." *National Bureau of Economic Research Working Paper Series*, No. 14472.

Anonymous. 2008. "Woodford shale play forms up in Oklahoma Anadarko basin." *Oil and Gas Journal*, 106(43): 38.

Anonymous. 2011. "Delaware Basin Delights." *Oil and Gas Investor*, 31(10): NA.

Bankers Petroleum. 2007*a*. *Bankers Petroleum Achieves First Natural Gas Production and Sales from the Woodford Shale in Oklahoma.* Calgary.

Bankers Petroleum. 2007*b*. *Bankers Petroleum Discovers Natural Gas in Oklahoma and Provides Update on US Exploration Program.* Calgary.

Boyd, Dan T. 2008. "Oklahoma 2007 Drilling Highlights." *Shale Shaker: The Journal of the Oklahoma Geological Society*, , (April): 173–181.

Brigham Exploration. 2009. *Brigham Exploration Completes ND Olson 10-15 #1H with Initial Production of 1,200 BO/d.*

Budzik, P. 2013. "The Application of Hydraulic Fracturing to US and World Shale Gas and Tight Oil Deposits." American Society of Mechanical Engineers Webinar.

Bureau of Labor Statistics, US Department of Labor. 2014. "Quarterly Census of Employment and Wages: County High-Level QCEW NAICS-Based Data Files from 1990-2013."

Cardott, Brian J. 2013. "Woodford Shale: from hydrocarbon source rock to reservoir." *Search and Discovery*, 50817(50817): 85pp.

Carter, Kristin M., John a. Harper, Katherine W. Schmid, and Jaime Kostelnik. 2011. "Unconventional natural gas resources in Pennsylvania: The backstory of the modern Marcellus Shale play." *Environmental Geosciences*, 18(4): 217–257.

Chesapeake Energy Corporation. 2008. *Chesapeake Announces Haynesville Shale Discovery and Seven Other New Unconventional Discoveries and Projects; Increases Capital Expenditures to Accelerate Development.* Oklahoma City.

Chesapeake Energy Corporation. 2011*a*. *Chesapeake Energy Corporation Discloses Initial Horizontal Well Drilling Results in Its Utica Shale Discovery and Announces Achievement of Corporate Production Milestones.* Oklahoma City.

Chesapeake Energy Corporation. 2011*b*. *Chesapeake Energy Corporation Reports Financial and Operational Results for the 2011 Second Quarter.* Oklahoma City.

Cimarex Energy. 2008. *Cimarex Energy 2007 Annual Report.*

Colorado Geological Survey. 2011. "Colorado's New Oil Boom - The Niobrara." *Rock Talk*, 13(1): 1–12.

Colorado Oil and Gas Conservation Commission Department of Natural Resources. 2015. *COGIS - Facility Inquiry.*

Connelly, Kelly, David Barer, and Yana Skorobogatov. 2015. *How Oil and Gas Disposal Wells Can Cause Earthquakes.*

Covert, Thomas. 2015. "Experiential and social learning in firms: the case of hydraulic fracturing in the Bakken Shale." Working Paper.

Crooks, Ed. 2015. "The US Shale Revolution." *Financial Times.*

Cubic Energy Inc. 2007. *Cubic Energy, Inc. Announces Positive Initial Analysis of Bossier/Haynesville Shale Formations Located in Johnson Branche Field, Caddo Parish, Louisiana.* Dallas.

Department of Natural Resources: State of Louisiana. 2015. *Haynesville Shale Wells.*

Devon Energy. 2009. *Devon Energy Reports 2008 Financial Results and Year-End Reserves.*

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill : 1980-2000." *American Economic Review*, 106: 1980–2000.

Drilling Info, Inc. 2012. *DI Desktop: US Oil and Gas Well Production Data.* Used with permission.

Durham, Louise. 2008. "Louisana Play a 'Company Maker'?" *American Association of Petroleum Geosciences: Explorer.*

Dutton, Shirley P., Eugene M. Kim, Ronald F. Broadhead, William D. Raatz, Caroline L. Breton, Stephen C. Ruppel, and Charles Kerans. 2005. "Play analysis and leading-edge oil-reservoir development methods in the Permian basin: Increased recovery through advanced technologies." *AAPG Bulletin*, 89(5): 553–576.

Ellsworth, William L. 2013. "Injection-Induced Earthquakes." *Science*, 341(July): 1–8.

Ellsworth, William L, Andrea L Llenos, Arthur F Mcgarr, Andrew J Michael, Justin L Rubinstein, Charles S Mueller, Mark D Petersen, and Eric Calais. 2015. "Increasing seismicity in the US midcontinent: Implications for earthquake hazard." *The Leading Edge*, 34(6): 618–626.

Emerald Oil and Gas NL. 2012. *Substantial oil production from Niobrara Shale wells near EMR Green River Basin acreage.*

Energy Information Agency. 2011. *Maps: Exploration, Resources, Reserves, and Production.*

Engelder, Terry, and Gary G. Lash. 2008. "Unconventional Natural Gas Reservoir Could Boost US Supply." Penn State University, University Park, Pennsylvania.

Entek Energy Limited. 2011. *2011 Niobrara Appraisal Program Update.*

Environmental Protection Agency, Office of Research and Development. 2015. "Assessment of the Potential Impacts of Hydraulic Fracturing for Oil and Gas on Drinking Water Resources Executive Summary." June.

EOG Resources. 2007. *EOG Resources Reports 2006 Results and Increases Dividend.*

EOG Resources. 2010. *EOG Resources Announces 2010 Analyst Conference Highlights.* Houston.

Federal Bureau of Investigation. 2015. *Uniform Crime Records (UCR).*

Fetzer, Thiemo. 2014. "Fracking Growth." CEP Working Paper No. 1278.

Finn, Thomas, and Ronald Johnson. 2005. "Niobrara Total Petroleum System in the Southwestern Wyoming Province." In *Petroleum Systems and Geologic Assessment of Oil and Gas in the Southwestern Wyoming Province, Wyoming, Colorado, and Utah.* Chapter 6. Denver, CO:United States Geological Survey.

Fletcher, Sam. 2000. *Mitchell Succeeds with Light Sand Fracing in Barnett shale.*

Frohlich, C. 2012. "Two-year survey comparing earthquake activity and injection-well locations in the Barnett Shale, Texas." *Proceedings of the National Academy of Sciences*, 109(35): 13934–13938.

Frohlich, C., J. I. Walter, and J. F. W. Gale. 2015. "Analysis of Transportable Array (USArray) Data Shows Earthquakes Are Scarce near Injection Wells in the Williston Basin, 2008-2011." *Seismological Research Letters*, 86(2A): 492–499.

Gold, Russell. 2014. *The Boom: How Fracking Ignited the American Energy Revolution and Changed the World.* New York, NY:Simon & Schuster.

Ground Water Protection Council, and ALL Consulting. 2009. "Modern Shale Gas Development in the United States: A Primer." , (April): 116.

Haines, Leslie. 2006. "The Woodford Shale." *Oil and Gas Investor*, 26(8): 79–81.

Harper, John A. 2011. *Activity and Potential of the Utica Shale in Pennsylvania.*

Harper, John, and Jaime Kostelnik. NA. "The Marcellus Shale Play in Pennsylvania: Part 3: The Modern Marcellus Shale Play." Department of Conservation and Natural Resources.

Hornbeck, Richard, and Enrico Moretti. 2015. "Who Benefits From Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality."

Imbens, Guido. 2015. "Matching Methods in Practice: Three Examples." *Journal of Human Resources*, 50(2): 373–419.

Imbens, Guido, and Donald Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* . First ed., New York, NY:Cambridge University Press.

Internal Revenue Service. 2015. *Statistics on Income Tax Stats - US Population Migration Data.* Washington, DC:IRS.

Kim, Won Young. 2013. "Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio." *Journal of Geophysical Research: Solid Earth*, 118(7): 3506–3518.

Klann, Susan. 2012. "Niobrara Challenge." *Oil and Gas Investor*, 32(3): NA.

Krauss, Clifford. 2009. *Natural Gas Prices Plummet to a Seven-Year Low.*

Langford, Sam. 2008. "The Woodford Shale: Newfield Investor Conference 2008." Newfield Investor Conference 2008.

Lewis Energy Group. 2015. *South Texas is Lewis Energy Territory.*

Marcil, Jean-Sebastien, Peter Dorrins, Jeremie Lavoie, Nabila Mechti, and Jean-Yves Lavoie. 2012. *Utica and Other Ordovician Shales: Exploration History in the Quebec Sedimentary Basins, Eastern Canada.* American Association of Petroleum Geologists: Eastern Section Meeting.

Martineau, David F. 2007. "History of the Newark East field and the Barnett Shale as a gas reservoir." *AAPG Bulletin*, 91(4): 399–403.

McCarthy, K., K. Rojas, M. Niemann, D. Palmowski, K. Peters, and A. Stankiewicz. 2011. "Basic Petroleum Geochemistry for Source Rock Evaluation." *Oilfield Review*, 23(2): 32–43.

Montana Board of Oil and Gas Conservation. 2015. *Wells Data.*

Montgomery, Carl, and Michael Smith. 2010. "Hydraulic Fracturing: A History of An Enduring Technology." *Journal of Petroleum Technology*, , (December): 26 –41.

National Agricultural Statistics Service (NASS), US Department of Agriculture. 2014. *Census of Agriculture.*

National Energy Technology Labratory. 2013. "Modern Shale Gas Development in the United States : An Update." 1–79.

Newfield Exploration. 2006. *Newfield Provides Update on Woodford Shale Drilling Program.* Houston.

Newfield Exploration Company. 2006. "Newfield Exploration Company: 2005 Annual Report."

New York State Department of Environmental Conservation. 2015. *New York's Oil and Gas Database.*

Nordeng, Stephen. 2010. "A Brief History of Oil Production from the Bakken Formation in the Williston Basin." *Geo News: North Dakota, Department of Mineral Resources Newsletter.*

Ohio Department of Natural Resources: Division of Oil and Gas Resources. 2015. *Utica/Point Pleasant Shale: Cumulative Permitting Activity.*

Oil and Gas Investor. 2013. "Eagle Ford Evolution." *Oil and Gas Investor*, 33(11).

Oklahoma Geological Survey. 2015. *Oklahoma Gas Shales Well Database.*

Oldham, Jennifer. 2015. *The Real Estate Crisis in North Dakota's Man Camps.*

Olesen, Neil. 2010. *Bakken Oil Resource Play, Williston Basin (US): Overview and Historical Perspective.*

Osborn, Stephen G, Avner Vengosh, Nathaniel R Warner, and Robert B Jackson. 2011. "Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing." *Proceedings of the National Academy of Sciences of the United States of America*, 108(20): 8172–8176.

Parallel Petroleum. 2005. *Parallel Petroleum Announces Operations Update.* Midland, TX.

Pennsylvania Department of Environmental Protection. 2015. *Oil and Gas Wells Formation Report.*

Phillips, Susan. 2014. *Pennsylvania Frack Ponds Now Number More than 500.*

Phish, Tim, Thomas McDermott, and Scotia Waterous. 2010. "Niobrara Oil: Regional Spotlight." *Oil and Gas Investor*, 30(8): NA.

Railroad Commission of Texas. 2015. *Oil and Gas Data Query.*

Range Resources. 2007. *Range Expands Barnett Shale Holdings and Provides Operations Update.*

Rex Energy. 2011. *Rex Energy Exceeds Third Quarter Production Guidance and Discloses Initial Horizontal Utica Shale Well Results.* State College, PA.

Reynolds, Cole. 2013. "The Cline Shale: Another Texas Giant?" *Oil and Gas Investor*, 33(9): NA.

Rubinstein, Justin L., and Alireza Babaie Mahani. 2015. "Myths and Facts on Wastewater Injection, Hydraulic Fracturing, Enhanced Oil Recovery, and Induced Seismicity." *Seismological Research Letters*, 86(4): 1060–1067.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. "Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]." Minneapolis.

Rystad Energy. 2014. *North American Shale Plays Maps: Q3, 2014.* Used with permission:http://www.rystadenergy.com/.

Sheehan, Chris. 2013. "The Niobrara Extends its Reach." *Oil and Gas Investor*, 33(5): NA.

Sidel, Robin, and Chip Cummins. 2001. *Devon Energy to Acquire Mitchell in $3.1 Billion Cash-and-Stock Deal.*

Silver, Jonathan. 2011. *Marcellus Boom, Origins: The story of a Professor, a Gas Driller, and Wall Street.*

Southwestern Energy Company. 2004. *Southwestern Energy Announces Fayetteville Shale Play in Arkoma Basin.*

State of Arkansas Oil and Gas Commission. 2015. *Wells Completed: Fayetteville Shale Gas Information (B-43 Field).*

Stell, Jeannie. 2009. "Bakken Breakout." *Oil and Gas Investor*, 29(10): NA.

Suarez Serrato, Juan Carlos, and Owen Zidar. 2016. "Who Benefits From State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review*, 106(9): 2582–2624.

Sutton, Leslie. 2015. *Permian Basin Production - Midland vs. Delaware Basins.*

Takersley, Jim. 2015. *After plunge in oil prices, hope fades for group of long-beleaguered workers.*

Taylor, Larry. 2015. *Fayetteville Shale.*

Toal, Brian A. 2007. “Fayetteville and Woodford Shales.” *Oil and Gas Investor*, 27(6): 38–48.

Toon, Steve. 2011*a*. “Eagle Ford Rising.” *Oil and Gas Investor*, 31(2).

Toon, Steve. 2011*b*. “Pioneer, EOG Lead Foray Into Emerging Horizontal Wolfcamp.” *Oil and Gas Investor*, 31(12): NA.

Trembath, Alex, Jesse Jenkins, Ted Nordhaus, and Michael Schellenberger. 2012. “Where the Shale Gas Revolution Came From: Government’s Role in the Development of Hydraulic Fracturing in Shale.” May.

United States Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. 2012. *Law Enforcement Agency Identifiers Crosswalk. Identification No. 35158.* Inter-University Consortium for Political and Social Research.

United States Geological Survey. 2008. *3 to 4.3 Billion Barrels of Technically Recoverable Oil Assessed in North Dakota and Montana’s Bakken Formation - 25 Times More than 1995 Estimate.* Reston, VA.

US Bureau of Economic Analysis (BEA). 2014. *Regional Economic Information System: Local Area Personal Income and Employment.* Washington, DC:BEA.

US Census Bureau. 2010. *2010 TIGER/Line Shapefiles (machine-readable data files).*

US Census Bureau. 2013. “American Community Survey Puerto Rico Community Survey 2013 Subject Definitions.” US Census Bureau.

US Census Bureau. 2014. *Building Permits Survey.*

Vaughn, Vicki. 2012. *Shale play built on expertise, risk, secrecy - and a simple handshake.*

Wang, Zhongmin, and Alan Krupnick. 2013. “A Retrospective Review of Shale as Development in the United States: What Led to the Boom?”

Warlick, Don. 2012. “Utica Shale Development: How do Early-Stage Characteristics Equate to Other North American Shale Plays.” *Oil and Gas Financial Journal.*

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

Williams, Peggy. 2004*a*. “The Bakken is Back.” *Oil and Gas Investor*, 24(4): 71–73.

Williams, Peggy. 2004*b*. “The Fayetteville Shale Debut.” *Oil and Gas Investor*, 24(12): 71.

Williams, Peggy. 2006. “The Wolfcamp.” *Oil and Gas Investor*, 26(1): 109.

Williams, Peggy. 2008. “The Wolfberry.” *Oil and Gas Investor*, 28(1): 113–116.

Williams, Peggy. 2009. "Fly Like an Eagle Ford." *Oil and Gas Investor*, 29(7): 51–54.

Williams, Peggy. 2010*a*. "Bone Spring Springs Forth." *Oil and Gas Investor*, 30(10): NA.

Williams, Peggy. 2010*b*. "Oklahoma's Woodford." *Oil and Gas Investor*, 30(9): NA.

Williams, Peggy. 2010*c*. "The Niobrara." *Oil and Gas Investor*, 30(8): NA.

Williams, Peggy. 2011. "Unconventional Exclusive: Unlocking the Utica?" *Oil and Gas Investor This Week*, 19(1): 3.

Zagorski, W. A., G.R. Wrightstone, and D.C. Bowman. 2012. "The Appalachian Basin Marcellus Gas Play: Its History of Development, Geologic Controls on Production, and Future Potential as a World-class Reservoir." In *Shale reservoirs - Giant resources for the 21st century. AAPG Memoir*, , ed. J. A. Breyer, 172 – 200.

Zuckerman, Gregory. 2013. *The Frackers: The Outrageous Inside Story of the New Billionare Wildcatters.* New York, NY:Penguin.

VII. Appendix Tables

Appendix Table 1—: First-frac dates

Play	Basin	First Production Date	First Publicity Date	Rystad Data Available
(1)	(2)	(3)	(4)	(5)
Barnett	Fort Worth	6/11/1998	7/19/2000	Yes
Permian	Permian Plays	1996	5/9/2005	Yes
Woodford-Arkoma	Arkoma	8/15/2004	3/1/2006	Yes
Fayetteville	Arkoma	5/13/2004	8/17/2004	Yes
Bakken	Williston	5/26/2000	2/1/2007	Yes
Marcellus	Appalachian	10/20/2004	12/9/2007	Yes
Woodford-Ardmore	Ardmore	5/19/2005	1/10/2007	Yes
Woodford-Anadarko	Anadarko	11/15/2007	2/28/2008	Yes
Hayesville/Bossier	TX-MS-LA-Salt	3/2/2007	3/24/2008	Yes
Eagle Ford	Western Gulf	8/28/2008	10/21/2008	Yes
Niobrara-Denver-Julesberg	Denver-Julesburg	9/5/2009	4/7/2010	Yes
Niobrara-Powder River	Niobrara	9/5/2009	4/7/2010	Yes
Utica	Appalachian	3/22/2010	7/28/2011	Yes
Niobrara-Greater Green River	Greater Green River	2011	12/21/2011	Yes

Notes: This Table shows first production, first publicity, and first frac dates for all plays listed in the NETL (2013) document as being actively developed shale gas plays, as well as the shale oil plays included in the Rystad data. In the analysis, we assign the first frac year to be the year of the first frac date if the month is June or earlier. We assign the year to be the first frac year plus one if the month is July or later.

Appendix Table 2—: Impact of fracing on building permits

	(1)	(2)	(3)
Log(Unit Permits)			
Top Quartile Effect at τ=5	0.201**	0.302*	0.341*
	(0.092)	(0.181)	(0.185)
Rystad Top Quartile Level Shift	Y	Y	Y
Rystad Top Quartile Trend	N	Y	Y
Rystad Top Quartile Trend Break	N	Y	Y
County Fixed Effects	Y	Y	Y
County-Specific Trends	N	N	Y
Year-Play Fixed Effects	N	Y	Y
Restricted to Balanced Sample	N	N	Y

Notes: This table reports regressions of housing permits on different Rystad top quartile variables. Housing supply measures come from the Census Bureau's "New Residential Construction" dataseries. To avoid dropping counties with 0 permits in a given year, we use log(permits + 1) as the outcome variable. The sample includes all counties in any shale basin from 1990 to 2013 with non-missing permit data in all years. Column (1) allows for a level shift in Rystad top quartile counties. Columns (2) and (3) allow for pre-trends, a post fracing level shift, and a post fracing trend break in Rystad top quartile counties. In Columns (1) and (2), all Rystad top quartile variables are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first frac date after 2008. The reported coefficients are for the balanced sample. Column (3) adds county-specific trends and restricts the sample to the balanced sample. The reported estimates and standard errors correspond to the top quartile level shift coefficient + 5 times the top quartile trend break coefficient. Standard errors clustered at the county level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Columns (1) and (2) include 9120 observations from 380 total counties, of which 61 Rystad top quartile and 237 outside top quartile counties are in the balanced sample. Column (3) includes 5,066 observations from 298 total counties, of which 61 Rystad top quartile and 237 outside top quartile counties are in the balanced sample.

Appendix Table 3—: Impact of fracing on demographics

	(1)
Panel A: Age/Sex Shares	
Male	
	0.001
	(0.001)
Prime Age Males	
	0.005*
	(0.003)
Prime Age Females	
	0.003
	(0.003)
All Not Working Age	
	-0.008
	(0.005)
Panel B: Social Characteristic Shares	
Never Married	
	0.005
	(0.007)
White	
	-0.005
	(0.006)
Panel C: Education Shares	
Less than High School	
	0.000
	(0.003)
High School Degree	
	-0.013***
	(0.005)
Some College	
	-0.005
	(0.004)
College Degree +	
	0.018***
	(0.007)
Play FE	Y

Notes: This table shows regressions on the change in demographic characteristics between 2000 and 2009-2013 on an indicator for the county being in the fourth quartile of the Rystad max prospectivity score among counties within the shale with a non-missing Rystad value. The Rystad top quartile indicator is included by itself as well as interacted with an indicator for being in the unbalanced sample, defined as having a first frac date after 2008. The reported estimates are for the balanced sample. Data come from the 2013-2009 ACS and 2000 Decenniel Census. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Includes observations from 404 total counties, of which 65 Rystad top quartile and 253 outside top quartile counties are in the balanced sample.

Appendix Table 4—: Robustness of hydrocarbon, total income, and housing price results

	Balanced Sample: 3 Years of Outcome Data or More		Balanced Sample: 2 Years of Outcome Data or More
	(1)	(2)	(3)
Panel A: Total Value of Hydrocarbon Production			
Top Quartile(Max Rystad)	409***	390***	439***
	(123)	(118)	(111)
	[133]	[122]	[151]
Top Tercile(Max Rystad)	350***	366***	367***
	(96)	(99)	(88)
Top Octile(Max Rystad)	521**	481**	567***
	(221)	(215)	(183)
Top Quartile(Mean Rystad)	287***	267***	332***
	(106)	(93)	(98)
Top Tercile(Mean Rystad)	246***	227***	262***
	(82)	(74)	(77)
Top Octile(Mean Rystad)	545***	503***	577***
	(168)	(150)	(147)
Panel B. Log(Total Income)			
Top Quartile(Max Rystad)	0.061**	0.080**	0.074**
	(0.031)	(0.033)	(0.031)
	[0.037]	[0.036]	[0.035]
Top Tercile(Max Rystad)	0.062**	0.089***	0.066**
	(0.027)	(0.030)	(0.027)
Top Octile(Max Rystad)	0.053	0.080*	0.056
	(0.043)	(0.043)	(0.045)
Top Quartile(Mean Rystad)	0.024	0.030	0.029
	(0.029)	(0.028)	(0.029)
Top Tercile(Mean Rystad)	0.040*	0.049**	0.037
	(0.024)	(0.024)	(0.025)
Top Octile(Mean Rystad)	0.034	0.044	0.057*
	(0.035)	(0.037)	(0.034)
Panel C: Log(Median Housing Value)			
Top Quartile(Max Rystad)	0.057***	0.027**	0.053***
	(0.012)	(0.013)	(0.011)
	[0.027]	[0.022]	[0.026]
Top Tercile(Max Rystad)	0.029***	-0.005	0.028***
	(0.010)	(0.011)	(0.010)
Top Octile(Max Rystad)	0.047***	0.022	0.042**
	(0.017)	(0.017)	(0.017)
Top Quartile(Mean Rystad)	0.049***	0.024**	0.048***
	(0.011)	(0.011)	(0.011)
Top Tercile(Mean Rystad)	0.044***	0.021**	0.042***
	(0.010)	(0.010)	(0.010)
Top Octile(Mean Rystad)	0.061***	0.023	0.061***
	(0.013)	(0.014)	(0.012)
Play(-Year) FE	Y	Y	Y
State(-Year) FE	N	Y	N

Notes: This table shows shows regressions of Hydrocarbon Production (Panel A), Log(Total Income) (Panel B), and Log(Median Housing Value) (Panel C) on different proxies for fracing potential. Panels A and B are specifications using annual data that allow for pre-trends, a level shift, and a trend break in exposed counties. The reported estimates and standard errors in Panels A(B) correspond to the mean shift coefficient + 3(4) times the trend break coefficient. Panel C reports results from long-difference specifications that allow for a level-shift in exposed counties. Different columns report estimates for different fixed effects and first-frac date restrictions. Columns (1) through (3) all include Play-Year (Play) fixed effects for Panels A and B(C). Column (2) adds state-year (state) fixed effects for Panels A and B (C). Column (3) changes the balanced sample definition to include all shale plays with a first-frac date in or before 2009. All Rystad variables are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as starting fracing after 2008 in columns (1) and (2) (or 2009 for column (3)). The reported coefficients are for the balanced sample. In Panels A and B, standard errors clustered at the county level are reported in parentheses. In Panel C, robust standard errors are reported in parentheses. Top Quartile(Max Rystad) rows also include Conley standard errors that allow for spatial correlation in brackets. *** p<0.01, ** p<0.05, * p<0.1

Sample: Panel A(B)[C] includes 8,100(9,246)[404] county-year observations from 405(402)[404] counties.

In all panels, the top tercile, quartile, and octile of maximum Rystad productivity include 88, 65, and 32 balanced sample counties in Columns (1) and (2) and 97, 72, and 39 balanced sample counties in Column (3). In Panels A(B)[C], the top tercile, quartile, and octile of mean Rystad productivity include 103(103)[103], 76(76)[76], and 38(38)[38] balanced sample counties in Columns (1) and (2) and 112(112)[112], 83(83)[83], and 42(42)[42] balanced sample counties in Column (3).

In all panels, outside of the top tercile, quartile, and octile of the maximum of Rystad productivity there are 230, 253, and 286 in Columns (1) and (2) and 249, 274, and 307 balanced sample counties in Column (3). In Panels A(B)[C], outside of the top tercile, quartile, and octile of mean prospectivity there are 215(214)[215], 242(241)[242], and 280(279)[280] balanced sample counties in Columns (1) and (2) and 234(233)[234], 263(262)[263], and 304(303)[304] balanced sample counties in Column (3).

Appendix Table 5—: Play specific estimates: additional outcomes

	All	Bakken	Barnett	Fayetteville	Haynesville	Marcellus	Woodford, Anadarko	Woodford, Ardmore	Woodford, Arkoma	Permian Plays		Joint F-test	Eagle Ford
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		(11)	(12)
Panel A: Average Characteristics of Top Quartile Counties													
Population (2000)	64,860	6,307	109,202	24,046	24,576	112,911	45,516	19,537	9,955	15,221			36,836
Oil Share of Hydrocarbon Production Value (2011)	0.33	0.94	0.42	0.00	0.01	0.07	0.34	0.48	0.01	0.64			0.65
Panel B: Hydrocarbon Production													
B1. New Wells													
	29	102***	51	39**	164***	6	12	19	65**	40**	F-stat	3.6	146***
	(21)	(33)	(39)	(17)	(48)	(22)	(24)	(13)	(29)	(19)	p-value	0.00	(27)
B2. Total Producing Wells													
	60	273***	388**	6	212	388***	-204***	150	-9	107	F-stat	10.4	205***
	(241)	(90)	(198)	(51)	(205)	(127)	(27)	(142)	(93)	(85)	p-value	0.00	(52)
Panel C: Labor Markets													
C1. Log(Mining Industry Employment)													
	0.140***	0.329**	0.340**	0.354	-0.024	0.089	0.029	0.183	-0.039	0.114	F-stat	1.4	0.104
	(0.050)	(0.144)	(0.169)	(0.371)	(0.178)	(0.077)	(0.375)	(0.240)	(0.287)	(0.124)	p-value	0.17	(0.156)
C2. Log(Total Employment)													
	0.055*	0.301***	0.033*	-0.065***	0.114	-0.024	-0.122***	-0.043	0.018	0.104**		7.1	0.326**
	(0.029)	(0.085)	(0.018)	(0.013)	(0.077)	(0.018)	(0.028)	(0.098)	(0.078)	(0.047)		0.00	(0.134)
C3. Unemployment Rate													
	-0.006**	-0.015	0.002	0.045	0.004	-0.009***	-0.017	0.006	0.017	-0.003	F-stat	1.2	-0.032***
	(0.003)	(0.023)	(0.006)	(0.033)	(0.015)	(0.003)	(0.023)	(0.022)	(0.039)	(0.013)	p-value	0.32	(0.012)
C4. Log(Median Household Income)													
	0.060***	0.300***	0.061**	0.041	0.032	0.055***	0.140	-0.001	0.020	0.094*	F-stat	5.1	0.050
	(0.010)	(0.084)	(0.026)	(0.112)	(0.054)	(0.012)	(0.086)	(0.080)	(0.136)	(0.049)	p-value	0.00	(0.047)
C5. Log(In-migration)													
	0.073*	0.379***	-0.144	-0.336***	0.057	0.046	-0.174***	-0.026	0.008	0.044	F-stat	11.3	0.421***
	(0.038)	(0.109)	(0.092)	(0.039)	(0.094)	(0.033)	(0.031)	(0.057)	(0.173)	(0.144)	p-value	0.00	(0.135)
C6. Log(Total Population)													
	0.027*	0.130***	0.071	-0.014	-0.045	0.018	0.060	0.042	-0.038	-0.007	F-stat	1.4	-0.090*
	(0.016)	(0.045)	(0.053)	(0.115)	(0.055)	(0.024)	(0.117)	(0.075)	(0.089)	(0.039)	p-value	0.20	(0.048)
Panel D. Quality of Life													
D1. Prime Age Male Total Population Share													
	0.005***	0.031**	-0.003	-0.002	-0.002	0.007***	0.006	-0.005	-0.026	0.008	F-stat	3.0	-0.001
	(0.002)	(0.013)	(0.004)	(0.017)	(0.008)	(0.002)	(0.012)	(0.012)	(0.019)	(0.007)	p-value	0.00	(0.006)
D2. Log(Violent Crimes)													
	0.208*	0.873***	0.006	-0.759**	0.058	0.206	0.384***	-0.234	-1.613***	-0.238	F-stat	11.7	1.148**
	(0.124)	(0.234)	(0.299)	(0.340)	(0.529)	(0.145)	(0.122)	(0.227)	(0.197)	(0.278)	p-value	0.00	(0.498)
D3. Log(Property Crimes)													
	-0.057	0.422	-0.261	0.137	0.151	-0.092	-0.283**	-0.133	-0.546***	0.098	F-stat	1.9	0.712
	(0.087)	(0.490)	(0.196)	(0.129)	(0.197)	(0.130)	(0.120)	(0.190)	(0.201)	(0.195)	p-value	0.06	(0.639)
Panel E. Rental Income and Housing													
E1. Log(Total Income from Rents)													
	0.109***	0.331***	0.068	-0.002	0.340***	0.055**	0.014	0.154**	0.082	-0.094	F-stat	4.4	0.608***
	(0.041)	(0.074)	(0.051)	(0.030)	(0.115)	(0.025)	(0.055)	(0.063)	(0.115)	(0.095)	p-value	0.00	(0.170)
E2. Log(New Housing Unit Permits)													
	0.302*	1.390***	0.331	-1.017***	0.825	0.005	-0.222	0.867	0.204	-0.305	F-stat	4.8	-0.452
	(0.181)	(0.354)	(0.632)	(0.199)	(1.270)	(0.162)	(0.268)	(0.581)	(0.814)	(0.309)	p-value	0.00	(1.446)
E3. Log(Total Housing Units)													
	0.011	0.073*	0.045	0.032	-0.009	0.017	0.034	0.040	0.020	-0.076**	F-stat	1.3	-0.092**
	(0.013)	(0.037)	(0.044)	(0.096)	(0.046)	(0.020)	(0.097)	(0.062)	(0.074)	(0.032)	p-value	0.21	(0.040)

Notes: This table shows estimates from regressions of outcome variables on Rystad top quartile variables interacted with dummies for being in particular shale plays. Column (1) shows the estimate for all counties with first frac dates in or before 2008. Columns (2)-(10) show play-specific results for all plays with first frac dates in or before 2008. Column (11) presents results from the Join F-test that the coefficients are equal for all plays with first frac dates in or before 2008. Column (12) reports results for the Eagle Ford, the one shale play with a first frac date in 2009. Panel A shows summary statistics on average county population and the oil share of hydrocarbon production. All specifications except for housing prices are time series estimates corresponding to column (2) in the main tables. Panels B and C allow for pre-trends, a level shift, and a trend break in the top quartile indicators, and also include play-year fixed effects. The reported estimates in Panels B, C2, C5, D2, D3, E1, and E2 correspond to the top quartile mean shift coefficient + tau (= T - 2008) times the top quartile trend break coefficient, where T is the latest year of data for the given outcome variable. In practice, this means evaluating the effect of being in a top quartile county 3 years after the start of fracing for Panel B, 4 years after the start of fracing for Panels C2, C5, and E1 and 5 years after the start of fracing for Panels D2 and D3. Panels C1, C3, C4, C6, D1, and E3 report long difference specifications of the change in the given outcome between 2000 and 2009-2013 on an indicator for being in the Rystad top quartile. In all panels, all Rystad top quartile variables are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first frac date after 2008. The reported coefficients are for the balanced sample. Panel B data come from HPDI well data aggregated to the county level. Panel C2 and E1 data come from the REIS data produced by the BEA. Panel C5 data come from the IRS county-to-county migration data. Panel E2 data come from the Census Bureau's New Residential Construction data series. Panel D2 and D3 data come from the FBI's Uniform Crime Records (UCR). Panel C1, C3, C4, C6, D1, and E3 data come from the 2000 Decennial Census and the 2009-2013 American Community Survey. In Panels B, C2, C5, D2, D3, E1, and E2 standard errors clustered at the county level are reported in parentheses. In Panels C1, C3, C4, C6, D1, and E3, robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 6—: Play specific estimate:s Welfare Estimates Under Alternative Assumptions

	All	Bakken	Barnett	Fayetteville	Haynesville	Marcellus	Woodford, Anadarko	Woodford, Ardmore	Woodford, Arkoma	Permian Plays	Eagle Fo
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(12)
Panel A: Using Change in Median Rents (dollars)											
A1. Change in Amenities and Welfare per Household (dollars)											
Change in amenities	-2,225***	-4,323*	-2,834	1,962***	-1,549*	-2,859***	-1,604*	357	-2,035**	-4,853***	805
	(842)	(2,572)	(2,568)	(578)	(922)	(938)	(965)	(2,092)	(924)	(817)	(1,722)
Change in welfare	1,716***	9,065***	-274	5,365***	1,265	1,028	741	18	1,037	1,216	1,324
	(610)	(2,099)	(2,187)	(857)	(1,346)	(640)	(871)	(1,391)	(1,072)	(1,558)	(1,596)
A2: Total Change in Amenities and Welfare (billions of dollars)											
Change in amenities	-74***	-1.76*	-11.37	0.40***	-1.43*	-72.92***	-0.56*	0.22	-0.30**	-5.75***	1.16
	(28)	(1.05)	(10.30)	(0.12)	(0.85)	(23.91)	(0.33)	(1.29)	(0.14)	(0.97)	(2.48)
Change in welfare	57***	3.70***	-1.10	1.09***	1.17	26.23	0.26	0.01	0.15	1.44	1.91
	(20)	(0.86)	(8.77)	(0.17)	(1.24)	(16.33)	(0.30)	(0.86)	(0.16)	(1.85)	(2.30)
Panel B: Using Change in Median Home Values (dollars)											
A1. Change in Amenities and Welfare per Household (dollars)											
Change in amenities	-1,405*	-2,579	-2,148	-294	-3,675***	-556	-1,284	1,175	-2,090*	-5,839***	-1,040
	(734)	(2,362)	(2,882)	(552)	(1,311)	(1,121)	(941)	(1,791)	(1,124)	(1,191)	(1,753)
Change in welfare	2,537***	10,809***	412	3,109***	-861	3,331***	1,061	836	981	230	-520
	(758)	(2,157)	(2,245)	(844)	(1,717)	(999)	(873)	(1,535)	(1,499)	(1,745)	(1,593)
A2: Total Change in Amenities and Welfare (millions of dollars)											
Change in amenities	-47*	-1.05	-8.62	-0.06	-3.39***	-14.19	-0.45	0.73	-0.31*	-6.92***	-1.50
	(24)	(0.96)	(11.56)	(0.11)	(1.21)	(28.60)	(0.33)	(1.11)	(0.16)	(1.41)	(2.53)
Change in welfare	85***	4.41***	1.65	0.63***	-0.80	84.96***	0.37	0.52	0.14	0.27	-0.75
	(25)	(0.88)	(9.01)	(0.17)	(1.59)	(25.47)	(0.30)	(0.95)	(0.22)	(2.07)	(2.30)
Top Quartile Counties	65	8	5	1	5	28	1	4	2	11	7
Outside Top Quartile Counties[a]	253	27	41	13	21	95	10	5	7	34	21

Notes: This table reports estimates of the effect of fracing on amenities and welfare in dollars for each shale play. The calculations are made using our preferred values of the share of wage and salary income spent on housing (²) and the standard deviation of idiosyncratic preferences for location (s) o .65 and s=.4 respectively. Different columns report results for different shale play while different panels report results for different estimamtes of the change in housing costs. Panel A reports results where the change in housing costs is measured using the estimated percent change in median rents, v Panel B shows estimates where the change in housings costs is measured using the estimated percentage change in median home prices. For each measure of the change in housing costs, we report both the estimated change in amenities and the estimated change in total welfare. The calculations converted to dollars using the mean household wage and salary income and mean household interest and dividend income in top quartile counties in each shale play. We aggregate these figures to the total impact of fracing in aggregate welfare in top quartile counties assuming a discount rate of 5 perc and using the mean number of households in top quartile counties and total number of top quartile counties in each shale play. Overall calculations are made excluding the Eagle Ford play.

Appendix Table 7—: Play specific estimates: State and play fixed effects

	All (1)	Bakken (2)	Barnett (3)	Fayetteville (4)	Haynesville (5)	Marcellus (6)	Woodford, Anadarko (7)	Woodford, Ardmore (8)	Woodford, Arkoma (9)	Permian Plays (10)		Joint F-test (11)	Eagle Ford (12)
Panel A: Average Characteristics of Top Quartile Counties													
Population (2000)	64,860	6,307	109,202	24,046	24,576	112,911	45,516	19,537	9,955	15,221			36,836
Oil Share of Hydrocarbon Production Value (2011)	0.33	0.94	0.42	0.00	0.01	0.07	0.34	0.48	0.01	0.64			0.65
Panel B: Hydrocarbon Production													
B1. Total Value of Hydrocarbon Production													
	390***	978**	322*	69	1,417*	162**	-452***	123*	199	166	F-stat	10.9	1,412***
	(118)	(420)	(185)	(79)	(838)	(64)	(65)	(71)	(161)	(132)	p-value	0.00	(273)
Panel C: Labor Markets													
C1. Log(Real mean total income per hh)													
	0.047***	0.296***	0.045**	0.099	0.071	0.029**	0.069	-0.013	0.000	0.200***	F-stat	4.7	-0.015
	(0.011)	(0.076)	(0.023)	(0.100)	(0.050)	(0.014)	(0.076)	(0.071)	(0.121)	(0.052)	p-value	0.00	(0.042)
C2. Log(Real mean wage and salary income per hh)													
	0.056***	0.292***	0.031	-0.014	0.086	0.049***	0.079	-0.028	0.075	0.204***	F-stat	3.9	-0.003
	(0.013)	(0.090)	(0.027)	(0.118)	(0.059)	(0.016)	(0.090)	(0.084)	(0.143)	(0.062)	p-value	0.00	(0.049)
C3. Log(Real mean rent, interest, and dividend income per hh)													
	0.084*	0.799**	0.061	0.671	-0.025	0.074	-0.171	0.116	0.495	0.007	F-stat	1.4	0.173
	(0.045)	(0.318)	(0.096)	(0.419)	(0.209)	(0.057)	(0.320)	(0.298)	(0.507)	(0.219)	p-value	0.20	(0.175)
C4. Log(Population)													
	0.028	0.128***	0.071	-0.014	-0.023	0.019	0.060	0.042	-0.038	-0.019	F-stat	1.2	-0.090*
	(0.017)	(0.047)	(0.053)	(0.116)	(0.063)	(0.028)	(0.117)	(0.075)	(0.089)	(0.039)	p-value	0.29	(0.049)
Panel D: Housing													
D1. Log(Median Home Values)													
	0.027**	0.239***	-0.052*	-0.008	-0.121**	0.058***	-0.031	-0.038	0.010	0.026	F-stat	3.4	0.002
	(0.013)	(0.081)	(0.028)	(0.104)	(0.055)	(0.016)	(0.086)	(0.077)	(0.129)	(0.056)	p-value	0.00	(0.052)
D2. Log(Median Rents)													
	0.007	0.080	-0.025	0.107	0.010	0.014	-0.072	0.018	0.066	0.031	F-stat	0.9	0.039
	(0.009)	(0.069)	(0.018)	(0.104)	(0.044)	(0.011)	(0.070)	(0.061)	(0.115)	(0.045)	p-value	0.52	(0.035)
Top Quartile Counties	65	8	5	1	5	28	1	4	2	11			7
Outside Top Quartile Counties[a]	253	27	41	13	21	95	10	5	7	34			21

Notes: This table shows estimates from regressions of outcome variables on Rystad top quartile variables interacted with dummies for being in particular shale plays. Column (1) shows the estimate for all counties with first frac dates in or before 2008. Columns (2)-(10) show play-specific results for all plays with first frac dates in

a. All panels include the same number of balanced sample top quartile and outside top quartile counties.

Appendix Table 8—: Robustness of wage, employment, and building permit results

	Balanced Sample: 3 Years of Outcome Data or More		Balanced Sample: 2 Years of Outcome Data or More
	(1)	(2)	(3)
Panel A. Log(Total Wage and Salary Income)			
Top Quartile(Max Rystad)	0.105***	0.122***	0.112***
	(0.035)	(0.037)	(0.036)
	[0.046]	[0.046]	[0.042]
Top Tercile(Max Rystad)	0.106***	0.133***	0.106***
	(0.031)	(0.034)	(0.032)
Top Octile(Max Rystad)	0.094**	0.115**	0.103**
	(0.047)	(0.046)	(0.047)
Top Quartile(Mean Rystad)	0.057*	0.058*	0.063*
	(0.033)	(0.032)	(0.034)
Top Tercile(Mean Rystad)	0.065**	0.068**	0.061**
	(0.028)	(0.028)	(0.029)
Top Octile(Mean Rystad)	0.071	0.083*	0.093**
	(0.043)	(0.045)	(0.047)
Panel B. Log(Employment)			
Top Quartile(Max Rystad)	0.055*	0.070**	0.071**
	(0.029)	(0.030)	(0.030)
	[0.030]	[0.031]	[0.030]
Top Tercile(Max Rystad)	0.074***	0.095***	0.081***
	(0.027)	(0.029)	(0.027)
Top Octile(Max Rystad)	0.044	0.065	0.066
	(0.042)	(0.041)	(0.043)
Top Quartile(Mean Rystad)	0.038	0.043	0.041
	(0.031)	(0.030)	(0.031)
Top Tercile(Mean Rystad)	0.049*	0.057**	0.042
	(0.026)	(0.026)	(0.026)
Top Octile(Mean Rystad)	0.082***	0.088***	0.100***
	(0.031)	(0.034)	(0.033)
Panel C. Log(Building Unit Permits)			
Top Quartile(Max Rystad)	0.302*	0.230	0.291
	(0.181)	(0.188)	(0.213)
	[0.206]	[0.210]	[0.221]
Top Tercile(Max Rystad)	0.294*	0.268*	0.294
	(0.151)	(0.161)	(0.179)
Top Octile(Max Rystad)	0.343	0.274	0.274
	(0.252)	(0.248)	(0.304)
Top Quartile(Mean Rystad)	0.246	0.167	0.194
	(0.160)	(0.169)	(0.181)
Top Tercile(Mean Rystad)	0.289**	0.226	0.264
	(0.140)	(0.148)	(0.161)
Top Octile(Mean Rystad)	0.125	0.038	-0.012
	(0.210)	(0.225)	(0.250)
Play(-Year) FE	Y	Y	Y
State(-Year) FE	N	Y	N

Notes: This table shows shows regressions of Log(Total Wage and Salary Income) (Panel A), Log(Employment) (Panel B), and Log(Building Units Permits) (Panel C) on different proxies for fracing potential. Data in Panel A and B come from the REIS data produced by the BEA. Data in Panel C come from the Census Bureau's "New Residential Construction" dataseries. All specifications use annual data and allow for pre-trends, a level shift, and a trend break in exposed counties. The reported estimates and standard errors correspond to the mean shift coefficient + 4[5] times the trend break coefficient in Panels A and B[C]. Different columns report estimates for different fixed effects and first frac date restrictions. Columns (1) through (3) all include play-year fixed effects. Column (2) adds state-year fixed effects. Column (3) changes the balanced sample definition to include all shale plays with a first frac date in or before 2009. All Rystad top quartile variables are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as starting fracing after 2008 in columns (1) and (2) (or 2009 for column (3)). The reported coefficients are for the balanced sample. Standard errors clustered at the county level are reported in parentheses. Top Quartile(Max Rystad) rows also include Conley standard errors that allow for spatial correlation in brackets. *** p<0.01, ** p<0.05, * p<0.1

Sample: Panel A and B[C] includes 9,246[9,120] county-year observations from 402[402] counties. The top tercile, quartile, and octile of maximum Rystad productivity include 88[83], 65[61], and 32[31] balanced sample counties respectively in Panels A and B[C]. The top tercile, quartile, and octile of mean Rystad productivity include 103[96], 76[69], and 38[36] balanced sample counties respectively in Panels A and B[C]

Appendix Table 9—: Impact of fracing on home values and rents across occupied and vacant homes

	(1)
Panel A: Log(Mean Home Values)	
Owner Occupied	0.056**
	(0.022)
Vacant-for-Sale	0.045
	(0.077)
All Owner Occupied and Vacant-for-Sale	0.071***
	(0.027)
Panel B: Log(Mean Rents)	
Renter Occupied	0.019
	(0.017)
Vacant-for-Rent	0.124***
	(0.046)
All Renter Occupied and Vacant-for-Rent	0.029
	(0.018)
Play FE	Y

Notes: This table shows regressions on the change in housing values and rents between 2000 and 2009-2013 on an indicator for the county being in the fourth quartile of the Rystad max prospectivity score among counties within the shale with a non-missing Rystad value. The Rystad top quartile indicator is included by itself as well as interacted with an indicator for being in the unbalanced sample, defined as having a first frac date after 2008. The reported estimates are for the balanced sample. Observations are weighted by the number of housing units in the county to address heteroskedasticity. Characteristics of vacant units for sale or for rent are not available in all years, so unlike our specifications in Table LINK, no covariates are included in this table. Data come from the 2000 Census and the 2009-2013 American Community Survey. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Includes observations from 345 total counties, of which 56 Rystad top quartile and 215 outside top quartile counties are in the balanced sample.

Appendix Table 10—: Differences in levels and changes of housing variable by universe definition

	Levels in 2000	Change Computed Using Specified Units as Base Year Compared to Full Universe	Change for Specified Units Compared to Change in Full Universe
	(1)	(2)	(3)
Panel A. Renter Occupied Units			
A1. Number of Units			
	-1.87	2.17	-2.93
A2. Mean Gross Rent			
	-0.10	0.12	0.05
A3. Median Gross Rent			
	-0.03	0.03	0.03
Panel B. Owner Occupied Units			
B1. Number of Units			
	-22.40	32.96	5.25
B2. Mean Gross Rent			
	1.34	-1.21	-1.32
B3. Median Gross Rent			
	4.34	-4.51	-1.19
Number of PUMAs	152	152	152

Notes: This table explores the differences in levels and changes in Public Use Microdata Area (PUMA) level housing outcomes using different universes. Panel A reports results for renter occupied units while Panel B reports results for owner occupied units. Column (1) reports the percentage difference in levels in 2000 for specified housing units compared to the full universe of housing units. Column (2) reports the percentage difference in the change from 2000 to 2007/11 when computed using the specified universe for 2000 and the full universe for 2007/11 compared to using the full universe for all variables. Column (3) reports the percentage difference in the change from 2000 to 2007/11 when using the specified universe for all variables compared to using the full universe for all variables. All differences are reported in percentage terms - i.e. a "1" is equal to a 1 percent difference in the given outcome. Data for 2000 come from the decennial census while data for 2007/11 come from the American Community Survey. All data were retrieved from IPUMS-USA (Ruggles, et al, 2015). The sample is restricted to PUMAs that intersect shale plays with first frac dates in or before 2008.

Appendix Table 11—: Impact of fracing on governmental employment and payroll

	(1)
Panel A: Log(Total Employment):	
	0.022
	(0.022)
A1. Log(Education Employment): 63 percent of direct local government expenditures.	
	-0.013
	(0.023)
A2. Log(Public Safety Employment): 10 percent of direct local government expenditures.	
	0.047
	(0.042)
A3. Log(Welfare and Hospital Employment): 9 percent of direct local government expenditures.	
	0.177
	(0.143)
A4. Log(Infrastructure Employment): 8 percent of direct local government expenditures.	
	0.089***
	(0.032)
A5. Log(Other Employment): 10 percent of direct local government expenditures.	
	0.064
	(0.054)
Panel B: Log(Total Real Payroll):	
	0.067***
	(0.023)
B1. Log(Education Payroll): 61 percent of direct local government expenditures.	
	0.038
	(0.025)
B2. Log(Public Safety Payroll): 10 percent of direct local government expenditures.	
	0.082*
	(0.045)
B3. Log(Welfare and Hospital Payroll): 9 percent of direct local government expenditures.	
	0.165
	(0.129)
B4. Log(Infrastructure Payroll): 8 percent of direct local government expenditures.	
	0.136***
	(0.043)
B5. Log(Other Payroll): 11 percent of direct local government expenditures.	
	0.082
	(0.052)
Play Fixed Effects	Y

Notes: This table shows regressions on the change in government employment and payroll between 2002 and 2012 on an indicator for the county being in the fourth quartile of the Rystad max prospectivity score among counties within the shale with a non-missing Rystad value. The Rystad top quartile indicator is included by itself as well as interacted with an indicator for being in the unbalanced sample, defined as having a first frac date after 2008. The reported estimates are for the balanced sample. Data come from the 2012 and 2002 Census of Governments. Employment is in terms of full time equivalent workers. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Includes observations from 405 total counties, of which 65 Rystad top quartile and 253 outside top quartile counties are in the balanced sample.

Appendix Table 12—: Impact of fracing on local government revenues and expenditures: 2012-1997

	(1)
Panel A: Log(Total Expenditures): 2012 - 1997	
	0.149***
	(0.037)
A. Log(Direct Expenditures)	
	0.141***
	(0.036)
A1. Direct Expenditures by Type	
A1a. Log(Current Operating Expenditure)	
	0.116***
	(0.034)
A1b. Log(Capital Outlays)	
	0.102
	(0.115)
A2. Direct Expenditures by Purpose	
A2a. Log(Education Expenditures)	
	0.034
	(0.034)
A2b. Log(Public Safety Expenditures)	
	0.214***
	(0.069)
A2c. Log(Welfare and Hospital Expenditures)	
	0.460**
	(0.189)
A2d. Log(Infrastructure and Utility Expenditures)	
	0.212***
	(0.070)
A2e. Log(Other Expenditures)	
	0.193***
	(0.068)
Panel B: Log(Total Revenues): 2012 - 2002	
	0.173***
	(0.034)
B1. Revenues by Type	
B1a. Log(Property Tax Revenues)	
	0.120**
	(0.048)
B1b. Log(Sales Tax Revenues)	
	0.316**
	(0.139)
B1c. Log(Other Tax Revenues)	
	0.145
	(0.145)
B1d. Log(Intergovernmental Revenues)	
	0.070
	(0.074)
B1e. Log(Charges Revenues)	
	0.098
	(0.092)
B1f. Log(Other Revenues)	
	0.226***
	(0.075)
Panel C: Government Balance Sheets	
C1. Log(Outstanding Debt)	
	0.127
	(0.153)
C1a: log(Long-Run Debt)	
	0.122
	(0.153)
C2a: log(Short-term debt)	
	0.830
	(0.622)
C2. Log(Cash and Securities)	
	0.180*
	(0.092)
C2a: Log(Insurance Trust Cash and Security Holdings)	
	0.374**
	(0.163)
C2b: Log(Non Insurance Trust Cash and Security Holdings)	
	0.155
	(0.095)
Panel D: Log(Elem/Sec Education Spending per Pupil)	
	0.025
	(0.037)
Play FE	Y

Notes: This table shows regressions on the change in government spending and revenues between 199 and 2012 on an indicator for the county being in the fourth quartile of the Rystad max prospectivity scor among counties within the shale with a non-missing Rystad value. The Rystad top quartile indicator i included by itself as well as interacted with an indicator for being in the unbalanced sample, defined a having a first frac date after 2008. The reported estimates are for the balanced sample. Data come fror the 2012 and 1997 Census of Governments. Robust standard errors are reported in parentheses. ** p<0.01, ** p<0.05, * p<0.1

Sample: Panels A, B, and C Include observations from 404 total counties, of which 65 Rystad top quartil and 253 outside top quartile counties are in the balanced sample. Panel D includes observations from a 385 counties over shale plays with non-missing school enrollment data from all districts in 1997, 2002, an 2012, of which 61 Rystad top-quartile, and 244 outside top quartile counties are in the balanced sample.

Appendix Table 13—: Impact of fracing on agricultural land acreage: 2012-2007

	(1)
Panel A: Aggregate Housing Supply and Land Use	
Acres of Agricultural Land	
	0.067
	(0.136)
Play FE	Y

Notes: This table reports regressions of agricultural land variables on an indicator for the county being in the fourth quartile of the Rystad max prospectivity score among counties within the shale with a non-missing Rystad value. The Rystad top quartile indicator is included by itself as well as interacted with an indicator for being in the unbalanced sample, defined as having a first frac date after 2008. The reported estimates are for the balanced sample. Data come from the census of agriculture. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Includes observations from 329 total counties, of which 51 Rystad top quartile and 198 outside top quartile counties are in the balanced sample.

Appendix Table 14—: Impact of fracing on the value of hydrocarbon production: Propensity-score estimates

	(1)	(2)
Panel A: Total Value of Oil and Gas Production		
1(Exposed)*1(Post)	52**	19**
	(23)	(8)
t*1(Exposed)		
t*1(Exposed)*1(Post)	136***	13*
	(36)	(7)
Fracing Exposure Effect at τ=3	459***	59***
	(108)	(23)
Fracing Exposure Group	Top Quartile	Quartiles 1-3
Control Group	Pscore Matched Sample	Pscore Matched Sample
Fracing Exposure Level Shift	Y	Y
Fracing Exposure Trend	Y	Y
Fracing Exposure Trend Break	Y	Y
County Fixed Effects	Y	Y
County-Specific Trends	Y	Y
Year-Play Fixed Effects	Y	Y
Restricted to Balanced Sample	Y	Y

Notes: This table reports regressions of oil/gas production variables on different Rystad top quartile variables. Oil and gas production data come from HPDI well data aggregated to the county level. Both columns allow for pre-trends, a post fracing level shift, and a post-fracing trend break in Rystad top quartile counties, as well as county-specific time trends. The sample is restricted to the balanced sample. Column (1) reports specifications where the exposed group is top quartile counties and the control group is the pscore-matched sample. Column (2) reports versions of the specifications in Column (1), but the exposed group is instead Rystad quartiles one through three. 1(Exposure) = 1 if the county is in the exposure group for the given column. 1(Post) = 1 if the year is after the first-frac date for the shale, defined as the first year that there is any fracing within the counties shale play. The coefficients and standard errors for the Fracing Exposure Effect at τ=3 correspond to the 1(Exposed)*1(Post) coefficient plus 3 times the t*1(Exposed)*1(Post) coefficeint. Standard errors clustered at the county level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Column (1) includes 13,195 observations from 1,015 county-play combinations, of which 65 Rystad top quartile and 558 unique pscore-matched counties are in the balanced sample. Column (2) includes 15,639 observations from 1,203 total county-play combinations, of which 253 Rystad quartiles one through three and 558 unique pscore-matched counties are in the balanced sample.

Appendix Table 15—: Impact of fracing on employment and aggregate income, time-series specifications: Propensity-score estimates

	(1)	(2)
Panel A: Log(Total Employment)		
Fracing Exposure Effect at τ=4	0.104***	0.059***
	(0.023)	(0.010)
Panel B: Income		
Log(Total Income)		
Fracing Exposure Effect at τ=4	0.085***	0.055***
	(0.029)	(0.011)
B1. Log(Total Wage/Salary Income): 56 percent of total personal income		
Fracing Exposure Effect at τ=4	0.146***	0.084***
	(0.036)	(0.012)
B2. Log(Total Rents/Dividends): 19 percent of total personal income		
Fracing Exposure Effect at τ=4	0.158***	0.079***
	(0.031)	(0.013)
B3. Log(Total Transfers): 10 percent of total personal income		
Fracing Exposure Effect at τ=4	0.007	0.012**
	(0.008)	(0.006)
B4. Log(Total Proprieter's Income): 18 percent of total personal income		
Fracing Exposure Effect at τ=4	0.014	0.009
	(0.070)	(0.045)
Panel C: Migration		
C1. Log(In Migration)		
Fracing Exposure Effect at τ=4	0.129***	0.131***
	(0.044)	(0.018)
C2. Log(Out Migration)		
Fracing Exposure Effect at τ=4	0.015	0.064***
	(0.031)	(0.016)
Fracing Exposure Group	Top Quartile	Quartiles 1-3
Control Group	Pscore Matched Sample	Pscore Matched Sample
Fracing Exposure Level Shift	Y	Y
Fracing Exposure Trend	Y	Y
Fracing Exposure Trend Break	Y	Y
County Fixed Effects	Y	Y
County-Specific Trends	Y	Y
Year-Play Fixed Effects	Y	Y
Restricted to Balanced Sample	Y	Y

Notes: This table reports regressions of aggregate economic outcomes on different Rystad fracing exposure measures. Employment and income variables in variables in Panels A and B come from the REIS data produced by the BEA. Migration measures in Panel C come from the IRS' county migration data. Both columns allow for pre-trends, a post fracing level shift, and a post fracing trend break in exposed counties as well as county-specific trends. The sample is restricted to the balanced sample. Column (1) reports specifications where fracing exposure is measured using an indicator for being in the top quartile of Rystad max prospectivity and the control group is the pscore-matched sample. Column (2) reports specifications where the fracing exposure measure is instead Rystad quartiles one through three. The reported estimates and standard errors correspond to the fracing exposure level shift coefficient + 4 times the fracing exposure trend break coefficient. Standard errors clustered at the county level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Panels A, B, B1, B2, and B3, Column (1) include 16,240 observations from 1,015 total county-play combinations, of which 65 Rystad top quartile and 558 unique pscore-matched counties are in the balanced sample. Column (2) include 19,232 observations from 1,202 total county-play combinations, of which 252 Rystad quartiles one through three counties and 558 unique pscore-matched counties are in the balanced sample.

Panel B4, Column (1) includes 16,016 observations from 1,001 county-play combinations, of which 59 Rystad top quartile and 553 unique pscore matched counties are in the balanced sample. Panel B4, Column (2) includes 18,864 observations from 1,179 county-play combinations, of which 237 Rystad quartile one through three counties and 553 unique pscore matched counties are in the balanced sample.

Panel C, Column (1) includes 13,078 observations from 1,006 county-play combinations, of which 63 Rystad top quartile and 556 unique pscore-matched counties are in the balanced sample. Panel C, Column (2) includes 15,483 observations from 1,191 county-play combinations, of which 248 Rystad quartile one through three and 556 unique pscore-matched counties are in the balanced sample.

Appendix Table 16—: Impact of fracing on employment and aggregate income, long-difference specifications: Propensity-score estimates

	(1)	(2)
Panel A: Employment Outcomes:		
A1. Log(Total Employment)		
	0.059***	0.011
	(0.019)	(0.011)
A2. Employment-to-Population Ratio		
	0.012	-0.007
	(0.008)	(0.005)
A3. Unemployment Rate		
	-0.015***	-0.011***
	(0.004)	(0.004)
Panel B: Household Income:		
B1. Log(Median Real Household Income)		
	0.066***	0.015
	(0.014)	(0.012)
B2. Log(Mean Real Household Wage and Salary Income)		
	0.077***	0.012
	(0.017)	(0.014)
B3. Log(Mean Real Rent and Dividend Income)		
	0.107***	0.066**
	(0.035)	(0.026)
Panel C: Population:		
C1. Log(Population)	0.008	-0.019**
	(0.017)	(0.010)
Fracing Exposure Group	Top Quartile	Quartiles 1-3
Control Group	Pscore Matched Sample	Pscore Matched Sample
Play Fixed Effects	Y	Y

Notes: This table reports long-difference regressions of the change in county aggregate economic outcomes between 2000 and 2009/2013 on an indicator for different measures of fracing exposure. The fracing exposure measures are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frac date after 2008. The reported estimates are for the balanced sample. Column (1) measures fracing exposure using an indicator for being in the top quartile of the Rystad max prospectivity measure, and the control group is the pscore-matched sample described in the text. Column (2) reports results where the measure of fracing exposure is an indicator for being in Rystad quartiles one through three and the control group is again the pscore-matched sample. 2013-2009 data come from the American Community Surveys. 2000 data come from the Decenniel Census. Specifications in panels A2, A3 and B1 are weighted by the county population, labor force, and number of households respectively. All income values are converted to 2010 dollars. Standard errors clustered at the county-level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Panels A1, B, and C, Column (1) include observations from 1,411 total counties, of which 65 Rystad top quartile and 558 pscore-matched sample counties are in the balanced sample. Panels A1, B, and C, Column (2) include observations from 1,627 total counties, of which 253 Rystad quartiles one through three and 558 pscore-matched sample counties are in the balanced sample.

Panels A2 and A3, Column (1) include observations from 1,410 county-play combinations, of which 64 Rystad top quartile and 558 pscore-matched counties are in the balanced sample. Panels A2 and A3, Column (2) include observations from 1,627 county-play combinations, of which 253 Rystad quartiles one through three and 558 unique pscore-matched counties are in the balanced sample.

Appendix Table 17—: Impact of fracing on crime

	(1)	(2)
Panel A. Log(Total Crime)		
Top Quartile Effect at τ=5	82.760	288.001*
	(153.177)	(160.348)
Panel B: Log(Violent Crime)		
Top Quartile Effect at τ=5	-12.636	-23.715
	(54.010)	(47.204)
Panel C: Log(Property Crime)		
Top Quartile Effect at τ=5	95.396	311.716*
	(113.958)	(161.042)
Fracing Exposure Group	Top Quartile	Quartiles 1-3
Control Group	Pscore Matched Sample	Pscore Matched Sample
Fracing Exposure Level Shift	Y	Y
Fracing Exposure Trend	Y	Y
Fracing Exposure Trend Break	Y	Y
County Fixed Effects	Y	Y
County-Specific Trends	Y	Y
Year-Play Fixed Effects	Y	Y
Restricted to Balanced Sample	Y	Y

Notes: This table reports regressions of crime rates on different fracing exposure measures. The fracing exposure measures are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frac date after 2008. The reported estimates are for the balanced sample. Crime data come from the FBI Uniform Crime Reporting (UCR) system. Crime reports law enforcement agencies are aggregated to the county level. Data from a law enforcement agency is only included if the agency reports crimes to the FBI UCR system in every year from 1990 to 2013. Both columns allow for pre-trends, a post fracing level shift, and a post fracing trend break in Rystad top quartile counties, as well as county-specific trends. The sample is restricted to the balanced sample of counties. Column (1) reports specifications where the fracing exposure measure is an indicator for being in the Top Quartile of Rystad max prospectivity and the control group is the pscore-matched sample. Column (2) reports versions of the specifications in Column (1), but the exposed group is instead Rystad quartiles one through three. The reported estimates and standard errors correspond to the top quartile level shift coefficient + 5 times the top quartile trend break coefficient. Standard errors clustered at the county level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Column (1) includes 9,570 observations from 638 county-play combinations, of which 56 Rystad top quartile and 359 unique pscore-matched sample counties are in the balanced sample. Column (2)

Appendix Table 18—: Impact of fracing on local government revenues and expenditures: Propensity-score estimates

	(1)	(2)
Panel A: Log(Total Expenditures): 2012 - 2002		
	0.141***	0.013
	(0.036)	(0.021)
A. Log(Direct Expenditures)		
	0.123***	0.000
	(0.036)	(0.021)
A1. Direct Expenditures by Type		
A1a. Log(Current Operating Expenditure):		
[84%]	0.096***	-0.011
	(0.029)	(0.018)
A1b. Log(Capital Outlays):		
[12%]	0.295**	0.114
	(0.143)	(0.081)
A2. Direct Expenditures by Purpose		
A2a. Log(Education Expenditures):		
[48%]	0.043	0.019
	(0.031)	(0.021)
A2b. Log(Public Safety Expenditures):		
[8%]	0.257***	0.062**
	(0.061)	(0.031)
A2c. Log(Welfare and Hospital Expenditures):		
[10%]	0.287**	0.044
	(0.133)	(0.108)
A2d. Log(Infrastructure and Utility Expenditures):		
[18%]	0.284***	0.041
	(0.073)	(0.039)
A2e. Log(Other Expenditures):		
[16%]	0.084	-0.038
	(0.063)	(0.038)
Panel B: Log(Total Revenues): 2012 - 2002		
	0.154***	-0.001
	(0.035)	(0.018)
B1. Revenues by Type		
B1a. Log(Property Tax Revenues):		
[24%]	0.164***	0.031
	(0.044)	(0.021)
B1b. Log(Sales Tax Revenues):		
[4%]	0.509***	-0.093
	(0.116)	(0.073)
B1c. Log(Other Tax Revenues):		
[2%]	0.273*	0.235***
	(0.143)	(0.089)
B1d. Log(Intergovernmental Revenues):		
[42%]	0.126	0.026
	(0.081)	(0.034)
B1e. Log(Charges Revenues):		
[14%]	0.167**	0.073
	(0.079)	(0.053)
B1f. Log(Other Revenues):		
[14%]	0.221***	-0.040
	(0.066)	(0.041)
Panel C: Government Balance Sheets		
C. Net Financial Position as Share of Revenues		
	-0.054	-0.034
	(0.072)	(0.042)
Panel D: Log(Elem/Sec Education Spending per Pupil)		
	0.034	0.034
	(0.035)	(0.022)
Fracing Exposure Group	Top Quartile	Quartiles 1-3
Control Group	Pscore Matched Sample	Pscore Matched Sample
Play Fixed Effects	Y	Y

Notes: This table shows regressions on the change in government spending and revenues between 2002 and 2012 on different indicators measuring fracing exposure. The fracing exposure measures are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frac date after 2008. The reported estimates are for the balanced sample. In Column (1) measures fracing exposure using an indicator for being in the top quartile of the Rystad max prospectivity measure and the control group is the pscore-matched sample described in the text. Column (2) reports results where the measure of fracing exposure is an indicator for being in Rystad quartiles one through three and the control group is again the pscore-matched sample. Data come from the 2012 and 2002 Census of Governments. Panels A1 and B1 show the share of total government revenues or expenditures represented by the given category in brackets below the category name. Both columns report standard errors clustered at the county-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Column (1), Panels A, B, and C include 1,412 county-play combinations, of which 65 Rystad top quartile and 558 unique pscore-matched counties are in the balanced sample. Panel D, Column (1) includes 1,225 county-play combinations with non-missing school enrollment data for all districts in 1997, 2002, and 2012, of which 61 Rystad top quartile and 484 unique pscore-matched counties are in the balanced sample.

Column (2), Panels A, B, and C, include 253 county-play combinations, of which 558 Rystad quartiles one through three and 12 unique pscore-matched counties are in the balanced sample. Panel D, Column (2) includes 244 county-play combinations with non-missing school enrollment data for all districts in 1997, 2002, and 2012, of which 244 Rystad quartiles one through three and 484 unique pscore-matched counties are in the balanced sample.

Appendix Table 19—: Impact of fracing on housing outcomes: Propensity-score estimates

	(1)	(2)
Panel A: House Values		
A1. Median House Value		
	-0.016	-0.036**
	(0.017)	(0.018)
A2. Mean Housing Value		
	0.004	-0.020
	(0.016)	(0.016)
A3. Mobile Housing Units: Median Housing Value		
	0.050	-0.011
	(0.042)	(0.028)
Panel B: Rental Prices		
B1. Median Rental Price		
	-0.008	-0.018*
	(0.015)	(0.010)
B2. Mean Rental Price		
	0.000	-0.020**
	(0.016)	(0.009)
Panel C: Housing Quantities		
C1. Total Housing Units		
	-0.019	-0.030***
	(0.014)	(0.008)
C2. Total Mobile Homes		
	0.039	0.017
	(0.028)	(0.017)
C3. Share of Housing Units Vacant		
	-0.014***	-0.004
	(0.005)	(0.003)
C4. Acres of Agricultural Land		
	-0.137	-0.008
	(0.142)	(0.078)
Fracing Exposure Group	Top Quartile	Quartiles 1-3
Control Group	Pscore Matched Sample	Pscore Matched Sample
Play Fixed Effects	Y	Y

Notes: This Table shows regressions of the change in different housing outcomes between 2000 and 2009-2013 (with the exception of acres of agricultural land, which is measured in 2002 and 2012) on different indicators measuring fracing exposure. The fracing exposure measures are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frac date after 2008. The reported estimates are for the balanced sample. Column (1) measures fracing exposure using an indicator for being in the top quartile of the Rystad max prospectivity measure and the control group is the pscore-matched sample described in the text. Column (2) reports results where the measure of fracing exposure is an indicator for being in Rystad quartiles one through three and the control group is again the pscore-matched sample. 2013-2009 housing data come from the American Community Survey. 2000 Housing data come from the Decenniel Census. 2002 and 2012 agricutlural land data come from the 2002 and 2012 Census of Agriculture respectively. All housing values are converted to 2010 dollars. Observations are weighted by the number of owner (renter) occupied units in the county. Non-mobile specific regressions are adjusted for changing owner (renter) occupied housing characteristics. Housing characteristics included are: fraction of units with 0, 1, 2, 3, or 5 or more bedrooms, fraction of units with full indoor plumbing, fraction of units with a complete kitchen, fraction of units that are mobile units, fraction of units by type of electricity, and fraction of units by age of unit. Both columns report standard errors clustered at the county-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Panels A1, A3, B1, B2, C1, C2, and C3 contain observations from 1,411 county-play combinations, of which 65 Rystad top quartile and 558 unique pscore-matched sample counties are in the balanced sample. Column (1), Panel C4 contains observations from 1,237 county-play combinations, of which 53 Rystad top quartile and 485 unique pscore-matched sample counties are in the balanced sample.

Column (2), Panels A1, A3, B1, B2, C1, C2, and C3 contain observations from 1,627 county-play combinations, of which 253 quartiles one through three and 558 unique pscore-matched sample counties are in the balanced sample. Column (2), Panel C4 contains observations from 1,420 county-play combinations, of which 211 quartiles one through three and 485 unique pscore-matched sample counties are in the balanced sample.

Appendix Table 20—: Welfare estimates under alternative outcome measures and assumptions

	Base Measures		Alternative Measures of Changes in Prices or Quantities					
	Δ in housing costs (rents) = 2.9% Δ in population (total) = 2.7%Δ in wa zges (wage inc/capita) = 7.5% Δ in rent/dividend income = 9.3%		Δ housing costs (home values) = 5.7%		Δ in population (18-55 population) = 5.3%		Δ in wages (wage/salary earnings for full-time/full-year workers) = 4.1%	
	WTP for change in:							
	Amenities (1)	Welfare (2)	Amenities (3)	Welfare (4)	Amenities (5)	Welfare (6)	Amenities (7)	Welfare (8)
Panel A: Base Parameters and Model								
A1. s = .3 and β = .65	-2,225***	1,716***	-1,405*	2,537***	-1,855*	2,086**	-560	1,529***
	(842)	(610)	(734)	(758)	(1,071)	(913)	(520)	(533)
Panel B. Base Parameters and Model with Pre-Trends								
B1. s = .3 and β = .65	-3,076***	848*	-1,288*	2,636***	-2,706**	1,217	-1,314***	792**
	(907)	(458)	(717)	(889)	(1,118)	(814)	(476)	(399)
Panel C: Alternative Parameters and Base Model								
C1. s = .1 and β = .65	-2,466***	1,475**	-1,646**	2,296***	-2,342**	1,599*	-764	1,325***
	(803)	(585)	(695)	(744)	(1,028)	(874)	(495)	(513)
C2. s = .5 and β = .65	-1,980**	1,961***	-1,160	2,782***	-1,360	2,581***	-352	1,737***
	(903)	(666)	(799)	(799)	(1,146)	(991)	(572)	(578)
C3. s = .5 and β = .33	-2,403***	1,538***	-1,987**	1,955***	-1,783	2,158**	-711	1,378***
	(921)	(595)	(861)	(660)	(1,158)	(940)	(544)	(516)

Notes: This table reports estimates of the effect of fracing on amenities and welfare in dollars per household per year. Different rows report values for different assumptions regarding the standard deviation of idiosynchratic preferences and the share of wage and salary income spent on housing or different assumptions about pre-trends. Columns (1) and (2) report results from our base measures, where change in housing costs is measured using the estimated percent change in median rents (.029), the change in wages is measured using the change in wage and salary income per capita (.075), the change in population is measured using the change in total population (.027) and the change in rent and dividend income is measured using the average change in rent and dividends per capita (.093). Columns (3)-(8) report results using alternative measures of changes in prices or quantities. Specifically, Columns (3) and (4) show estimates where the change in housings costs is measured using the estimated percentage change in median home prices (.057), Columns (5) and (6) report results where the change in population is measured using the change in the 18-55 year old population (.054), and Columns (7) and (8) report results where the change in wage and salary income is measured using the change in wage and salary income for full-time, full-year workers (.041). In columns (3)-(8), the base measures are used for all prices and quantities other than the specified quantity. For each measure, we report both the estimated change in amenities (Columns (1), (3), (5), and (7)) and the estimated change in total welfare (Columns (2), (4), (6), and (8)). Panel A uses our preferred parameter values for the standard deviation of household idiosynchratic preferences and the share of income spent on housing of s=.30 and ^{2}= .65 respectively. Panel B reports results from models that also allow for differential pre-trends in high-fracking potential counties. Panel C reports results for alternative parameter values. The calculations are converted to dollars using the mean household wage and salary income in top quartile counties of of $45,668 and mean household interest and dividend income in top quartile counties of $3,822. Panel B aggregates these figures to the total impact of fracing in aggregate welfare in top quartile counties assuming a discount rate of 5 percent, and using the mean number of households in top quartile counties of 25,650 and the total number of top quartile counties of 65. Standard errors incorporate both sampling variance in the estimated parameters, as well as uncertainty in the values of s and ^{2}. For these calculations, the assumed standard deviation of the standard deviation of idiosyncratic preferences (s) is .09 and for the share of income spent on housing (^{2}) is .1. For more information, see discussion in the text.

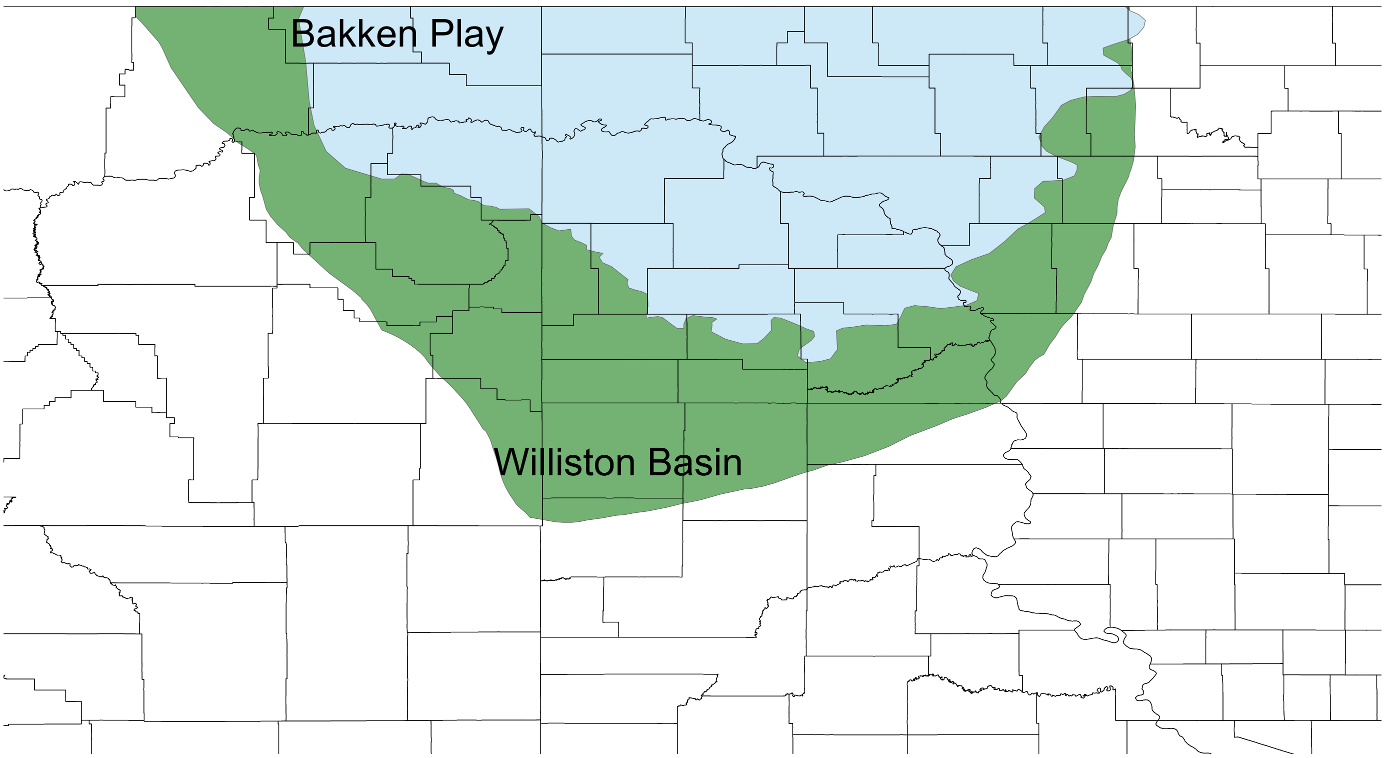
Appendix Table 21—: 2SLS Estimates of the Impact of Hydrocarbon Production on Local Economic Outcomes

	OLS	2SLS
	(1)	(2)
Panel A: First Stage - Instrument 1(Rystad Top Quartile)		
A1. Hydrocarbon Production per Capita (tens of thousands of dollars per household)		
	-	2.220**
	-	(0.886)
Panel B: Endogenous Variable - Hydrocarbon Production (tens of thousands of dollars per hh)		
B1. Log(Mean Home Values)		
	0.002***	0.024*
	(0.001)	(0.012)
B2. Log(Mean Rental Values)		
	0.002***	0.013***
	(0.000)	(0.005)
B3. Log(Wage and salary income per household)		
	0.002***	0.031**
	(0.001)	(0.014)
B4. Log(Rent and Dividend Income per HH)		
	0.009***	0.038**
	(0.002)	(0.016)
B5. Log(Population)		
	0.000	0.017
	(0.000)	(0.017)
Fracing Exposure Group	Top Quartile	Top Quartile
Control Group	Quartiles 1-3	Quartiles 1-3
Play Fixed Effects	Y	Y

Notes: This table reports estimates of the relationhip between hydrocarbon production per household (in tens of thousands of dollars per household) on various economic outcomes. Column (1) reports OLS estimates, while Column (2) reports Two-Stage Least Squares estimates using an indicator for the county being in the top quartile of Rystad prospectivity as an instrument for hydrocarbon production per household. Panel A reports the first-stage results, while Panel B reports the OLS and 2SLS estimates. All estimates come from long-difference regressions (OLS or 2SLS) of the change in county aggregate economic outcomes between 2000 and 2009/2013 on the given RHS variable (an indicator for being in the top quartile of Rystad prospectivity in Panel A and the change in total hydrocarbon production per household between 2000 and 2011 in tens of thosueands of dollars of dollars . The fracking and oil and gas measures are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frac date after 2008. The reported estimates are for the balanced sample. 2013-2009 data come from the American Community Surveys. 2000 data come from the Decenniel Census. Oil and gas production data come from HPDI well data aggregated to the county level. All specifications are weighted by the number of households in the county in 2000. All income values are converted to 2010 dollars. Heteroskedasticity robust tandard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

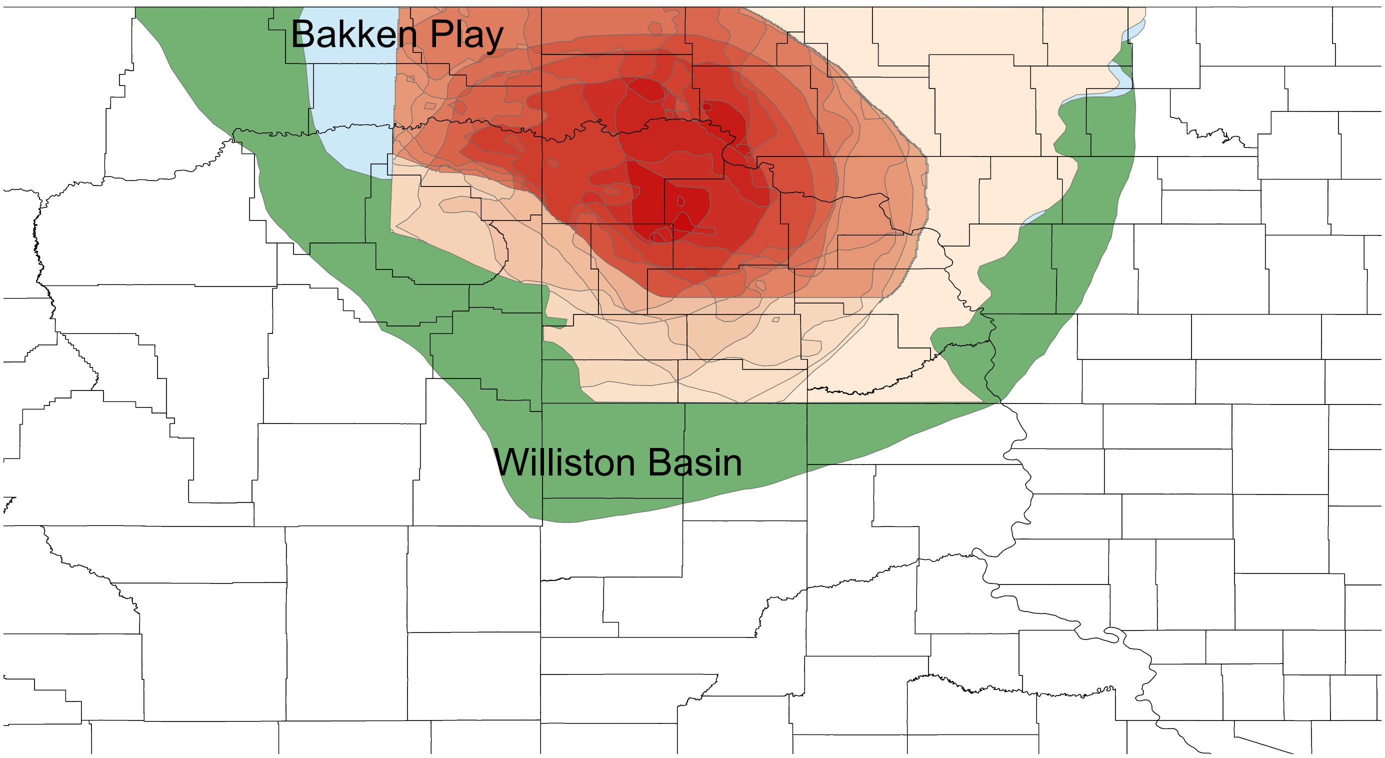
Sample: All Panels contain observations from 403 total counties, of which 64 top quartile counties and 339 Rystad quartiles one through three.

VIII. Appendix Figures



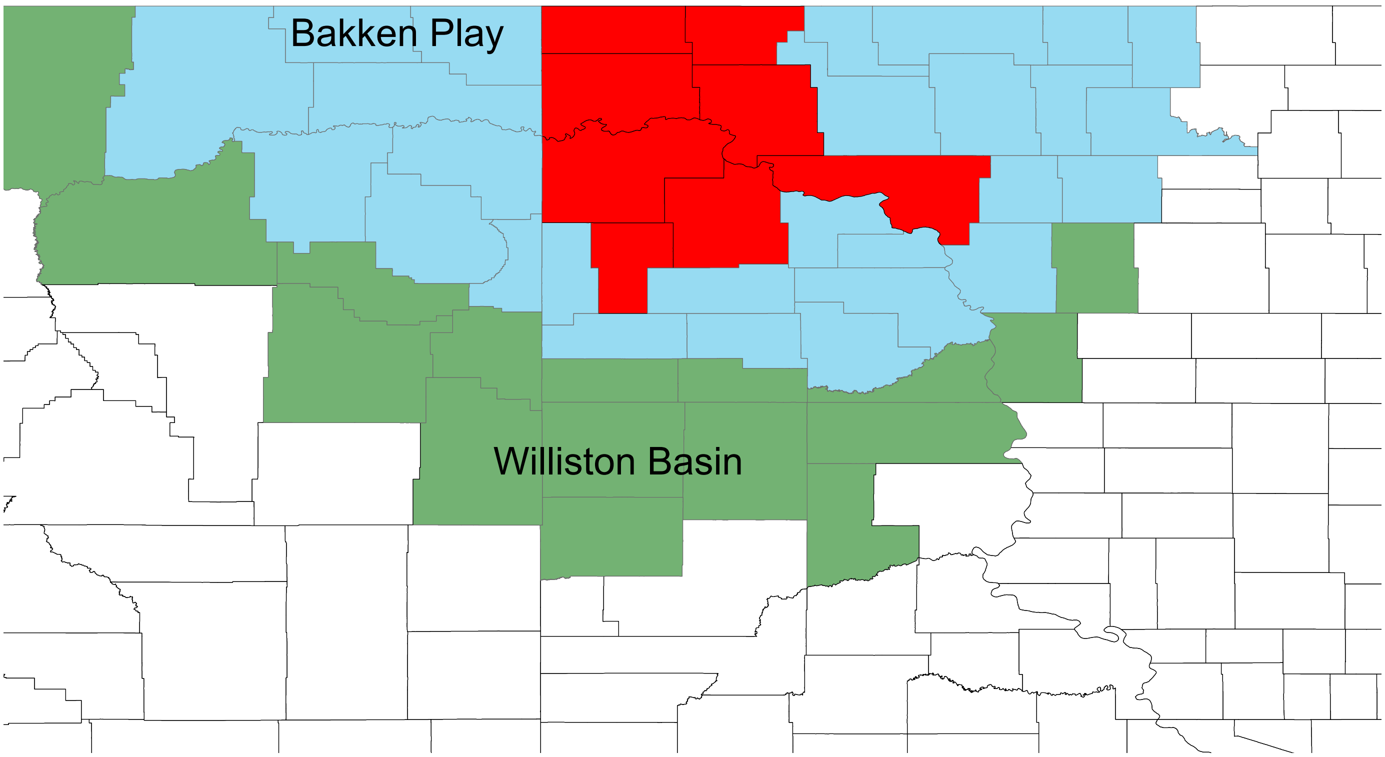
Appendix Figure 1. : Bakken play and Williston Basin

Notes: This figure overlays the Williston Basin and Bakken shale play over nearby counties in North Dakota, Montana, Wyoming, and South Dakota. The Williston Basin is shown in green and the Bakken shale play is shown in blue. Shapefiles for US shale basins and plays comes from the Energy Information Agency (2011).



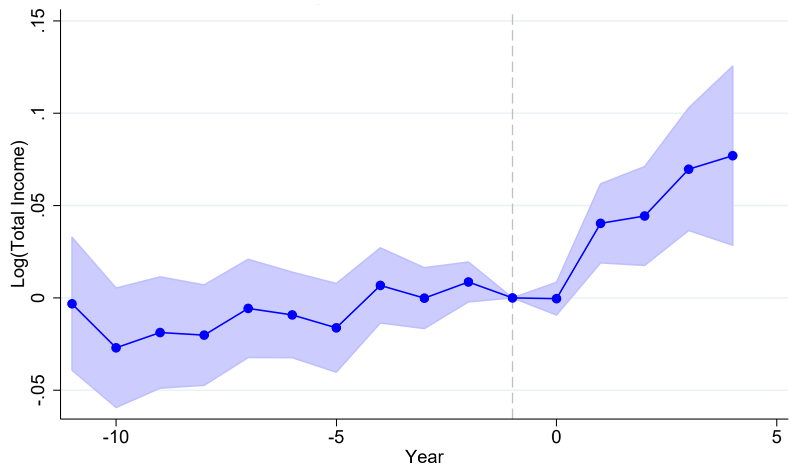
Appendix Figure 2. : Bakken play and Williston Basin with Rystad Prospectivity scores

Notes: This figure overlays the Williston Basin, Bakken shale play, and Rystad Prospectivity scores over nearby counties in North Dakota, Montana, Wyoming, and South Dakota. The Williston Basin is shown in green and the Bakken shale play is shown in blue. Darker shades of red correspond to higher Rystad Prospectivity scores. Shapefiles for US shale basins and plays comes from the Energy Information Agency (2011), while prospectivity scores were purchased from Rystad Energy (2014).



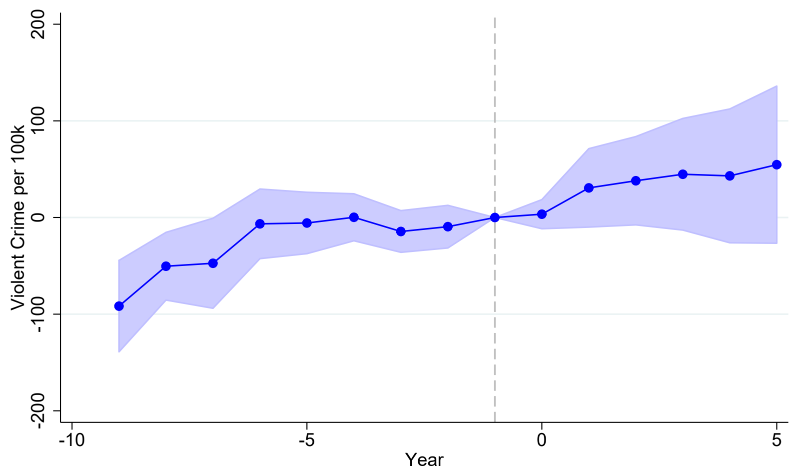
Appendix Figure 3. : Bakken play with county assignments

Notes: This figure shows county-assignments for counties in the area around the Williston Basin. Counties shaded red are in the top quartile of Rystad Prospectivity for the Bakken play. Counties shaded blue are outside of the Rystad top quartile, but within the Bakken shale play boundaries. Counties shaded green are within the Williston Basin, but outside the Bakken shale play and the top quartile of Rystad Prospectivity. Shapefiles for US shale basins and plays comes from the Energy Information Agency (2011), while prospectivity scores were purchased from Rystad Energy (2014).



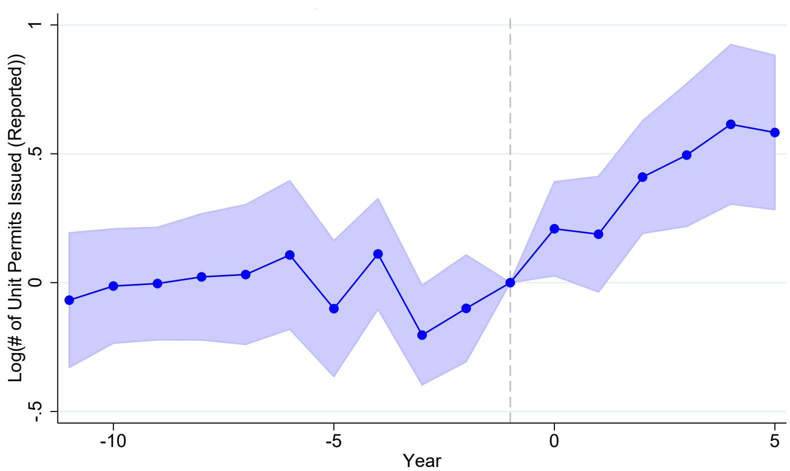
Appendix Figure 4. : Event study analysis of county-level total income

Notes: This figure plots results from an event-study analysis of the difference in log(total income) between high-fracing potential counties and other counties in shale plays before and after fracing began. The reported coefficients come from fitting a modified version of Equation 7 where we interact 1[Rystad Top Quartile]$_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frac year in the relevant shale play. These coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. The model also includes play-year and county fixed effects. All Rystad Top Quartile-event year interactions are interacted with an indicator for being in the unbalanced sample. The reported coefficients correspond to the balanced sample. Consequently, the results in the figure correspond to shale-plays that began fracing in or before 2008 and event-years common to all these shale plays (i.e. event-years observed for all shale plays that began fracing in or before 2008). Data on county-level total income from 1990 to 2012 come from the Local Area Personal Income (LAPI) data from the Regional Economic and Information Systems (REIS) data produced by the US Bureau of Economic Analysis (BEA). Specifically, we use the the variable CA04-10. The shaded blue region shows 95 percent confidence intervals calculated using standard errors clustered at the county level.



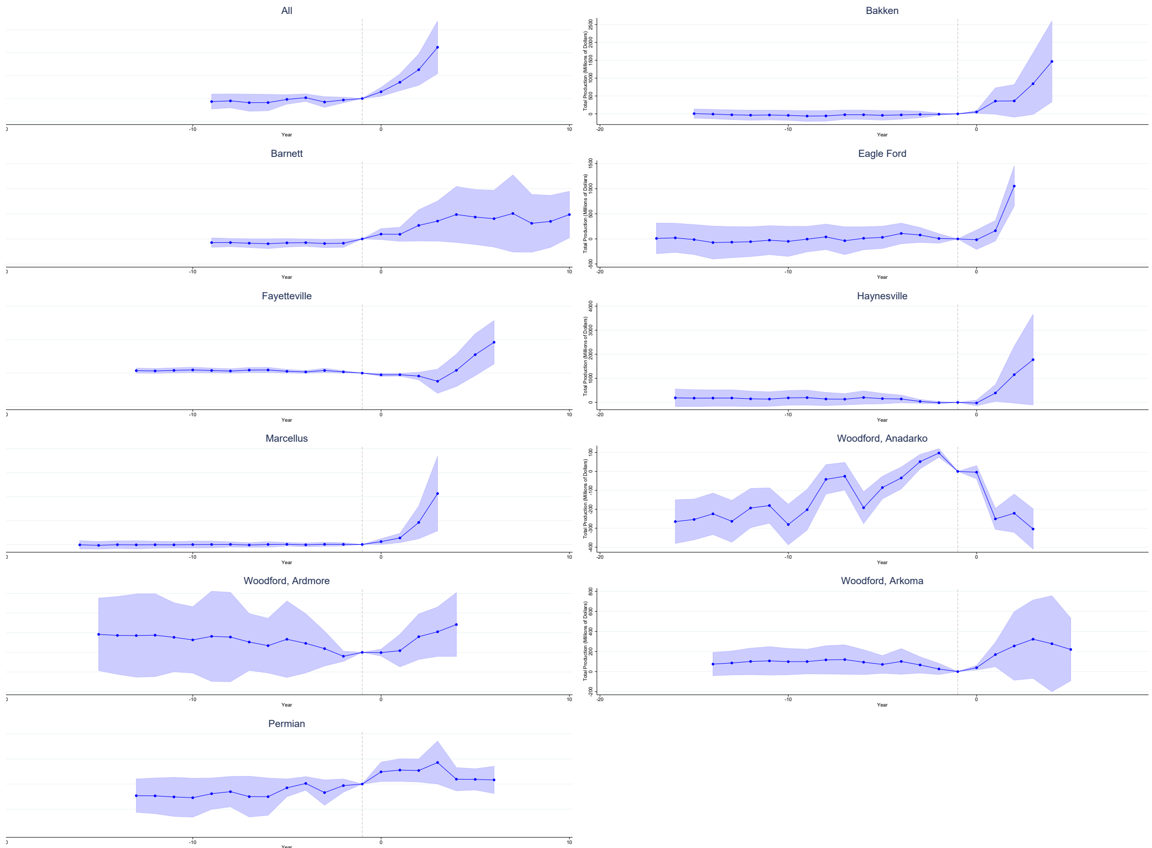
Appendix Figure 5. : Event study analysis of county-level violent crime

Notes: This figure plots results from an event-study analysis of the difference in the county-level violent crimes per 100 thousand residents between high-fracing potential counties and other counties in shale plays before and after fracing began. The reported coefficients come from fitting a modified version of Equation 7 where we interact 1[Rystad Top Quartile]$_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frac year in the relevant shale play. These coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. The model also includes play-year and county fixed effects. All Rystad Top Quartile-event year interactions are interacted with an indicator for being in the unbalanced sample. The reported coefficients correspond to the balanced sample. Consequently, the results in the figure correspond to shale-plays that began fracing in or before 2008 and event-years common to all these shale plays (i.e. event-years observed for all shale plays that began fracing in or before 2008). Data on violent crimes from 1992 to 2013 come from the Uniform Criminal Records (UCR) produced by the Federal Bureau of Investigation (2015). The UCR data consistent of police agency reports of total crime to the FBI in each year. Because participating in the UCR program is voluntary, not all agencies report in all years. To avoid our results being affected by changes in the set of agencies that report crimes to the UCR program, we use the total number of crimes within a county reported by police agencies that consistently report crime in almost all years from 1992 to 2013, interpolating crimes in years when the agencies do not report. Violent crimes include murder, rape, assault, and robbery. The shaded blue region shows 95 percent confidence intervals calculated using standard errors clustered at the county level.



Appendix Figure 6. : Event study analysis of county-level value residential construction permit activity

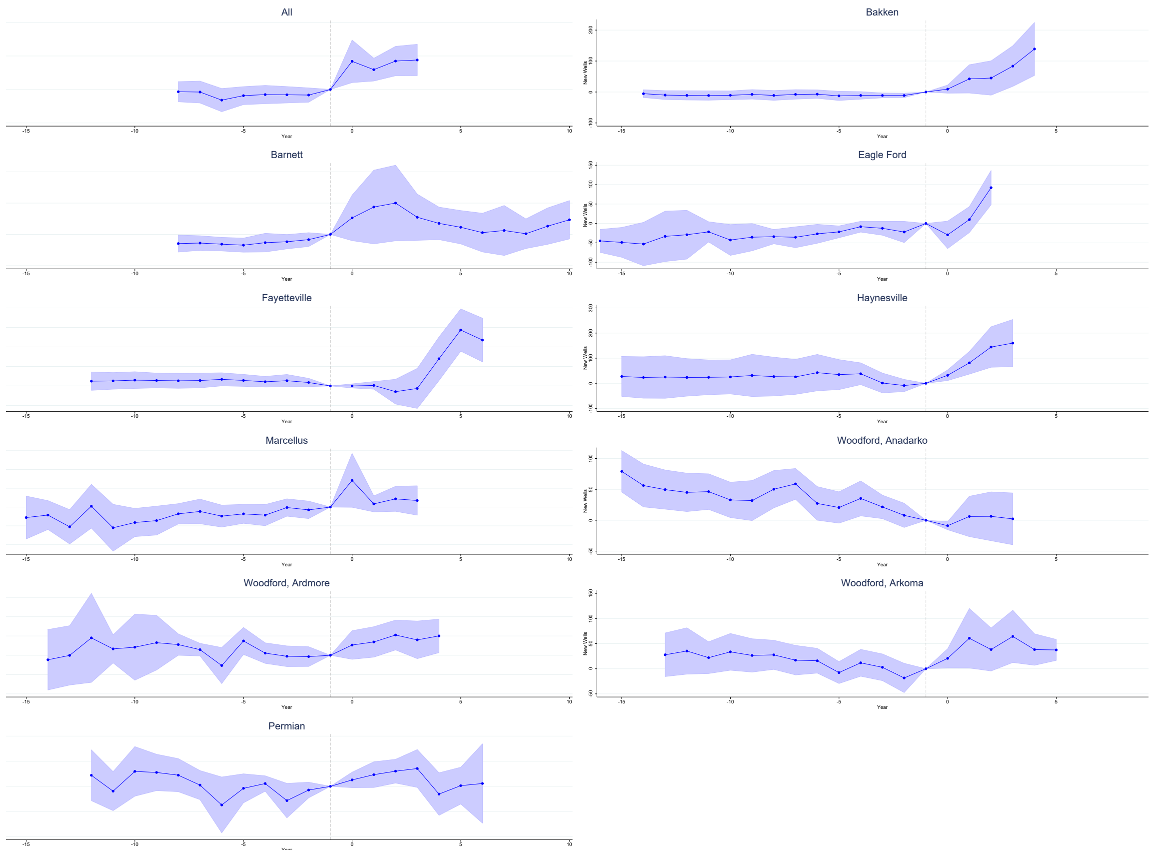
Notes: This figure plots results from an event-study analysis of the difference in the county-level log(residential unit construction permits) between high-fracing potential counties and other counties in shale plays before and after fracing began. The reported coefficients come from fitting a modified version of Equation 7 where we interact 1[Rystad Top Quartile]$_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frac year in the relevant shale play. These coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. The model also includes play-year and county fixed effects. All Rystad Top Quartile-event year interactions are interacted with an indicator for being in the unbalanced sample. The reported coefficients correspond to the balanced sample. Consequently, the results in the figure correspond to shale-plays that began fracing in or before 2008 and event-years common to all these shale plays (i.e. event-years observed for all shale plays that began fracing in or before 2008). Data on residential unit construction permits 1990 to 2013 come from US Census Bureau (2014). The shaded blue region shows 95% confidence intervals calculated using standard errors clustered at the county level.



Appendix Figure 7. : Event study analysis of county-level hydrocarbon production by play

Notes: This figure plots results from an event-study analysis of the difference in the county-level value of hydrocarbon production between high-fracing potential counties and other counties in shale plays before and after fracing began separately by shale play. The top-left figure reports results for all shale plays where fracing began in or before 2008, while the remaining figures report results separately by shale play.[a] The reported coefficients come from fitting a modified version of Equation 7 where we interact 1[Rystad Top Quartile]$_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frac year in the relevant shale play. These coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. The model also includes play-year and county fixed effects. For shale-play specific figures, we also interact the Rystad Top Quartile-event-year dummies with shale-play indicators. Data on hydrocarbon production from 1991 to 2011 come from Drilling Info, Inc (2012). The shaded blue region shows 95 percent confidence intervals calculated using standard errors clustered at the county level.

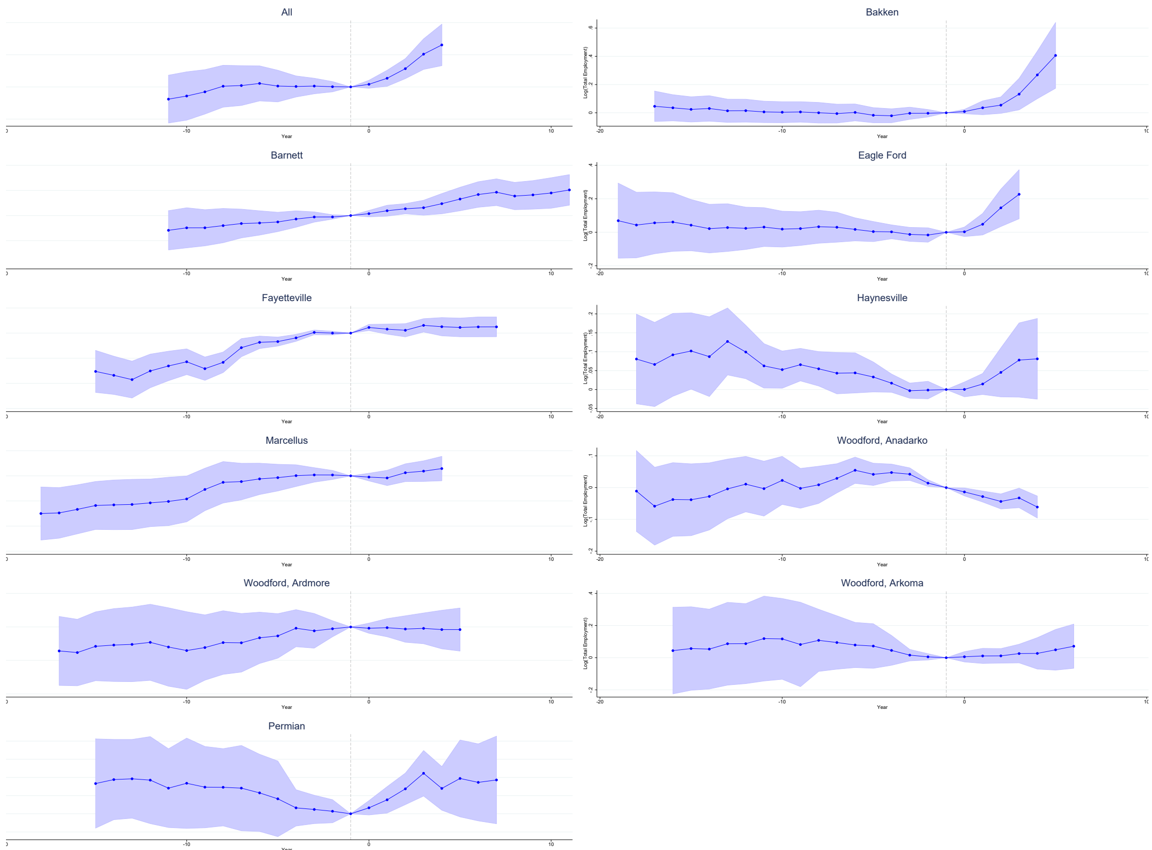
[a]The first-frac year for the Eagle Ford is 2009. Consequently, it does not contribute to the overall results because it does not have at least three years of post-data for all variables. However, because this is an especially prominent play we report the Eagle Ford specific results for reference.



Appendix Figure 8. : Event study analysis of new producing well by play

Notes: This figure plots results from an event-study analysis of the difference in the county-level total number of new wells drilled between high-fracing potential counties and other counties in shale plays before and after fracing began separately by shale play. The top-left figure reports results for all shale plays where fracing began in or before 2008, while the remaining figures report results separately by shale play.[a] The reported coefficients come from fitting a modified version of Equation 7 where we interact 1[Rystad Top Quartile]$_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frac year in the relevant shale play. These coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. The model also includes play-year and county fixed effects. For shale-play specific figures, we also interact the Rystad Top Quartile-event-year dummies with shale-play indicators. Data on new oil and gas wells from 1993 to 2011 come from Drilling Info, Inc (2012). A well is coded as being "new" if in year t if year t is the first year that the well is observed producing oil or gas in the data. The shaded blue region shows 95% confidence intervals calculated using standard errors clustered at the county level.

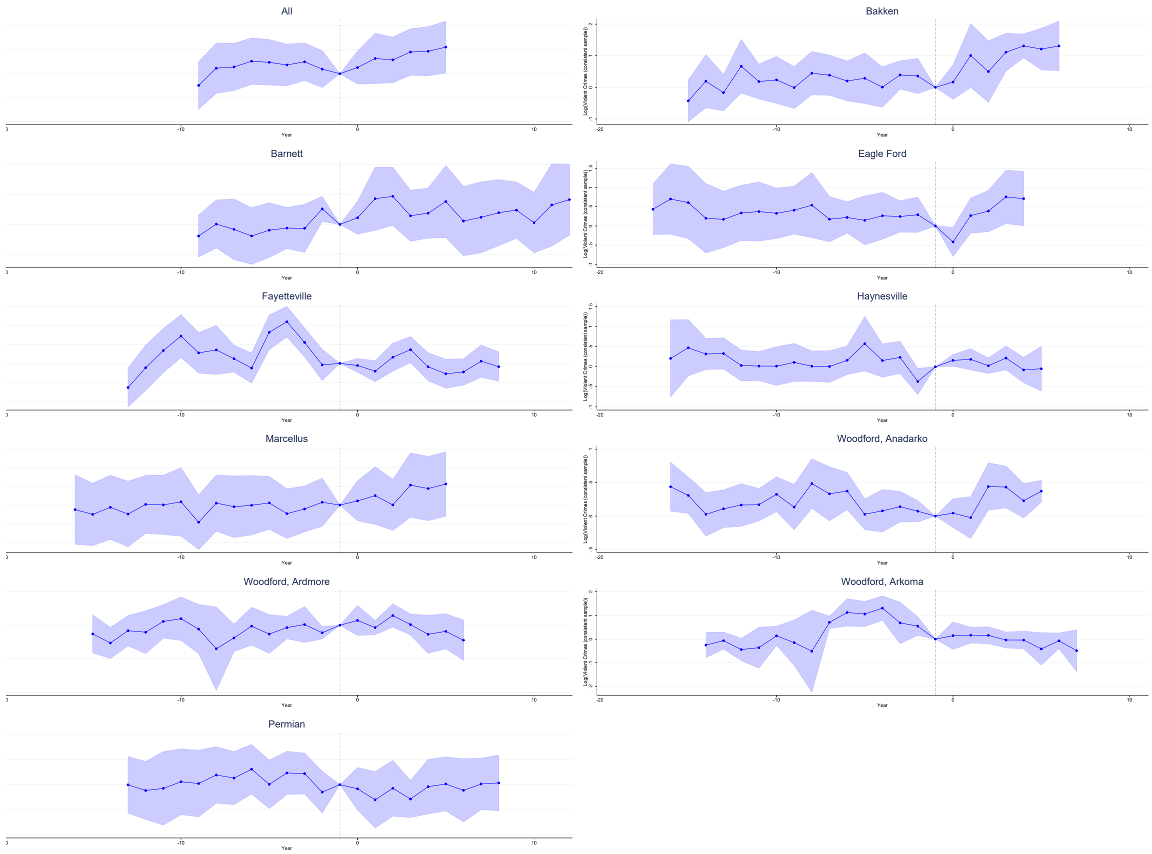
[a]The first-frac year for the Eagle Ford is 2009. Consequently, it does not contribute to the overall results because it does not have at least three years of post-data for all variables. However, because this is an especially prominent play we report the Eagle Ford specific results for reference.



Appendix Figure 9. : Event study analysis of county-level employment by play

Notes: This figure plots results from an event-study analysis of the difference in log(total employment) between high-fracing potential counties and other counties in shale plays before and after fracing began separately by shale play. The top-left figure reports results for all shale plays where fracing began in or before 2008, while the remaining figures report results separately by shale play.[a] The reported coefficients come from fitting a modified version of Equation 7 where we interact 1[Rystad Top Quartile]$_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frac year in the relevant shale play. These coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. The model also includes play-year and county fixed effects. For shale-play specific figures, we also interact the Rystad Top Quartile-event-year dummies with shale-play indicators. Data on county-level total employment from 1990 to 2012 come from the Local Area Personal Income (LAPI) data from the Regional Economic and Information Systems (REIS) data produced by the US Bureau of Economic Analysis (BEA). Specifically, we use the the variable CA25-10. The shaded blue region shows 95% confidence intervals calculated using standard errors clustered at the county level.

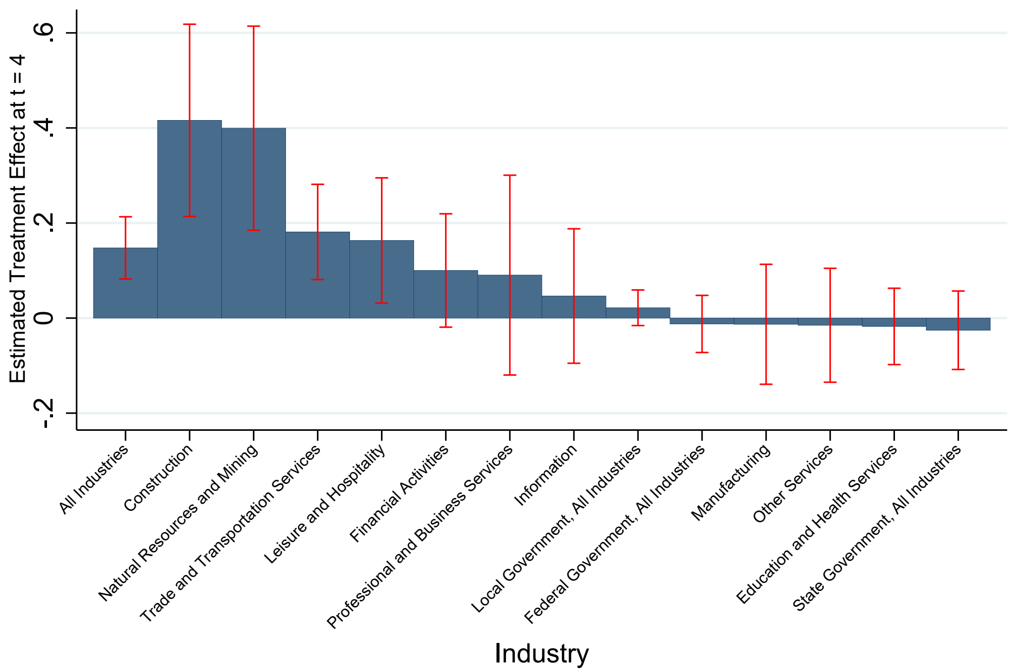
[a]The first-frac year for the Eagle Ford is 2009. Consequently, it does not contribute to the overall results because it does not have at least three years of post-data for all variables. However, because this is an especially prominent play we report the Eagle Ford specific results for reference.



Appendix Figure 10. : Event study analysis of county-level log(violent crime) by play

Notes: This figure plots results from an event-study analysis of the difference in county-level log(violent crimes) between high-fracing potential counties and other counties in shale plays before and after fracing began separately by shale play. The top-left figure reports results for all shale plays where fracing began in or before 2008, while the remaining figures report results separately by shale play.[a] The reported coefficients come from fitting a modified version of Equation 7 where we interact 1[Rystad Top Quartile]$_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frac year in the relevant shale play. These coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. The model also includes play-year and county fixed effects. For shale-play specific figures, we also interact the Rystad Top Quartile-event-year dummies with shale-play indicators. Data on violent crimes from 1992 to 2013 come from the Uniform Criminal Records (UCR) produced by the Federal Bureau of Investigation (2015). The UCR data consistent of police agency reports of total crime to the FBI in each year. Because participating in the UCR program is voluntary, not all agencies report in all years. To avoid our results being affected by changes in the set of agencies that report crimes to the UCR program, we use the total number of crimes within a county reported by police agencies that consistently report crime in almost every year from 1992 to 2013, interpolating crime for years in which the agency does not report. Violent crimes include murder, rape, aggravated assault, and robbery. The shaded blue region shows 95% confidence intervals calculated using standard errors clustered at the county level.

[a]The first-frac year for the Eagle Ford is 2009. Consequently, it does not contribute to the overall results because it does not have at least three years of post-data for all variables. However, because this is an especially prominent play we report the Eagle Ford specific results for reference.



Appendix Figure 11. : Employment effects by industry: pscore-matching

Notes: This figure plots estimates of the effect of fracing on log(employment) by industry five years after the start of fracing. Each bar reports results of fitting Equation 8 for the given industry, which corresponds to Column (4) in the tables. Equation 8 allows for differential pre-trends in event time, as well as a trend break in outcomes and a mean shift for Rystad top-quartile counties. The model also includes play-year and county fixed effects. All Rystad Top Quartile variables are interacted with an indicator for being in the unbalanced sample. The reported estimates correspond to the balanced sample. Data on employment by industry from 1990 to 2013 come from the Quarterly Census of Employment and Wages (QCEW) produced by the Bureau of Labor Statistics, US Department of Labor (2014). Counties are included in the sample if the given employment variable is non-missing in all years from 1990-2013. Red bars report 95 percent confidence intervals calculated using standard errors clustered at the county level.