The Local Economic and Welfare Consequences of Hydraulic Fracturing

(Revised December 2016)

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

February 2016 CEEPR WP 2016-002

The Local Economic and Welfare Consequences of Hydraulic Fracturing

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

Revision of 12/21/16

Abstract

Exploiting geological variation within shale deposits and timing in the initiation of hydraulic fracturing, this paper finds that allowing fracing leads to sharp increases in oil and gas recovery and improvements in a wide set of economic indicators. At the same time, estimated willingness-to-pay (WTP) for the decrease in local amenities (e.g., crime and noise) is roughly equal to -$1000 to -$1,600 per household annually (-1.9% to -3.1% of mean household income). Overall, we estimate that WTP for allowing fracing equals about $1,300 to $1,900 per household annually (2.5% to 3.7%), although there is substantial heterogeneity across shale regions.

---

*Bartik: Department of Economics, MIT, abartik@mit.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 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.

1 Introduction

The discovery of hydraulic fracturing is considered the most important change in the energy sector since the commercialization of nuclear energy in the 1950s. Fracing,[1] as it is known colloquially, has allowed for the recovery of vast quantities of oil and natural gas from shale deposits that were previously believed to be commercially inaccessible. The result is increases in US production of oil and natural gas to levels unimaginable, even five years ago, substantial reductions in energy prices that have greatly aided consumers both domestically and abroad, and fundamentally altered global geopolitics that are likely to benefit the United States (e.g., reducing the power of OPEC and Russia). Further, while the US has been the focus of early fracing activity, large shale deposits of both natural gas and oil exist around the world, posing tremendous challenges to the planet's climate.[2]

Ultimately, access to these energy resources rests on the willingness of the local communities that sit atop these shale deposits to allow fracing within their jurisdictions. On the one hand, the drilling brings royalty payments and economic activity. On the other hand, there are substantial concerns about the impacts on the quality of life, including water, air, and noise pollution, traffic congestion, and crime.[3] Indeed, there has been substantial heterogeneity in communities' reactions with Pennsylvania, Texas, and North Dakota embracing fracing, while other localities, like New York, Vermont, and internationally some countries such as Germany and France, have banned it. However, in making these decisions about allowing fracing, policymakers and their communities have not had systematic evidence on its benefits or costs, and certainly not on net benefits.

This paper empirically characterizes the effects of fracing on local communities across a wide variety of dimensions, including a plausible measure of the net welfare impacts. A challenge for measuring these impacts is that the communities where fracing has taken root differ from other parts of the country both in levels and trends of economic variables. Consequently, we develop an identification strategy that is based on geological variation within shale plays across the US and variation in the timing of the onset of fracing. Specifically, several factors, including thickness, depth, and thermal maturity of the shale deposit, determine the accessibility and quantity of hydrocarbons. Rystad Energy, an international oil and gas consulting company, has created an index of these factors that is a strong predictor of the variation in the application of fracing techniques within US shale deposits. We purchased a GIS file from Rystad that maps this index. Thus, our identification strategy compares counties over shale deposits in the same shale play with

---

[1] Hydraulic fracturing has been abbreviated in a number of ways, including "fracing," "fracking," "frac'ing," and "fraccing." We use "fracing" throughout the paper.

[2] For example, the global supply of natural gas has increased by more than 70 years, based on current consumption levels, and oil reserves have increased by more than 10 years of current consumption (EIA). The newfound abundance of fossil fuels may also have reduced incentives to invest in low carbon energy technologies.

[3] The Environmental Protection Agency (EPA) has devoted an entire website to the issues surrounding fracing. `http://www2.epa.gov/hydraulicfracturing`.

high potential for fracing to counties with lower values.[4] The second source of variation is the difference in the timing of the onset of fracing across shale plays; these differences are also due to geological variation, among other factors. Together, these two sources of variation are the basis for a difference-in-differences-style identification strategy.

There are four primary findings. First, counties with high-fracing potential produce roughly an additional $400 million of oil and natural gas annually three years after the discovery of successful fracing techniques, relative to other counties in the same shale play. Second, these counties experience marked increases in economic activity with gains in total income (4.4 - 6.9 percent), employment (3.6 - 5.4 percent), and salaries (7.6 - 13.0 percent). Further, local governments see substantial increases in revenues (15.5 percent) that are larger than the average increases in expenditures (12.9 percent).

Third, there is evidence of deterioration in the quality of life or total amenities. We find marginally significant estimates of higher violent crime rates, despite a 20 percent increase in public safety expenditures. Building on the work by by Moretti (2011) and Hornbeck and Moretti (2015), who allow for moving costs and elastic housing supply in a Roback (1982) style model, we develop a model that allows us to calculate both the change in welfare and the change in the value of amenities from the reduced form estimates. These calculations suggest that annual willingness-to-pay (WTP) for fracing-induced changes in local amenities are roughly equal to -$1000 to -$1,600 per household annually (i.e., -1.9 to -3.1 percent of mean annual household income).

Fourth, we use the model to develop a measure of the overall change in welfare among households that lived in these communities before fracing's initiation. The expression is a function of the decline in amenities and observed changes in incomes (4.4 -6.9 percent), population (2.7 percent), housing values (5.7 percent), and housing rental rates (2.7 percent).[5] Overall, we estimate that WTP for allowing fracing equals about $1,300 to $1,900 per household annually (2.5 to 3.7 percent of mean household income), although there is substantial heterogeneity across shale regions.

This paper makes several contributions. First, the focus on net welfare consequences provides a broad picture of fracing's overall impacts.[6] Of course, these estimates are only as good as the information on impacts of fracing that households have at their disposal; and as new information emerges about potential health consequences and other impacts, this effect may change.[7] Second,

---

[4] A "shale play" is an area where oil and gas producing firms have targeted a particular shale formations or set of shale formations.

[5] Although it is very demanding of the data, we also estimate play-specific housing prices effects and find estimates on housing prices that range between an increase of 25 percent to no statistically significant change. The largest housing price gains are in the Bakken (primarily in North Dakota and Montana) and the Marcellus (largely in Pennsylvania, West Virginia and Ohio) shale plays.

[6] Due to the use of county-level information on housing prices, this paper is not able to provide a detailed assessment of the distributional consequences of fracing on the housing market. In an important paper, Muehlenbachs et al. (2014a) find that in a sample of roughly 1000 Marcellus region houses, proximity to a fracing site reduces prices by 20 percent for houses that rely on well water, relative to those that utilize piped water. Nor does our paper deal with the more global issue of how fracing affects global greenhouse gas emissions and geopolitics.

[7] The EPA released a preliminary report on a wide-ranging study on the health and environmental risks of fracing

the examination of 9 different shale plays provides near comprehensive measures of the impacts of fracing across the United States.[8] In contrast, much of the previous research has focused on single plays, especially the Marcellus in Pennsylvania (Gopalakrishnan and Klaiber (2013); Muehlenbachs et al. (2014a)).

Third, the paper demonstrates that areas of the country with abundant opportunities for fracing differ from the rest of the country in important ways. As a solution to this identification problem, this paper offers a credible identification strategy based on the geological characteristics of shale deposits. In contrast, we are unaware of any other papers in the rapidly growing literature on fracing's impacts that have both a research design that applies to a wide range of shale plays and address this problem of confounding. Fourth, we have collected data on a wide set of outcomes, ranging from measures of local economic activity to crime to housing market outcomes, which together with the locational equilibrium model that we set out provides a fuller picture of fracing's impacts than has been available previously. In this respect, it expands our understanding of resource booms (see, e.g., Wynveen (2011)), although it does not shed light on the potential for the "Dutch disease" (see, e.g., Allcott and Keniston (2014) and Fetzer (2015) for recent work on this topic) or our understanding of how these effects propagate (see, e.g., Feyrer et al. (2015)). In the most closely related work, Jacobsen (2016) finds that fracing has benefited local communities economically as measured by wages and housing rental rates.

For several reasons, this paper's estimates are likely to be relevant going forward for communities making decisions about whether to allow fracing. First, there are vast shale deposits around the globe that have not yet been accessed due to a mix of legal, institutional, and economic reasons. As some of the non-economic barriers are removed and drilling technologies continue to advance, many jurisdictions will be confronted with decisions about whether to allow fracing.[9] Second, the estimates are based on a period when natural gas prices were historically low, stable, and near current levels. Thus for shale deposits that can be fraced to deliver natural gas, the paper's results are self-evidently relevant. Third, although the paper's results come from a period when petroleum prices were higher than they are currently, petroleum prices have a long history of volatility and multiple "new normals" over the last several decades (Baumeister and Kilian (2016)).

The paper proceeds as follows. Section 2 outlines our conceptual framework. Section 3 discusses hydraulic fracturing and how it differs from conventional oil and natural gas recovery. Section 4 discusses the data used in the analysis, while section 5 describes our identification strategy. Section 6 provides preliminary evidence, our econometric estimates, and the robustness of those results. Section 7 presents evidence of local welfare implications of our estimates. Finally, Section

---

(Environmental Protection Agency, Office of Research and Development (2015)). Regulations also continue to evolve.

[8] We restrict the sample to 9 plays to ensure enough post-fracing data to identify the effects. We come back to this later in the paper.

[9] See Covert et al. (2016) for a discussion of these issues and `http://www.eia.gov/todayinenergy/detail.cfm?id=14431` for a map of world resources.

8 concludes the paper.

2 Conceptual Framework

The aim of the paper is to understand the impacts of fracing on local communities, with an eye toward developing a summary measure of welfare. We follow a stylized model that builds upon the insights of the canonical Roback (1982) model, which is often used as a signpost for assessing the welfare consequences of changes in local amenities (see, e.g., Chay and Greenstone (2005); Greenstone et al. (2010); Kline and Moretti (2015)). The model is a slightly modified version of Moretti (2011) and Hornbeck and Moretti (2015), who incorporate the possibility of moving costs and elastic housing supply into a Roback (1982) style model.[10] The model is explained in detail in Appendix Section A.

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 and rents (which are assumed to be an index for locally produced goods)[11], household wage and salary income, and population $\widehat{\Delta \ln r_t}$, $\widehat{\Delta \ln w_t}$, and $\widehat{\Delta \ln N_t}$ respectively. Using these estimates, and values of the standard-deviation of idiosyncratic location preferences or moving costs, s, and the share of household income spent on housing, 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.[12] Specifically, differentiation of Equation A.4 and re-arrangement yields an expression for household willingness-to-pay for the amenity changes caused by fracing:

$$\Delta \text{WTP for Amenities} = \quad \alpha \Delta \ln A_{at} = \quad s \Delta \ln N_{at} - (\Delta \ln w_{at} - \beta \Delta \ln r_{at}) \tag{2.1}$$

Thus, WTP for the change in amenities, expressed as a percentage of income, is equal to the difference between the change in population, adjusted for the magnitude of moving costs, and the change in real wages.

This is a remarkably useful expression because it provides an estimate of willingness-to-pay for the full set of amenity changes,[13] even though a data set with the complete vector of amenities or

---

[10] The only difference between the model we present here and Hornbeck and Moretti (2015) is that they are focused on the effects of a pure productivity shock, whereas we allow the introduction of fracing to shift both local productivity and amenities.

[11] 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 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.

[12] In the canonical Roback (1982) model, s is effectively assumed to be equal to zero.

[13] A_{at} is the full vector of amenities and α measures the willingness to pay for log-changes in those amenities.

information on willingness-to-pay for these amenities are unlikely to ever be available. The intuition behind this approach comes from the fact that, in spatial equilibrium, the marginal resident must be indifferent to relocating, which means that local housing prices will respond to changes in local wages. The strength of this response will depend on both the elasticity of local housing supply and moving costs.

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.[14] This is the population that has the greatest influence on whether fracing should be allowed in a community. Specifically, let $\bar{W}_a$ be average baseline household wage and salary income, $\bar{Y}_a$ be the average household rental, dividend and interest income, 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 $\bar{R}_a \times \widehat{\Delta \ln r_{at}} + \bar{Y}_a^{\text{owner}} \times \Delta \widehat{\ln y_{at}^{\text{owner}}}$[15]. This expression for WTP is more realistic than the workhorse expression from the canonical Roback (1982) model that is simply equal to the change in property values. Thus, the expression for the total change in welfare for all individuals that either reside or own property in location a before the change in amenities is:

$$\text{WTP for Allowing Fracing} = \Delta\widehat{V_{at}} \approx N_{at} \times \left(\bar{W}_a\widehat{\Delta \ln w_{at}} + \bar{Y}_a \times \widehat{\Delta \ln y_{at}} + \bar{W}_a\alpha\widehat{\Delta \ln A_{at}}\right) \tag{2.2}$$

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.[16]

Nevertheless, this model is still stylized and there are three caveats worth highlighting. First, the model assumes that workers are homogenous, and relaxing this assumption would lead to additional welfare consequences. An especially vulnerable population is workers with skills that are not well-suited for fracing-related employment (e.g., the elderly) who rent homes; this group could experience declines in utility due to continued residence in a jurisdiction that allows fracing and face moving costs that, in principle, could lock them in their current location. Additionally, some homeowners may not own the mineral rights to their homes, meaning that they will not benefit from lease payments even if there is drilling on or near their property. While these benefits obviously

---

[14] 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 average change in household income is attained by original residents, and is not due to high earnings by immigrants. Finally, the expression omits profits of landowners who develop new housing units or rent previously vacant housing units. However, we believe it is the correct expression for WTP for allowing fracing in a community.

[15] Where $\bar{Y}^{\text{owner}}$ is the average interest and dividend income for home-owners.

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

accrue to someone, our estimates of fracing on the change of housing prices will not capture these benefits. Second, the model assumes that households have knowledge of and rational expectations about fracing's impact on all present and future changes in household income and amenities. If households are misinformed or uninformed about current or future changes, then the true welfare impacts of fracing will be more complicated. Of course, as new information about fracing's impacts (e.g., health effects) emerge, then households will update their willingness-to-pay for local amenities. Finally, it must be emphasized that this model provides expressions solely for *local* welfare changes. The model is silent on the many potential regional, national, or global effects of fracing, including reductions in petroleum, natural gas, or electricity prices, ambiguous effects on global warming, adoption of renewable technologies, and changes in geopolitics resulting from America's growing role as a fossil fuel producer.

The below analysis develops estimates of the impacts of allowing fracing on housing prices, as well as on household incomes. In Section 7 we combine the expressions developed in this section with these estimates to develop estimates of the willingness-to-pay for the amenity changes associated with the introduction of fracing and the overall change in household welfare for the people that either resided or owned property in these locations.

3 A Primer on Hydraulic Fracturing and a New Research Design

The development of hydraulic fracturing of shale formations is widely considered the most important change in the energy sector since the commercialization of nuclear energy in the 1950s. It has led to massive increases in North American production of natural gas and petroleum that have disrupted energy markets and geopolitics, and, depending on the rate of innovation in low carbon technologies, has either increased or decreased the probability of disruptive climate change. The new production has also greatly altered local economies and communities in a few short years. In North Dakota, the flaring of methane by-product at that state's more than 8,000 fraced wells can be seen from outer space.[17] While fracing has rocketed across many regions of the United States, technological and political constraints have slowed its adoption around the world although most of the resources are buried in shale formations outside the US. This section provides a brief primer on hydraulic fracturing. It also describes how geological variation in the suitability of shale for drilling within shale plays and variation in the timing of the spread of fracturing techniques across US shale formations provide the basis for a research design. The appendix provides more details.

---

[17] `http://geology.com/articles/oil-fields-from-space/`

3.1 A Primer on Hydraulic Fracturing

3.1.1 A Layman's Description of Conventional and Hydraulic Fracturing Drilling

The traditional approach to gas and oil recovery involves drilling into the earth in search of a "pool." These "pools" exist in permeable reservoir rocks such as limestone or sandstone. The oil and gas migrates to these pools from deeper source rocks (such as shale) where the hydrocarbons were formed. The hydrocarbons migrate until they reach a impermeable "cap" or "seal" rock which traps them.[18] During this process, a drill bit drills through the ground and once the drill bit reaches where the pool is believed to be located (typically 1,000 - 5,000 feet below the surface for an on-shore well), the bit is removed, and casing-pipe is placed into the hole. Once the well is cased, the casing is perforated toward the bottom of the casing so that the deposits, being under pressure, will flow up through the pipe on their own. If the underground pressure is insufficient for the deposits to naturally flow up the pipe, pumping equipment is installed at the bottom of the tubing.

For unconventional wells, drilling often continues to lower depths than are typically reached with conventional wells—sometimes exceeding 10,000 feet and generally significantly below the water table. Once the drill bit nears the shale formation, the bit begins to turn sideways. This point is known as the kick-off point. Drilling continues in a horizontal fashion often for more than 10,000 feet. This portion of the well is then cased and then perforated.[19] Although the pipe is perforated, the deposits do not flow because they are trapped in small pockets within the shale formation and the surrounding rock is not sufficiently permeable to allow the hydrocarbons to flow to the well-head. To break the pockets, a mixture of water, sand, and chemicals is pumped into the well under high pressure. The pressure of the liquid fractures the pockets and the sand keeps them from closing once the pressure is relieved. Once the shale is fractured, the hydrocarbons can escape up through the piping to the surface.

There are noteworthy differences in the economics of conventional and unconventional drilling. A typical conventional well requires an investment of roughly $1 to 3 million to determine whether the resources below the ground can be recovered. Fracing is more expensive with an investment cost of approximately $5 to 8 million per well.[20] There are, however, dramatic differences in the success of these two approaches. Fracing has been dubbed farming for the relative certainty of producing hydrocarbons.

It should not be surprising that the fraced wells account for a rapidly growing share of new wells. Although national data on the number of wells that are fraced are unavailable, we can gain a sense for the emergence of fracing from the share of new wells that are drilled horizontally over shale formations; this share has increased from 0.7 percent in 2000 to 25 percent in 2011 (the year

---

[18] Because shale rocks have such low permeability, they can also be a "seal" rock.

[19] Proper casing also plays the important role of preventing reserves and other chemicals produced during drilling from leaking into the groundwater. There has been substantial debate about the frequency of improper casing.

[20] `https://blogs.siemens.com/measuringsuccess/stories/688/`.

with the most recent data that we have purchased from Drilling Info, Inc (2012)).[21] In part because of this rapid increase in the amount of fracing, the fraction of successful exploratory wells in the US has risen from 41 percent in 2000 to 62 percent in 2010 (EIA, 2014).[22]

3.1.2 Shale Terminology

Throughout the paper, we refer to shale basins and shale plays. A basin is a geological concept that refers to a region where geological forces have caused the rock layers to form roughly a bowl shape, where the central part is deeper than the outside portions, with the center then filled in by layers of sediment. If one of the layers is a shale layer, the basin can sometimes be referred to as a "shale" basin. Note that a basin can contain many different rock layers and formations, and that in a "shale" basin, many of the rock layers will not be shale.

A shale play is a region of a shale basin where oil and gas producing firms have targeted a specific formation or group of formations that exhibit similar geological and drilling characteristics. Importantly, the definition of a shale play often depends on where drilling has occurred or may occur. For example, a widely used 2011 Energy Information Administration map[23] defined shale plays by drawing a line around the parts of shale formations with the highest density of wells. Additionally, a shale play usually refers to one formation (for example, the Marcellus shale), while shale basins often contain several different shale formations. For example, the Appalachian Basin contains both the Marcellus shale and the Utica shale, which overlap for much of their extent but at different depths.

3.1.3 Local Impacts of Hydraulic Fracturing Activity

Shale deposits are located in a relatively small number of communities and, as the bans on fracing in multiple jurisdictions indicate, these communities ultimately determine access to the resources. As the numbers at the end of the previous section underscore, unconventional drilling has produced substantial economic value. This paper will develop measures of the economic benefits to local communities in terms of hydrocarbon production, employment, income, net migration, etc. However, these benefits come bundled with a number of impacts that are less desirable. The claimed negative impacts include water and air pollution, increased traffic, crime, and damage to otherwise largely unperturbed physical environments (see e.g. Environmental Protection Agency, Office of Research and Development (2015), Phillips (2014), Ground Water Protection Council and ALL Consulting (2009), National Energy Technology Labratory (2013), Rubinstein and Mahani (2015)).

---

[21] The fraction of wells that are drilled over shale formations has increased from 41 percent in 2000 to 48 percent in 2011, while the share of horizontal wells has increased from 1.7 percent in 2000 to 33 percent in 2011.

[22] This improvement in the success rate for exploratory wells cannot be entirely attributed to fracing, as advances in 3D-imaging have also reduced dry holes for conventional wells.

[23] `http://www.eia.gov/analysis/studies/usshalegas/` We, as well as much of the growing economics literature on fracing, use this map to define the boundaries of shale plays.

The below analysis will measure as many of these local impacts as is possible with available data sources. Ultimately, they cannot be all measured and even if they could their net impact on social welfare is unknowable. As the conceptual framework outlines, we develop estimates of the WTP for the total change in amenities and the net welfare impacts of allowing fracing in the community.

3.2 A New Research Design

This paper's empirical analysis aims to determine the consequences for a local community of allowing fracing. The empirical challenge is to identify a valid counterfactual for jurisdictions that allowed fracing. That is, it is necessary to identify jurisdictions that are identical, except for the presence of fracing; otherwise the empirical analysis may confound fracing with the other differences across jurisdictions. The difficulty is that places with fracing activity may differ from those that do not for a variety of reasons that also affect key outcomes. Places that have a more extensive history of oil and gas development, a lower value of land, or different local economic shocks may be more likely to experience fracing.

The growing fracing literature has relied on a variety of identification strategies. Perhaps, the most widely used one is to compare areas over shale formations to areas without shale formations underneath them (see e.g., Cascio and Narayan (2015); Fetzer (2015); Maniloff and Mastromonaco (2014); Weber (2012); Weinstein (2014)). As we demonstrate below, these places differ on many dimensions in both levels and trends undermining the validity of this approach. Others have taken advantage of a border discontinuity design, based on comparing border areas in Pennsylvania where fracing has been embraced versus New York where it has been banned (Boslett et al. (2015)). This design may be appealing for reasons of internal validity, but its results are specific to just one of the more than ten shale plays in the country, leaving important questions of external validity unanswered.[24]

As an alternative to these approaches, this paper's identification strategy is based on differences in geology within shale plays and the rate at which the basic principles of hydraulic fracturing were successfully applied across US shale formations. The remainder of this subsection describes these two sources of variation that underlie our difference-in-differences-style research design.

3.2.1 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, as well as the thermal maturity,

---

[24] As our results below highlight, there is substantial heterogeneity across shale plays in the effects of fracing.

porosity, permeability, clay content, and total organic content of the local shale rock (Zagorski et al. (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.[25][26]

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. In practice, 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. The appendix illustrates in greater detail how the Rystad prospectivity measure was used to assign counties into top quartile and the bottom three quartiles.

3.2.2 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

---

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

[26] See Budzik (2013) for a general discussion of the role played by different geological characteristics in determining the effectiveness of fracing. Zagorski et al. (2012) describes the geological features of the Marcellus and their role in drilling productivity, Covert (2014) 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.

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. The appendix provides more details on the development of the dates and the implications for identification.

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.

3.3 Alternative Identification Strategies

While our identification strategy provides a plausible control group for top-quartile counties, there are two potential shortcomings of this design. The first is that our strategy does not yield estimates of the impact of fracing in counties other than the top quartile. Second, and related, if fracing has local economic effects on non-top quartile counties, our estimates of the impacts on top-quartile counties will be biased. This might occur for two reasons. First, counties in physical proximity to top quartile counties may benefit from an increase in drilling activity because of either economies of geographic scope associated with drilling, or because these counties themselves have deposits of newly economically accessible hydrocarbons. Second, many of the economic outcomes that we measure can increase as a result of an increase in nearby drilling activity. For example, workers may commute to nearby counties, or workers living in top-quartile counties may travel and spend money in nearby counties.

In principle, these two shortcomings could be overcome by matching all counties in shale plays to counties that are outside of shale plays. Following the procedure in Imbens and Rubin (2015), we used propensity score matching to match counties within shale plays to counties outside shale plays. The appendix describes this matching strategy in more detail. However, as we show in Table 2, this matching strategy was unable to provide plausible control counties, especially for the housing-price measures. This creates a tension between developing a comprehensive measure of fracing's impacts and what can be estimated credibly. Because matching does not appear to be a solution to the confounding problem here, this paper focuses on estimate the impact of fracing on top-quartile counties, but we note that this is likely an underestimate of the full impacts of fracing across the United States.

4 Data Sources and Summary Statistics

The analysis is conducted with the most comprehensive and detailed data set ever assembled on fracing and its consequences. Clearly, it would be impossible to estimate the effects of fracing on every potential outcome; however, we collected data on a large set of effects and will use these results to estimate the net welfare effects of fracing. This section briefly describes the data sources, with more details provided in the Data Appendix. It then provides some evidence on the validity of the research design.

4.1 Data Sources

4.1.1 Fracing Data

Shapefiles of the locations of shale plays and basins, as well as historic oil and gas prices, come from the Energy Information Agency (EIA).[27] Oil and gas production data for 1992 through 2011 come from data purchased from Drilling Info, Inc (2012). The research design depends on the prospectivity estimates from Rystad Energy's NASMaps product purchased from Rystad Energy (Rystad Energy (2014)).

4.1.2 Economic Outcomes

We measure the effect of fracing on a variety of county-level economic outcomes. The Bureau of Economic Analysis' Regional Economic and Information Systems (REIS) data are the source for data on total employment and total annual earnings by type (US Bureau of Economic Analysis (BEA) (2014)). These data are complemented by the Quarterly Census of Employment and Wages' (QCEW) data on wages by industry (Bureau of Labor Statistics, US Department of Labor (2014)).

Housing price data for 2009 through 2013 come from the American Community Survey (ACS), while housing price data for previous decades (2000 and 1990), as well as data on the total number of housing units, come from the decennial Census.[28] In some of our specifications, we also draw on economic data from the decennial Census and 2009 - 2013 pooled ACS, including employment, per capita income, population, and population broken down by age and sex.[29] The 2009 - 2013 ACS data need to be pooled to precisely estimate average county outcomes, so, for a given county, these data

---

[27] For oil prices we use the Cushing, OK, spot price for West Texas Intermediate (Energy Information Agency (2011)) and for natural gas we use the city-gate price. Shapefiles for the boundaries of shale plays and basins come from the EIAs Maps: Exploration, Resources, Reserves, and Production site (Energy Information Agency (2011)).

[28] Alternatives to Census data on housing outcomes do exist, such as Zillow or RealtyTrac data. However, for many of the counties affected by fracing, these data are either missing or interpolated. In addition, these data would not have information on rental markets.

[29] All Census and ACS data were retrieved from the National Historical Geographical Information System (Minnesota Population Center (2011)).

are treated as a single year's observation.[30] Housing permit data come from the Census Bureau's New Residential Construction data-series (US Census Bureau (2014a)). Monetary variables are inflation adjusted using the Consumer Price Index (CPI) produced by the BLS (Bureau of Labor Statistics, U.S Department of Labor (2015)).

Migration data come from the Internal Revenue Service's county-county migration dataset, released as part of the Statistics on Income (Internal Revenue Service (2015)).

4.1.3 Crime

Crime data come from the Federal Bureau of Investigation (2015) Uniform Crime Reporting program (UCR). Individual police 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. Reporting is non-mandatory,[31] and consequently not all agencies report all index crimes in all years. To prevent within-county sample composition changes over time from influencing our results, we define a consistently reporting series using agencies that report[32] index crimes in most years[33] from 1992 through 2013. To ensure that the consistently reporting agencies are representative of the county as a whole, we only include counties in our sample if the consistent sample agencies account for at least 20 percent of total crimes in a given county between 2011 and 2013.[34] 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.

4.1.4 Public Finance

Data on local government spending and revenues come from the Census of Governments conducted every 5 years (years ending in 2 and 7) by the US Census Bureau (US Census Bureau (2014b)). We aggregate direct expenditures and revenues to the county level by summing the values for all local governments within the county. These outcomes are inflation adjusted using the same CPI as

---

[30] The Census Bureau suppresses data for many counties in the 1-year and 3-year ACS releases. Data from very few counties are suppressed in the 5-year ACS estimates.

[31] Some federal grants are conditioned on reporting UCR data, so there is an incentive to report.

[32] Some agencies report crime for only a few months in some years, while others report 0 crime in some years despite covering a large population and reporting high levels of crime in other years, while still others report some crime types but not others. We discuss how we handle these and other misreporting or insufficient reporting in the appendix.

[33] To avoid throwing out data from agencies that report crime in all years except for one or two, we interpolate each crime type for an agency in year t if the agency reports the given crime type in year $t+1$ and $t-1$ and the crime type is missing for the agency for no more than three years from 1990 to 2013. The consistent sample is then agencies for which we have either a reported or an interpolated crime value for each crime type in every year from 1992 to 2013.

[34] Unfortunately, a few counties do not have any agencies that report crimes in most years, 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 Rystad top-quartile counties and 405 total counties in the full sample.

above. We supplement these data using school district-level enrollment data from the Common Core (National Center for Education Statistics (2015)), which allow us to create measures of spending per pupil. Specifically, for all counties in which every school district reports enrollment data in 1997, 2002, and 2012[35] we total county-level primary and secondary enrollment and divide elementary and secondary direct expenditures from the Census of Governments by this enrollment number to compute spending per pupil.

4.2 Summary Statistics

Column (1) of Table 2 reports on the county-level means of key variables. Panel A reports on the values of these variables in 2000, which predates the widespread development of fracing shale plays with horizontal wells in all areas of the US, while Panel B reports on the change between 2000 and 1990. The entries in the first column are intended to provide a sense of the economic magnitude of the differences in means between pairs of counties that are reported in the remaining columns. These comparisons provide an opportunity to gauge the credibility of the paper's quasi-experimental research design, as well as alternative potential designs. Because the crime data have many more missing observations than the data for the other variables, we perform this exercise separately for the crime and non-crime variables. We first discuss the non-crime variables and then the crime variables.

Column (2) compares counties over shale basins with counties across the remainder of the United States and finds that there are important differences between these two sets of counties. Counties within a shale basin have worse economic outcomes; for example, per capita income in 2000 is almost 30 percent (0.279 natural log points) lower in these counties. Indeed, 9 of the 10 reported variables are statistically (and economically) different between the two sets of counties. This is summarized by the p-value of 0.00 associated with the F-test for the hypothesis that the differences in the 10 variables are jointly equal to zero. Panel B reveals that shale basin counties were growing more slowly than the rest of the country from 1990 to 2000; just as in Panel A, 9 of the 10 variables would be judged to be statistically different across the two sets of counties by conventional criteria. Overall, the results in column (2) cast doubt on the validity of a difference-in-difference specification that is based on comparing shale basin counties with the rest of the United States which has become a prevalent identification strategy in the literature.

Column (3) explores the validity of an alternative identification strategy that compares counties in shale plays versus the remaining counties in the same shale basin but not necessarily in the same shale play. [36] (Recall that basins are larger than plays in general.) The differences in income levels

---

[35] We don't use 2007 data because we estimate long-difference models of the change in public finance outcomes between 2002 and 2012. We include 1997 data because, in Appendix Table 12, we also report the robustness of our results to estimating long-difference models of the change between 1997 and 2012.

[36] The entries report the results from regressions of the variable in the row against an indicator for whether the

and income changes are even larger than in column (2), and across the 10 variables there are again statistically and economically large differences between these other two sets of counties. The entries suggest that this comparison is also unlikely to be the basis for a credible quasi-experiment.

In contrast, the entries in column (4) support the validity of this paper's identification strategy that relies on comparing changes in counties within a play that have a Rystad prospectivity measure in the top quartile to the other counties within the same play.[37] A comparison of pre-treatment levels and trends finds little evidence of differences in these two categories of counties. For example, the large differences in levels and trends of housing values and per capita income in columns (2) and (3) are not evident, either statistically or economically, in these two sets of counties. More broadly, the null of equality of the reported variables cannot be rejected in either levels or trends.[38]

The last two columns compare top-quartile counties and non-top quartile counties to their p-score matching counterparts. Column (5) shows that the p-score technique performs well for top-quartile counties in terms of statistical significance. However, a number of the differences are large in magnitude. Column (6) matches quartiles 1 through 3. Here, the matching perform significantly worse; all but hydrocarbon production are statistically different across the two groups.

Turning to the crime variables and pre-trends in Panels A2 and B2, we can see in column (2) that there are large differences in levels of crime, but only small differences in trends, in counties within shale basins compared to the rest of the US. In particular, counties within shale basins have lower levels of violent and property crime. Column (3) shows that comparing counties within shale plays to other counties within the same shale basin reduces the magnitude of the difference between crime levels in Panel A2 markedly, but actually increases the magnitude of the differences in crime trends. Column (4) shows that when comparing Rystad top-quartile counties to other counties within the same shale play, we cannot reject the joint null that property and violent crime do not differ between top-quartile and other shale play counties in either levels or trends. However, it must be noted that the estimated difference in trends for property crime is statistically significant. Furthermore, for both trend variables, our point estimates are large and positive, and the standard errors are extremely large, meaning that we cannot rule out quite large pre-trends in crime in

---

county is in a shale play, an indicator for whether a county is in a shale play interacted for an indicator for whether the shale play is in the balanced sample of shale plays, and basin fixed effects on the subset of counties in shale basins. The coefficient and standard error associated with the shale play indicator are reported in the table and are based on the balanced sample of counties.

[37] The entries report the results from regressions of the variable in the row against an indicator for whether the county has landmass with a top-quartile Rystad prospectivity score, this Rystad top quartile indicator interacted with an indictor for whether the shale play that lays under the county is in the balanced sample of shale plays, and play fixed effects on the subset of counties in plays. The coefficient and standard error associated with the top-quartile indicator are reported in the table and are based on the balanced sample of counties.

[38] Interestingly, one of the few variables that remains different in levels across all columns is total hydrocarbon production. This is not too surprising because shale formations were often source or seal rocks for conventional hydrocarbon production. Consequently, some locations with high potential for fracing also had high potential for earlier, conventional production. Reassuringly, these differences are dramatically reduced when we look at trends in hydrocarbon production, which are not economically or statistically significantly different between top quartile and other counties within shale plays.

top-quartile counties. Consequently, our crime results must be interpreted cautiously.

Although the column (4) results fail to undermine the validity of contrasting these two sets of counties, all reported specifications will control for all permanent differences between them. Further, we will also report on some specifications that adjust for county-specific time trends. The next section discusses the estimation details.

Finally, we turn to the matching comparisons. Each of the shale play county groups exhibit statistically significantly lower crime rates compared to their p-score matching counterparts. In terms of pre-trends, the comparison of top-quartile counties performs somewhat well, although the F-statistic is significant at conventional levels. As with levels, the non-top quartile counties are significantly different from their p-score matching control group. These findings suggest that the pscore-matching procedure is not successful in generating an adequate match for counties exposed to fracing.

5 Empirical Strategy

This section describes the paper's two approaches to implementing the research design based on variation in geology within shale plays and timing in when fracing techniques were adapted to individual plays. Depending on whether the economic variable of interest is measured annually or decennially, we estimate difference-in-differences and long-difference specifications.

5.1 Estimation: Time-Series Difference-in-Differences

When annual data are available, we estimate the following equation for outcome variable y_{cpt}, where the subscripts refer to county (c), shale play (p), and year (t):

$$\begin{aligned} y_{cpt} = \quad & \mu_{pt} + \gamma_c \\ & + \delta\Big(1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c\Big) + \epsilon_{cpt}. \end{aligned} \tag{5.1}$$

The specification includes year-by-play, μ_{pt}, and county fixed effects, γ_c. The two key covariates are: 1) $1[\text{Post Fracing}]_{pt}$, which is an indicator that equals 1 in the year that fracing is initiated in shale play p and remains 1 for all subsequent years;[39] 2) $1[\text{Rystad Top Quartile}]_c$ is an indicator for whether the maximum prospectivity value within county c is in the top quartile for counties in shale play p. The model is fit on the sample of counties that intersect at least one of the 14 US shale plays listed in Table 1.

The parameter of interest, δ, is a difference-in-differences estimator of the effect of fracing. It measures the change in the difference in y_{cpt} between high and low Rystad prospectivity counties

---

[39] This variable equals one for all counties that intersect a shale play after its first-frac date.

within shale plays, after fracing was initiated, relative to before its initiation. Two limitations to this approach are that δ could confound any treatment effect with differential pre-trends in the Rystad top-quartile counties[40] and that it assumes that fracing only affects the level of economic activity, rather than the growth rate. With respect to the latter issue, the possibility of adjustment costs, as well as capital and labor frictions, means that the effect of fracing on economic and other outcomes may evolve over time in ways that a pure mean shift model fails to capture.

Additionally, we fit event study-style versions of equation (5.1), where the indicator variable, $1[\text{Post Fracing}]_{pt}$, is replaced by a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year (e.g., 2006) minus the first-frac year in the relevant shale play. In the subsequent analysis, we plot the coefficients associated with the interaction of this vector and $1[\text{Rystad Top Quartile}]_c$; these coefficients measure the difference in outcomes between top-quartile and other counties within a play, by event years. These figures provide an opportunity to visually assess whether differential pre-trends pose a challenge to causal inference and examine the evolution of the treatment effect over time.

We also estimate a richer specification that directly confronts these two potential shortcomings of equation (5.1). Specifically, we estimate:

$$\begin{aligned} y_{cpt} = \quad & \mu_{pt} + \gamma_c \\ & + \beta_1(\tau_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) \\ & + \delta_0(1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) \\ & + \delta_1(\tau_{pt} \cdot 1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) + \epsilon_{cpt}. \end{aligned} \tag{5.2}$$

This model allows for differential pre-trends in event time for Rystad top-quartile counties, which are captured by the parameter β_1. Moreover, it allows for a trend break in outcomes, δ_1, as well as a mean shift, δ_0. Thus, the estimated effect of fracing τ years after the start of fracing is then $\delta_0 + \delta_1 \times \tau$. Finally, we will also report on models where we include trends in the calendar year t that are allowed to vary at the county level.[41]

Details about the variance-covariance matrix are also noteworthy. First, several of the outcome variables, for example mean housing prices, are county-level estimates. Observations on counties with values estimated on a smaller sample will mechanically have error terms with higher variance. To account for this heteroskedasticity, we weight the equations for these outcomes with the square root of the sample size used to compute the value (e.g., the total number of owner occupied housing

---

[40] Although we are not able to reject the joint null hypothesis there are no overall differences in pre-trends between Rystad top-quartile and other counties for all of our outcome variables, a few important outcomes, such as income and employment, exhibit economically large pre-trends. Allowing for differential pre-trends reduces concerns that these pre-trends in income and employment are biasing our results.

[41] The variable $\tau_{pt} \cdot 1[\text{Rystad Top Quartile}]_c$ is collinear with the county-specific time trends, so that variable is dropped in these specifications.

units for the county-level mean housing price).[42] Second, the reported standard errors are clustered at the county level to allow for arbitrary serial correlation in residuals from the same county. Third, there may be spatial correlation between the error terms in nearby counties. In the robustness Tables 4 and A8 we report Conley standard errors in brackets under the first row, which allow for spatial correlation in the error terms between nearby counties. We discuss these results in more detail in Section 6.4.

Finally, it is important to underscore that the variation in the year of development across shale plays has implications for estimation. In particular, there are differences in the number of pre- and post-fracing years across shale plays, including some that have none or very few post-fracing years. To avoid introducing compositional bias in the estimation of the treatment effects, we focus estimation on a balanced sample throughout the analysis; this sample is restricted to county-year observations with corresponding event years that range from -11 through 3, 4, or 5 (depending on the data source), from the 9 shale plays with first-frac dates that occur in 2008 or before. The subsequent analysis reports both treatment effects that are estimated using all available data and treatment effects where the sample is restricted to the balanced sample. In the former sample, the years outside the balanced sample contribute to the identification of the county fixed effects.[43] Among these 9 shale plays, there are a total of 65 top-quartile counties and 310 counties outside the top quartile.[44] We report estimates of fracing's impact on outcomes evaluated 3, 4, or 5 years (depending on the data source) after fracing's initiation from this balanced sample.

---

[42] The variables for which we implement this weighted least squares approach are: mean housing prices, median housing prices, mean rents, median rents, mean mobile home rental price, mean mobile home value, salary income per worker, income-per-capita, median household income, employment-to-population ratio, unemployment rate, sex by age population shares, manufacturing employment share, and mining employment share.

[43] The unbalanced sample is comprised of observations from shale plays with first-frac dates after 2008 and observations from shale plays with first-frac dates before 2009, for the years corresponding to less than -11 or -10 years or greater than 3, 4, or 5 years (depending on the data source) in event time. In practice, the models are estimated on the full sample so, for example, the specification corresponding to equation (5.2) takes the following form to ensure that the treatment effects are identified from the balanced sample only:

$$\begin{aligned} y_{cpt} = \quad & \mu_{pt} + \gamma_c + \beta_1 \tau \cdot 1[\text{Rystad Top Quartile}]_c \\ & + \beta_2 (1[\text{Unbalanced Sample}]_{ct} \cdot \tau \cdot 1[\text{Rystad Top Quartile}]_c) \\ & + \delta_0 (1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) \\ & + \delta_1 (\tau \cdot 1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) \\ & + \delta_2 (1[\text{Unbalanced Sample}]_{ct} \cdot 1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) \\ & + \delta_3 (1[\text{Unbalanced Sample}]_{ct} \cdot \tau \cdot 1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) \\ & + \epsilon_{cpt}. \end{aligned} \tag{5.3}$$

The reported estimate of the treatment effects is then based on δ_0 and δ_1.

[44] For outcomes with annual data, we restrict the sample to counties with non-missing data in all years since 1990 (1992 for the drilling variables). For some variables, this reduces the sample size slightly.

5.2 Estimation: Long-Differences

For a number of outcomes, such as housing values, population, and demographic variables, well-measured county-year level data are not available nationally. For these outcomes, we turn to the Decennial Census and the American Community Survey (ACS) to estimate long-difference models using the pooled 2009 - 2013 ACS as the post-period and 2000 decennial census as the pre-period.[45] The long difference specification may be especially appealing in the case of housing prices: as discussed in Section 3.2.2, asset prices very quickly reflect information about the future, so with annual housing data assigning a first fracing data after information about fracing potential was known would lead to an understatement of the effect on housing prices. Consequently, a long-difference specification, where the first year of the period is before fracing information is available anywhere in the country and the last year is after our estimated first fracing date for the shale play in our sample where fracing arrived last, is likely to solve this problem. Our estimating equation is derived by first differencing equation (5.2), which gives:

$$y_{cp,2013/09} - y_{cp,2000} = \gamma_p + \delta(1[\text{Post Fracing}]_{pt} \cdot 1[\text{Rystad Top Quartile}]_c) + \epsilon_{cpt}. \tag{5.4}$$

The parameter δ is a difference-in-differences mean shift estimate of the effect of fracing and maps directly to δ in equation (5.2).

Three details about the long-difference approach are worth noting. First, the below event-study graphs suggest that fracing increases the growth rate of many economic variables, rather than simply affecting their levels. Thus, for many economic variables, such as income or total housing units, we might expect the difference-in-differences estimator to understate the impact of fracing several years after its initiation. This concern is ameliorated in the case of asset prices, such as house prices, that may rapidly incorporate the expected future impact of fracing. Second, the long-difference approach is unable to adjust the estimates for differences in pre-existing trends in outcomes between the top-quartile and other counties within a play. Third, we expect that the initiation of fracing will affect the quality of the housing stock, in addition to the price of land, so specifications for prices and rents adjust for housing characteristics of both rental and owner-occupied housing units. Appendix Section E.3 describes which housing characteristics we use in more detail.

---

[45] For long-difference results using the Census of Governments or the Census of Agriculture, the post-year is 2012 and the pre-year is 2002.

6 Results

6.1 Oil and Natural Gas Production Effects

The analysis begins with Figure 4, which is derived from the event-study regression for the total value of hydrocarbon production, measured in millions of dollars. There is little evidence of a trend in hydrocarbon production in advance of the successful application of fracing techniques in the top-quartile counties, relative to the other counties. Additionally, the figure makes clear that following the initiation of fracing, the average top-quartile Rystad county experiences a significant gain in the value of hydrocarbon production, increasing by more than \$400 million from year $\tau = -1$ to year $\tau = 3$.

Table 3 more parsimoniously summarizes the findings from Figure 4. It reports the results from three alternative specifications, each building upon the previous specification. The column (1) specification includes county and year-by-play fixed effects and reports the mean increase in oil and gas production in the post-fracing years. Column (2) allows for differential pre-fracing event time trends in top-quartile counties and then includes a term to test whether these potentially differential top-quartile trends change after fracing is initiated. Column (3) makes two changes, relative to Column (2); it restricts the data file to the balanced sample described above and replaces the top-quartile, pre-fracing event time trend variable with county-specific calendar time trend variables. The bottom of the table reports the estimated treatment effect from each of these models three years after fracing begins.

It is apparent that the initiation of fracing led to substantial increases in hydrocarbon production in top-quartile Rystad counties. The column (1) estimate that does not allow for a trend break suggests that fracing increases the value of production by about \$242 million per year in top-quartile counties. Columns (2) and (3) confirm the visual impression that the change in hydrocarbon production is better characterized by a specification that allows for a trend break, rather than only a mean shift; these specifications suggest that hydrocarbon production was about \$410 million higher in each county three years after the initiation of fracing in top-quartile counties. To put this estimated effect into context, the median population in top-quartile counties prior to fracing activity is about 22,000, indicating an increase of hydrocarbon production of roughly \$19,000 per capita.

6.2 Labor Market and Amenity Effects

Figures 5 and G.4 are event study plots of county-level natural log of total employment and total income for Rystad top-quartile counties, respectively, after adjustment for county and play-by-year fixed effects. Both total employment and total income increase substantially in top-quartile counties following fracing's initiation. Additionally, there is evidence of positive pre-trends for both outcomes, especially for income. These graphs suggest that the more reliable specifications for these outcomes will allow for differential pre-trends and a trend break post-initiation of fracing.

Table 4 reports the results of estimating the same three specifications used in Table 3 for a series of measures of local economic activity and population flows. For reasons of brevity, the table only reports the estimated treatment effect 4 years after the initiation of fracing, rather than the fuller set of individual regression parameters reported in Table 3. Panels A and B are derived from the REIS data file and report on total employment, total income, and income subcategories, while Panel C uses the Internal Revenue Service (IRS) county-county migration flows data.[46]

Panels A and B indicate that Rystad top-quartile counties experience sharp improvements in economic activity after the initiation of fracing, relative to other counties in the same play. In the more reliable specifications presented in columns (2) and (3) specifications, the estimates indicate increases in employment of about 4.9 - 5.4 percent. The income results reveal gains of 4.4 - 6.9 percent that are driven by increases in wages/salaries and rents/dividends (this includes royalty payments from natural resource extraction). The migration results in Panel C are not stable across specification but qualitatively point to modest increases in net migration.

Table 5 reports on tests of the robustness of these results by fitting the long difference-in-differences specification with data from the 2009-2013 American Community Survey and 2000 Census of Population and Housing. This specification is most comparable to the column (1) specification in Table 4, because it is not possible to adjust for differential pre-trends with just two years of data per county. However, the qualitative conclusions about economic activity are unchanged from the trend-break specification described above, as the estimates in Panels A and B suggest a 4.8 percent increase in employment, 2.6 percentage point gain in the employment to population ratio, 0.6 percentage point decline in the unemployment rate, and 5.8 percent rise in mean household income.[47] Finally, Panel C indicates that there was 2.7 percent increase in population although this is only statistically significant at the 10 percent level.

We next turn to the QCEW data to obtain a more nuanced picture of the changes in the local labor market. Figure 6 plots the implied treatment effect four years after fracing begins in Rystad top-quartile counties, along with 95-percent confidence intervals. Across all industries, the estimates indicate that employment increases by an average of roughly 10 percent and this would be judged to be statistically significant by conventional criteria. This is larger than the 4 - 5 percent increase in employment in Tables 4 and 5, but the QCEW assigns employment to a county based on the place of work, not the place of residence as is the case for the data files used in Tables 4 and 5.[48] Natural resources and mining is the industry with the largest increase in employment, more than 40 percent.

---

[46] The IRS data track county-to-county migration flows using the addresses of income tax filers.

[47] The estimate for median household income is an increase of 6.0 percent with a standard error of 1.2 percent.

[48] Furthermore, we use QCEW data through 2013, whereas we only use REIS data through 2012, which one might also expect to decrease the estimated employment effect using REIS data if the effect of fracing on employment is increasing over time.

There are also statistically significant increases in employment in construction and transportation. No industry has a decline that would be judged to be statistically significant.[49]

Hydraulic fracturing is also likely to lead to changes in the composition of the workforce and population, because many of the jobs associated with fracing are held by men in their 20s and 30s. The increase in demand for these workers may lead to in-migration of young males, but could also lead to out-migration of other age groups and women. Appendix Table 3 explores how the population's demographics change. While many of the estimates are imprecise, we find some evidence of an increase in the share of prime-age males and a decrease in the non-working aged population (both young and old). The Panel C results indicate that there is an increase in the share of people with college degrees, perhaps underscoring the sophistication of these drilling operations.

There is a close connection between the labor market and criminal activity and there have been several media reports suggesting that fracing is associated with increases in crime rates.[50] Furthermore, as we see in Appendix Table 3, fracing is associated with increases in the population share of prime age males, which some evidence suggests may result in higher crime rates (for example, see Edlund et al. (2013)). We investigate this possibility with the FBI Uniform Crime Reporting program data, which is the most comprehensive, standardized data available on crime rates. Figure G.5 shows the event-study plot for log violent crime. The estimates are imprecise, but are suggestive of an increase in violent crime. Panels A, B, and C of Table 6 report the results of the same three specifications used in Tables 3 and 4 for log total-crime, log violent crime, and log property crime respectively. Consistent with Figure G.5, the estimates for violent crime are positive across all three geological-based specifications, but imprecise.

Finally, we note that we attempted to measure whether air quality in top-quartile counties was affected by fracing-related activity. The EPA air pollution monitoring network is sparse in the countries covered by shale plays and it was not possible to develop reliable estimates. Even when using the air quality measure with the broadest coverage,[51] only 13 of 65 top quartile counties and 66 of 370 shale play counties have non-missing data in all years between 2000 to 2011.

---

[49] Interestingly, despite the large estimated increase in wage and salary income in Table 4, which we would expect would make manufacturing firms less competitive in fracing counties, the estimated change in manufacturing employment is very small. There are a few possible explanations for this finding. One is that, given capital adjustments costs and other frictions, any effect on manufacturing may appear only a number of years after fracing starts. Alternatively, lower natural gas prices may help keep local manufacturers competitive despite the rise in wages. Fetzer (2015) proposes this channel and finds evidence consistent with lower natural gas prices being an important mechanism in keeping manufacturing in fracing counties.

[50] `http://geology.com/articles/oil-fields-from-space/`.

[51] Average Total Suspended Particulate Matter (TSP), imputed using PM10 or PM2.5 when TSP is not available.

6.3 Local Public Finance

The influx of hydraulic fracturing may also lead to changes in the composition and levels of local government's public finances, specifically revenues and expenditures, in ways that affect public well-being. Table 7 reports the estimated treatment effects for local government expenditures and revenues, based on the fitting of equation 5.2. The estimates suggest that fracing is largely budget neutral; county-wide local government expenditures increase by 12.9 percent, while revenues increase by 15.5 percent. The specific sources of the increases in expenditures and revenues follow intuitive patterns. We estimate that public safety expenditures increase by about 20 percent, infrastructure and utility expenditures went up by roughly 24 percent, and welfare and hospital expenditures increased by about 24 percent, too (although this increase would not be judged statistically significant by conventional criteria). Interestingly, we only find a small, and noisily estimated, 2.5 percent increase in education expenditures. Looking at Panel D, which reports the change in log elementary and secondary education per pupil, we see that spending per pupil is virtually unchanged. The increase in total revenues is largely a result of increases in property tax revenues of 13 percent and other revenues of 26 percent. Panel C reveals that the overall financial position (i.e., debt minus cash and securities as a percentage of annual revenue) of local governments in top-quartile counties is essentially unchanged.[52]

Overall, the Table 7 results indicate that fracing leads to important changes in the character of local governments. Most obviously, these governments grow in size as the local economies grow. On the spending side, many of the new public resources are devoted to infrastructure investments with much of this spending likely aimed at accommodating and/or supporting the new economic activity. The increase in expenditures on public safety is telling and underscores that a full accounting of the impact on crime must include this additional effort to prevent crime. Put another way, the full effect of fracing on crime includes both the potential increase in criminal activity described above and the increase in resources devoted to preventing crime.[53] A topic of considerable interest is whether public education spending is affected and the available evidence suggests that the rise in local revenues does not lead to higher per pupil school spending. Finally, it is noteworthy that the net financial position of governments in top-quartile counties appears unchanged.[54]

---

[52] Appendix Table 12 reports long difference results using 1997 as the base-year instead of 2002 (our first-frac date for the Barnett is in late 2001, so in theory the 2002 local public finance outcomes could already have incorporated some of the effect of fracing). The results for local government spending and revenues are qualitatively unchanged when using 1997 as the pre-year instead of 2002. Appendix Table 11 reports on the impacts of fracing on local government employment and payroll.

[53] This is analogous to Deschenes et al. (2012) which demonstrates that the full welfare effects of a reduction in air pollution include changes in health outcomes and expenditures on medicines that protect individuals' health from exposure to air pollution.

[54] This is consistent with recent case-study evidence from Newell and Raimi (2015), although they find important heterogeneity across municipalities.

6.4 Robustness

We gauge the robustness of the results to alternative definitions of fracing exposure and approaches to controlling for local economic shocks. Panels A and B of Table 4 and Panel B of Appendix Table 8 report on these exercises for hydrocarbon production, employment, and income, respectively. Column (1) reports the results from fitting specifications that were used in column (2) of Tables 3 and 4. Column (2) adds state-by-year fixed effects to the column (1) specification. Column (3) returns to the specification in column (1), but here the balanced sample is defined to include shale plays that have at least two years of post data for all outcome variables (rather than three years) although the treatment effect is still reported at $\tau = 3$. In practice, this allows the Eagle Ford shale play to contribute to the reported treatment effects. All three columns use the same sample used throughout the paper.

The entries in the rows of each Panel report on alternative definitions of counties that are highly amenable to fracing. The first row repeats the definition that we have utilized throughout the paper. That is, a county must have some land area with a Rystad prospectivity score that is in the top quartile for its shale play. For the entries in this row, we report standard errors clustered at the county-level (in parentheses) as is done throughout the rest of the paper and standard errors that allow for spatial correlation (in square brackets) in the error terms (Conley (1999)).[55] The next two rows alter the definition so that it is based on land area with a Rystad score in the top tercile and quartile, respectively. Rows 4-6 base the definition on the mean value of the Rystad prospectivity score across all of a county's land area, using the top quartile, tercile, and octile, respectively.

The Panel A results suggest that the conclusions about the effect of fracing on hydrocarbon production are qualitatively unchanged by these alternative approaches. It is reassuring that the estimated effect is increasing in the stringency of the indicator definition for fracing amenability in the cases of both the maximum- and mean-based definitions. Further, the estimates are larger for the maximum-based definition. The standard errors tend to be larger with the Conley assumptions about the variance-covariance matrix, but these assumptions do not appreciably affect the statistical significance of the results. Additionally, the estimates are essentially unchanged by replacing the play-year fixed effects with the state by year ones. Finally, it is noteworthy that the estimated effects in column (3) are modestly larger, reflecting the Eagle Ford's boom in petroleum production since 2009.

The results in Panel B broadly support the conclusions from the preferred results in Table 4. They are qualitatively unchanged by the use of state by year fixed effects or allowing the Eagle Ford to influence the estimated treatment effect. When the maximum Rystad prospectivity score is

---

[55] To implement Conley standard errors, we use code from Hsiang (2010). We compute the centroids of counties using GIS software and allow for spatial correlation between counties whose centroids fall within 200 km of a given county. Nearby counties are uniformly weighted until the cutoff distance is reached. These standard errors also allow for serial correlation in the error terms of a given county.

used, fracing is estimated to increase total income by 6 - 9 percent and the effect would be judged statistically significant by conventional criteria in 7 of the 9 specifications. When the mean Rystad prospectivity score is used, the estimated effects tend to be smaller and statistically insignificant, although the 95 percent confidence intervals overlap the analogous intervals associated with the maximum based variables.[56] Panel B of Appendix Table 8 reveals that the employment-based results have the same pattern in that the estimated effects tend to be larger with the maximum-based definitions of a county's suitability for fracing. The broader lesson here seems to be that even within shale plays, the economic benefits of fracing are concentrated in the subset of counties that are most suitable for drilling, although the imprecision of the estimates makes definitive conclusions unwarranted.[57]

An issue that is related to the question of the robustness of the estimated treatment effects is the degree of spillovers between top-quartile counties and other counties in the same play. The full local effects of fracing include these spillovers, which may involve individuals living in a non-top-quartile county but working in one and the resulting knock-on effects in their home county. If there are fixed local costs of drilling, neighboring counties might also experience increases in hydrocarbon production; for example, it is costly to move rigs and other infrastructure long distances.

While these effects are likely real and cause the paper's estimates to understate the full local economic benefits, our identification strategy is not well suited to measure them. The ideal experiment would provide random variation in the suitability of fracing in adjoining, either geographically or economically, counties. Since our empirical approach rests on comparing different sets of counties, both of which sit atop the same shale formation, this violates the ideal.

6.5 Heterogeneity Across Shale Plays

Our empirical design also allows us to estimate play-specific effects from fracing. We report on the 9 shale plays included in the pooled results. Additionally, we also include the Eagle Ford shale play although fracing began there in 2009 which is beyond the cutoff for our pooled results; however, the Eagle Ford, located in the southern part of Texas, has attracted a lot of attention.

The 10 event study plots for hydrocarbon production (Figure G.7) suggest that in 9 of the shale plays, hydrocarbon production in top-quartile counties prior to fracing was largely flat and then took off after the commencement of fracing. The lone exception is the Woodford Anadarko play, which for largely idiosyncratic reasons experienced an increase in production in advance of fracing and decline afterwards.[58]

---

[56] Panel A of Appendix Table 8 reports estimates from the same specifications for total wage and salary income and also suggests that the results for this outcome are robust.

[57] The number of top-quartile counties with the maximum- and mean-based definitions are 65 and 75, respectively. The analogous numbers of counties for the octile variables are 32 and 39, and 88 and 102 for the tercile ones.

[58] Two factors explain the patterns in the Woodford Anadarko. First, there is only one top-quartile county in the Anadarko play. Therefore, we are essentially measuring how this county compares to the rest of the play. Conse-

Table 9 reports the econometric results across the ten shale plays. Here, we focus on three outcomes: hydrocarbon production, wage and salary income, and housing prices.[59] Column (1) reproduces the overall estimate for the relevant outcome from previous tables. The play-specific estimates are in columns (2) through (10) and the Eagle Ford estimates are in column (12). Column (11) reports the F-statistic and associated p-value from a test that the 9 shale estimates in columns (2) through (10) are equal. The Eagle Ford is not included in the F-test or in the overall estimates for Column (1). Although it is demanding to estimate shale-specific treatment effects, this exercise is still able to produce results with substantial empirical content.

As suggested by the event study graphs, we estimate large increases in hydrocarbon production in 9 of the 10 plays; the estimates are statistically significant in 6 of the 9. Similarly, we estimate sizable increases in income per household in 7 of 10 plays; the estimates would be judged statistically significant by conventional criteria for 4 of the 7. In contrast, the gains in housing prices appear to be concentrated in two of the 10 plays. Specifically, the house price gains in the Bakken and Marcellus shale plays—the two shale plays that have generally received the most media attention—are 23 percent and 9 percent, respectively.

It is noteworthy that we can reject the null of equal effects for all three outcome variables. With only 10 observations, it is difficult to make precise statements about the sources of the observed heterogeneity. However, we note that the estimated effects on income are (weakly) positively correlated with the hydrocarbon effect (0.11), positively correlated with the share of oil production (0.49), and negatively correlated with pre-fracing population (-0.30).[60] That is, places with large changes in hydrocarbon and small baseline populations experience larger labor demand shifts and have fewer workers in other sectors who can switch into oil and gas production, increasing the impacts on incomes. It is not surprising that the results are imprecise for higher population plays, because there is less statistical power to detect reasonable effect sizes for aggregate outcomes in these areas; further, it seems reasonable to expect smaller effect sizes in heavily populated areas with larger economies..

---

quently, even if top-quartile counties are expected to have much more fracing than others, with only one draw there is a non-trivial probability that the top-quartile county will not have higher hydrocarbon production. Second, the Anadarko play had considerable conventional drilling activity prior to hydraulic fracturing. Therefore, our estimation conflates the decline in conventional production and the increase in fracing, possibly beginning as a response to the reduction in conventional production. See, for example, http://www.ogj.com/articles/print/volume-93/issue-10/in-this-issue/exploration/partial-us-oil-gas-resource-volumes-termed-39astonishing39.html.

[59] Given the substantial heterogeneity suggested by these results, it is also interesting to explore whether this heterogeneity extends to other outcomes. Appendix Table 5 reports play-specific results for a broad set of additional hydrocarbon, labor market, quality of life, and housing variables. The results also show substantial heterogeneity on these dimensions, and like our other results, suggest that the effects of fracing on the Bakken have been much larger than the effects on other plays.

[60] The estimated housing price effects also show this pattern, although the correlations are weaker—0.06 for hydrocarbon production and -0.18 for population.

7 Interpretation and Local Welfare Consequences of Fracing

What are the net local welfare consequences of fracing? To this point, the paper has reported on a wide range of outcomes with some indicating that, on average, Rystad top-quartile counties have benefited from the initiation of fracing, while others reveal less positive impacts. Guided by the conceptual framework outlined in Section 2, this section develops measures of willingness to pay for the change in local amenities and for the net local welfare consequences of the initiation of fracing based on estimated changes in housing prices and rents, income, and population. The section begins with an examination of the impacts of fracing?s initiation on housing markets, which is a key input into both willingness to pay expressions.

7.1 Housing Price and Quantity Estimates

Panel A of Table 8 reports on the impact of fracing?s initiation in Rystad top-quartile counties from the estimation of the long difference-in-differences specification detailed in equation (5.4). The estimates indicate that median and mean housing values for owner-occupied homes increased by 5.7 percent due to fracing. Further, the median price of mobile homes increased by almost 8 percent. Panel B indicates that rental prices for renter-occupied units increased by 2 to 3 percent.[61]

Returning to Appendix Table 4, Panel C explores the robustness of the estimated effect on log median housing values. The estimates are generally unchanged by the use of alternative Rystad measures (e.g., quartile versus octile and maximum versus mean). The models that add state-by-year fixed effects in column (2) tend to produce smaller point estimates, although the 95 percent confidence intervals of these estimates overlap with those in column (1).[62] In total, 17 of the 18 estimates fall in a range of roughly 2 percent to 6 percent and 15 of those 17 estimates would be judged to be statistically significant by conventional criteria. Allowing for spatial correlation, which is done in brackets below row 1, roughly doubles the standard errors, but the estimates in columns (1) and (3) still remain significant at a 95 percent level. Overall, we conclude that the initiation of fracing led to meaningful increases in housing prices in counties especially amenable to fracing, relative to other counties in the same shale play.

---

[61] Appendix Table 9 demonstrates that the housing price results are robust to including vacant homes and rentals in the calculation of mean home values and mean rents.

[62] Given that adding state fixed effects tends to reduce the estimated effect of fracing on housing prices, we explore how adding state fixed effects influences the play-specific results in Appendix Table 7. This table shows that adding state fixed effects does not dramatically influence many of the point estimates. The most notable change is that the estimate of the impact of fracing on housing prices for the Marcellus is reduced from roughly 9 percent to about 6 percent. It is perhaps not too surprising that the Marcellus estimates are influenced more than the estimates for other plays because the Marcellus overlaps 5 states in our sample. The only other play-specific estimate that changes markedly is the Haynesville estimate, which changes from a 7 percent estimated reduction in house prices to a 12 percent reduction. The Haynesville is roughly half in Texas and half in Louisiana, so it is also not surprising that adding state fixed effects influences the Haynesville results.

It is noteworthy that there is an extensive literature documenting the capitalization of various amenities into local housing prices and that 5.7 percent is a large effect for a county-level one.[63] For example, Chay and Greenstone (2005) find that the dramatic air quality improvements induced by the implementation of the Clean Air Act increased housing prices by just 2.5 percent in counties that faced strict regulation. Further, Currie et al. (2010) find that school facility investments lead to 4.2-8.6 percent increases in house prices but over the smaller geographic unit of school districts. While Currie et al. (2015) find that the opening of an industrial plant leads to 11 percent declines in housing prices, this effect is limited to houses within 0.5 miles of the plant.

Returning to Table 8, Panel C examines the impact on housing supply and land use. Contrary to the conventional wisdom, the data do not reveal a substantial increase in the number of housing units or even mobile homes. The point estimate for acres of agricultural land is large and negative, suggesting that some of this land is converted to residential usages; however, its associated t-statistic is less than 1.[64] It is noteworthy, however, that the vacancy rate for housing units declined by 1.0 percentage point.

A shortcoming of the housing supply data is that the end of period data is an average calculated from 2009-2013, and this includes several years where fracing was only in its early stages for multiple shale plays. As a complement and means to peek further into the future, Appendix Figure G.6 is an event study graph that examines the impact of the initiation of fracing in Rystad top-quartile counties on the number of housing unit construction permits issued. The figure suggests that there has been an increase in permits with the introduction of fracing but this increase does not become apparent until three years after fracing was initiated. The fitting of the column (2) version of equation (5.2) indicates that five years after fracing's initiation in these counties, the annual number of housing unit permits are about 30 percent higher; this is only statistically significant at the 10 percent level, which is not surprising in light of the noisiness in the event study figure (Panel C of Appendix Table 2).

7.2 Local Welfare Estimates

While there is little question that fracing increases local productivity, a central question in the debate about fracing is the magnitude of its negative aspects or its net impact on local amenities, and how large these negative aspects are relative to the increases in local income. With some assumptions, it is possible to develop a back-of-the-envelope estimate of the total local welfare

---

[63] It seems reasonable to presume that the 5.7 percent average effect obscures important within-county variation in housing price changes, and indeed this is an important finding in Muehlenbachs et al. (2014b).

[64] As for local public finance, the Census of Agriculture is reported in every year ending in 7 or 2. Consequently, it is unclear whether 2002 or 1997 is the best base year for the Barnett play because our first-frac date for the Barnett is in late 2001. In Appendix Table 13 we report specifications where we replace 2002 with 1997 as the base year. The point estimate for the effect of fracing on agricultural land quantities becomes 0.067 and is, again, imprecisely estimated. The sensitivity of the agricultural land results suggest that they must be interpreted with caution.

change caused by fracing, as well as the willingness-to-pay for the change in amenities. We use the local labor market model in developed in Section 2 above, that relaxes the assumptions of the canonical Roback (1982) to derive both estimates of the WTP for an amenity change and the total change in welfare. As we noted above, the intuition behind this approach comes from the fact that, in spatial equilibrium, the marginal resident must be indifferent to relocating, which means that local housing prices will respond to changes in local wages. The strength of this response will depend on both the elasticity of local housing supply and moving costs. Using estimates from the literature on the relationship between pure productivity shocks and house prices, we can then back out the change in local amenities and use these estimates to infer the total change in local welfare.

Specifically from Equation 2.1, WTP for the change in amenities can be expressed as:

$$\alpha\widehat{\Delta \ln A}_{at} = \quad s\widehat{\Delta \ln N}_{at} - (\widehat{\Delta \ln w}_{at} - \beta\widehat{\Delta \ln r}_{at}), \tag{7.1}$$

where $\Delta \ln N$ is the change in local population and s is the standard deviation of idiosyncratic location preferences or moving costs and the term in parentheses is the change in real income, which is measured as the difference between the change in wage and salary income per household, $\Delta \ln w$, and the product of the share of locally produced goods in the consumption basket, β, and the change in housing prices or rents (a proxy for a price index for local goods), $\Delta \ln r$.[65] Thus, WTP for the change in amenities, expressed as a percentage of income, is equal to the difference between the change in population, adjusted for the magnitude of moving costs, and the change in real wages. With the estimated WTP for the change in amenities, it is straightforward to develop an estimate for the WTP for allowing fracing (i.e., the net welfare change for original residents) by using 2.2, which also incorporates income from lease payments received by households.

Before proceeding, we further explore the expression for willingness to pay for amenities to provide further intuition. For example, consider the case where WTP for amenities is zero?here, the change in real income is equal to the adjusted change in population. Alternatively, when the population change is larger than the change in real income normalized by s, i.e., $\frac{\widehat{\Delta \ln w}_{at} - \beta\widehat{\Delta \ln r}_{at}}{s}$, then amenities must have risen (fallen); that is, at the margin, people are exchanging reductions in real incomes for higher amenity levels. Finally, higher values of s mean that location decisions are less responsive to changes in real wages.

Table 10 reports empirical estimates of the annual WTP for the change in amenities and annual WTP for allowing fracing using these equations, the above estimates, and a range of assumptions. The entries in Panel A report the mean annual WTP measures for original households in top-quartile

---

[65] The model discussed above is based on rents. If the housing market is perfectly competitive and the change in rents is constant after the introduction of fracing, then $\Delta \ln p_j = \frac{1}{1-\beta}\Delta \ln r$ and the percentage change in rents and house prices will be identical. In practice, we do not find an identical increase in house prices and rents. This result could be due to several factors, including the fact that homeowners receive oil and gas lease royalty payments while renters do not. Alternatively, the larger increase in house prices could reflect expectations about future growth associated with fracing.

counties. The entries in Panel B report the present value of WTP for permanently allowing fracing for original residents in these counties when the estimated annual changes in amenities, income, housing costs, etc are assumed to be constant and to last forever and a 5 percent discount rate is assumed. Columns (1) - (2) use the change in rental prices as the measure of the change in housing costs and columns (3) - (4) use the change in housing prices.

In both panels, the first row reports on estimates that assume that $\beta = 0.65$, the share of household wage and salary income spent on locally produced goods, following Albouy (2008) and $s = 0.40$, the standard deviation of idiosyncratic location preferences or moving costs, which is in the mid point of the range from 0.27 to 0.57 estimated by Diamond (2016).[66] The subsequent rows in each panel are based on alternative assumptions for β and s, although we believe the first row's assumptions are the most defensible. Throughout, we assume a 7.5 percent change in mean wage and salary income, a 9.3 percent change in interest and dividend income, and a 2.7 percent change in population (based on the Table 5 results)

The estimates suggest that the initiation of fracing decreases local amenities. Using the preferred assumptions, the estimated annual WTP is -$964 per household when the change in housing prices is used as a proxy for local prices and -$1,582 with the change in rental rates. Alternative assumptions about β and s do not greatly alter these estimates, supporting the conclusion that local amenities decline appreciably after fracing's initiation. If we assume that the decline in amenities is permanent, then the present value of the decline in local amenities is -$32 billion with housing prices and -$53 billion with rental rates. [67] Finally, we note that, in principle, these estimates captures all of the changes in positive and negative amenities, including any changes in truck traffic, criminal activity, noise and air pollution from drilling activity, and household beliefs regarding expected health impacts.

The full WTP for allowing fracing accounts for both the decline in amenities and the greater economic opportunities (i.e., it is the difference between the gross benefits and the gross costs). The estimates in columns (2) and (4) suggest that the net effect is positive meaning that on average the benefits exceed the costs. Specifically, we estimate that WTP for allowing fracing equals about $1,300 to $1,900 per household annually (i.e., 2.5 to 3.7 percent of annual income). If the changes in amenities and economic opportunities are permanent, Panel B suggests that the increase in welfare is in the neighborhood of $44 billion to $64 billion in the top quartile Rystad counties. As a basis of comparison, the estimated welfare gain is $10.4 billion when the canonical Roback model with its assumptions of inelastic housing supply and zero moving costs is combined with the paper's

---

[66] The 65% share of income spent on housing is significantly higher than the 30-40% usually found in the Consumer-Expenditure Survey. This difference is driven by two primary factors. First, as mentioned above, the 65% number incorporates the correlation between local rents and the prices of other locally traded goods, such as retail services, etc... Second, this 65% is in terms of household wage and salary income rather than total income.

[67] This calculation uses the 2000 Census population for each county.

estimates.[68] It is evident that with both models, and this paper's empirical estimates that the value of the greater economic opportunities outweighs the decline in local amenities.

Are these estimates plausible? Recall that our estimate of the impact of the introduction of fracing on local hydrocarbon production is roughly \$400 million per year, which, if it represented a permanent change, would have a present discounted value of \$8 billion dollars per county. There are 65 top-quartile counties, so the estimated national welfare gain of \$44 to \$64 billion is approximately 10% of the national increase in hydrocarbon production of \$520 billion. Thus, at least with this basis of comparison, these estimates seem reasonable.

It is worth underscoring that Table 10 has reported average estimates of WTP and it is unlikely that all residents are made better off by allowing fracing. For example, individuals who are not in the labor force will not benefit from the increase in local productivity. Renters who are not in the labor force are likely to fare especially poorly because they will face higher rents and no change in income. Additionally, homeowners who do not own the mineral rights to their property will not benefit from the drilling royalties, but may experience the negative impacts of drilling activity. The extent of the heterogeneity in the impacts of local productivity shocks and of changes in local amenities is a promising area for future research that requires more detailed micro data.

It is possible, however, to explore the heterogeneity in the WTP measures across shale plays. In Table 9, Panel E, we report the estimated change in WTP for amenities and local welfare separately by shale play. The estimates are qualitatively consistent across shale plays, with 8 of 10 shale plays experiencing declines in amenities or quality of life and 7 of 10 benefiting from welfare improvements. The largest estimated welfare gains are in the Bakken, which has received a lot of attention in the popular media, the Fayetteville and Marcellus plays.

It is natural to wonder about the sources of heterogeneity in the welfare impacts across the plays. Panel A reports the average population in top quartile counties and the share of hydrocarbon production value that comes from oil as we had ex ante assumed that these two variable would be important predictors of WTP to allow fracing. Among the three largest gainers one is dominated by petroleum (Bakken) and the other two (Fayettevile and Marellus) are dominated by natural gas production, underscoring that there this explanatory variable is imperfect. Besides observable predictors, it seems plausible that there is heterogeneity across shale plays in moving costs, s and the share of income spent on housing, β, due to differences in proximity to other labor markets, demographic composition, or tastes; such heterogeneity would lead to different estimates of the heterogeneity in the welfare impacts of fracing across shale plays than indicated in Table 9. Overall,

---

[68] Using the Roback model, the increase in housing prices of 5.7 percent implies an increase in each county's welfare of approximately \$160 million on average; this is a total welfare gain of roughly \$10.4 billion across the 65 Rystad top-quartile counties. The reason for this much smaller estimated welfare effect in the canonical Roback model is that when there are zero-moving costs and inelastic housing supply, large changes in income would cause very large rises in rents if amenities were unchanged. The fact that there is only a small rise in rents, despite the rise in wage and salary income, implies that there must have been a large decline in local amenities.

it is apparent that the question of where fracing offers the largest net benefits cannot be answered decisively with just ten data points.[69]

Two final points are noteworthy. First, these revealed preference estimates of WTP to allow fracing (and for amenity changes) are ultimately determined by households' knowledge. If new information causes households to update their estimates of fracing?s environmental and quality of life impacts, then this paper?s WTP estimates will necessarily change. Second, this paper?s estimates of WTP to allow fracing only reflect local changes in welfare. The global welfare effects of fracing include potentially very important consequences for petroleum, natural gas and electricity prices, local air pollution, global warming, and geopolitics. All of these impacts are outside the scope of this paper; however, none of them become relevant if local communities do not allow fracing within their jurisdictions.

8 Conclusions

Using a new identification strategy based on geological variation in shale deposits within shale plays, we estimate the effects of fracing on local communities. There are four primary findings. First, counties with high fracing potential produce roughly an additional $400 million of oil and natural gas annually three years after the discovery of successful fracing techniques, relative to other counties in the same shale play. Second, these counties experience marked increases in economic activity with gains in total income (4.4 - 6.9 percent), employment (3.6 - 5.4 percent), and salaries (7.6 - 13.0 percent). Further, local governments see substantial increases in revenues (15.5 percent) that are larger than the average increases in expenditures (12.9 percent) though the increased expenditures seem largely aimed at supporting the new economic activity, with little effect, for example, on per pupil expenditures in public schools. Third, there is evidence of deterioration in the quality of life or total amenities, perhaps most notably marginally significant estimates of higher violent crime rates, despite a 20 percent increase in public safety expenditures. We estimate that annual willingness-to-pay (WTP) for fracing-induced changes in local amenities are roughly equal to -$1,000 to -$1,600 per household annually (i.e., -1.9 to -3.1 percent of annual mean household income). Fourth, we estimate that mean WTP for allowing fracing equals about $1,300 to $1,900 per household annually (2.5 to 3.7 percent of median household income) among original residents of counties with high fracing potential.

The discovery of hydraulic fracturing is widely considered the most important change in the energy sector since the commercialization of nuclear energy in the 1950s. To date, almost all of the fracing activity has been confined to North America, yet even so it has upended many features of the global economy, global environment, and international relations. There are substantial shale

---

[69] In Appendix Table 6 we report play-specific estimates instead using the change in rents to measure house prices. This table also reports aggregate affects of fracing on welfare by play.

deposits both in North America and other parts of the world that have not been exploited to date so there is potential for further change. This paper demonstrates that to date local communities that have allowed fracing have benefited on average, although there is evidence of important heterogeneity in the local net benefits.

References

Albouy, D. (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.

Allcott, H. and D. Keniston (2014): “Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America,” Tech. rep., NYU Working Paper.

Baumeister, C. and L. Kilian (2016): “Forty Years of Oil Price Fluctuations: Why the Price of Oil May Still Surprise Us,” *Journal of Economic Perspectives*, 30, 139–60.

Boslett, A., T. Guilfoos, and C. Lang (2015): “Valuation of Expectations: A Hedonic Study of Shale Gas Development and New York’s Moratorium,” Working paper, University of Rhode Island.

Budzik, P. (2013): “The Application of Hydraulic Fracturing to US and World Shale Gas and Tight Oil Deposits,” Tech. rep., 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,” Tech. rep.

Bureau of Labor Statistics, U.S Department of Labor (2015): *CPI Detailed Report, Table 24: Historical Consumer Price Index for All Urban Consumers (CPI-U): US city average, all items.*

Cascio, E. U. and A. Narayan (2015): “Who Needs a Fracking Education? The Educational Response to Low-Skill Biased Technological Change,” Working paper, Dartmouth College.

Chay, K. Y. and M. Greenstone (2005): “Does Air Quality Matter? Evidence from the Housing Market,” *Journal of Political Economy*, 113, pp. 376–424.

Conley, T. (1999): “GMM Estimation with Cross Sectional Dependence,” *Journal of Econometrics*, 92, 1–45.

Covert, T., M. Greenstone, and C. R. Knittel (2016): “Will We Ever Stop Using Fossil Fuels?” *Journal of Economic Perspectives*, 30, 117–38.

Covert, T. R. (2014): "Experiential and Social Learning in Firms : The Case of Hydraulic Fracturing in the Bakken Shale JOB MARKET PAPER," 1–43.

Currie, J., L. Davis, M. Greenstone, and R. Walker (2010): "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *The Quarterly Journal of Economics*, 125, 215–261.

——— (2015): "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings," *American Economic Review*, 105, 678–709.

Deschenes, O., M. Greenstone, and J. S. Shapiro (2012): "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program and Ozone Reductions," Working Paper 18267, National Bureau of Economic Research.

Diamond, R. (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.

Edlund, L., H. Li, J. Yi, and J. Zhang (2013): "Sex Ratios and Crime: Evidence from China," *Review of Economics and Statistics*, 95, 1520–1534.

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

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," Tech. Rep. June.

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

Fetzer, T. (2015): "Fracking Growth," Working paper, London School of Economics.

Feyrer, J., E. T. Mansur, and B. Sacerdote (2015): "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution," Working Paper 21624, National Bureau of Economic Research.

Gopalakrishnan, S. and H. A. Klaiber (2013): "Is the Shale Energy Boom a Bust for Nearby Residents? Evidence from Housing Values in Pennsylvania," *American Journal of Agricultural Economics.*

Greenstone, M., R. Hornbeck, and E. Moretti (2010): "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 118, pp. 536–598.

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

Hornbeck, R. and E. Moretti (2015): "Who Benefits From Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality," Tech. rep.

Hsiang, S. (2010): "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America," *Proceedings of the National Academy of Sciences*, 107, 15367–15372.

Imbens, G. and D. Rubin (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, New York, NY: Cambridge University Press, first ed.

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

Jacobsen, G. D. (2016): "Who Wins in an Energy Boom? Evidence from Wage Rates and Housing," Tech. rep.

Kline, P. and E. Moretti (2015): "People, Places and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics.*

Maniloff, P. and R. Mastromonaco (2014): "The Local Economic Impacts of Hydraulic Fracturing and Determinants of Dutch Disease," Tech. rep.

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

Minnesota Population Center (2011): *National Historical Geographic Information System: Version 2.0*, Minneapolis, MN: University of Minnesota.

Moretti, E. (2011): "Local Labor Markets," Elsevier, vol. 4B, chap. 14, 1237–1313, 1 ed.

Muehlenbachs, L., E. Spiller, and C. Timmins (2014a): "The Housing Market Impacts of Shale Gas Development," Working Paper 19796, National Bureau of Economic Research.

——— (2014b): "The Housing Market Impacts of Shale Gas Development," Working Paper 19796, National Bureau of Economic Research.

National Center for Education Statistics (2015): "Common Core Data: Local Education Agency (School District) Universe Survey Data," Tech. rep.

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

NEWELL, R. G. AND D. RAIMI (2015): “Shale Public Finance: Local Government Revenues and Costs Associated with Oil and Gas Development,” Working Paper 21542, National Bureau of Economic Research.

PHILLIPS, S. (2014): *Pennsylvania Frack Ponds Now Number More than 500.*

ROBACK, J. (1982): “Wages, Rents, and the Quality of Life,” *Journal of Political Economy*, 90, pp. 1257–1278.

RUBINSTEIN, J. L. AND A. B. MAHANI (2015): “Myths and Facts on Wastewater Injection, Hydraulic Fracturing, Enhanced Oil Recovery, and Induced Seismicity,” *Seismological Research Letters*, 86, 1060–1067.

RYSTAD ENERGY (2014): *North American Shale Plays Maps: Q3, 2014*, Used with permission: http://www.rystadenergy.com/.

SUAREZ SERRATO, J. C. AND O. ZIDAR (2016): “Who Benefits From State Corporate Tax Cuts ? a Local Labor Markets Approach with Heterogeneous Firms,” *American Economic Review*, Forthcoming.

US BUREAU OF ECONOMIC ANALYSIS (BEA) (2014): *Regional Economic Information System: Local Area Personal Income and Employment*, Washington, DC: BEA.

US CENSUS BUREAU (2014a): *Building Permits Survey.*

——— (2014b): *Census of Governments: Finances.*

WEBER, J. G. (2012): “The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming,” *Energy Economics*, 34, 1580–1588.

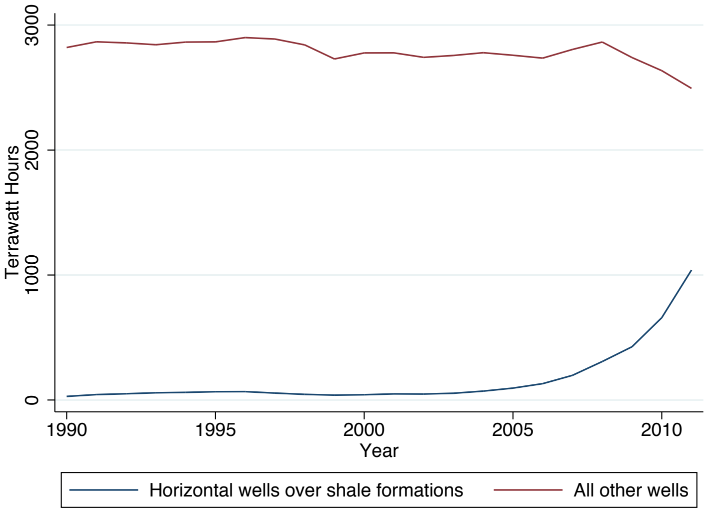
WEINSTEIN, A. (2014): “Local Labor Market Restructuring in Shale Booms,” *Journal of Regional Analysis and Policy*, 44, 71–92.

WYNVEEN, B. J. (2011): “A Thematic Analysis of Local Respondents Perceptions of Barnett Shale Energy Development,” *Journal of Rural Social Sciences*, 26, 8–31.

ZAGORSKI, W. A., G. WRIGHTSTONE, AND D. 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*, ed. by J. A. Breyer, no. 97 in AAPG Memoir, 172 – 200.

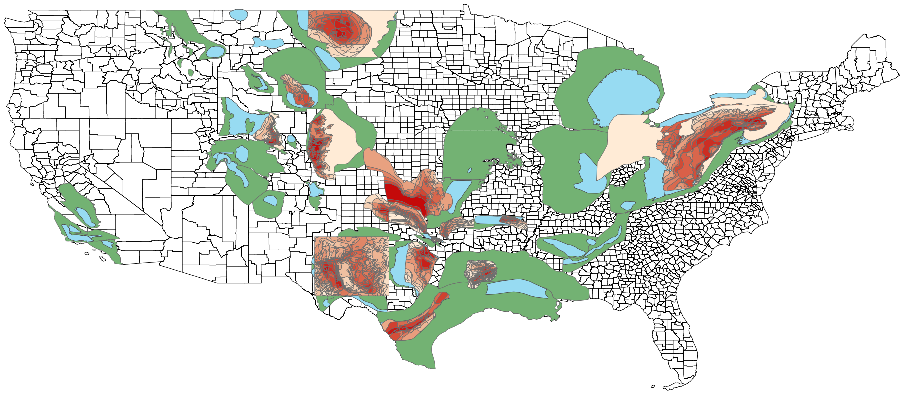
9 Figures

Figure 1: Hydrocarbon production from horizontal wells over shale play



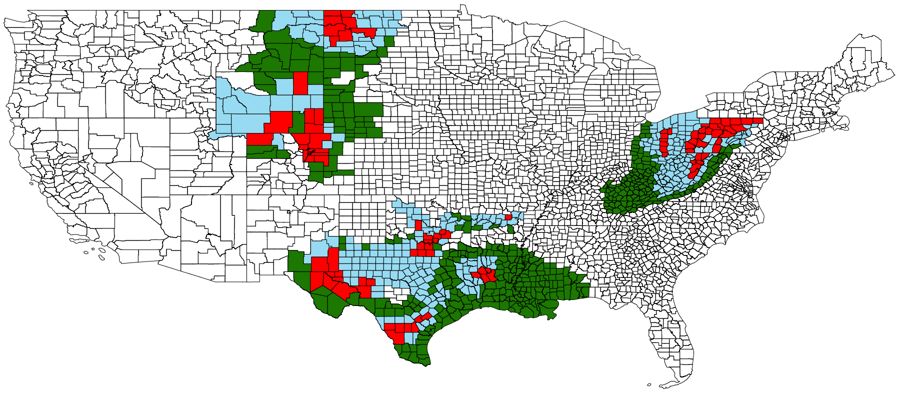
Notes: This figure plots the total energy content of hydrocarbons produced from horizontal wells over shale plays over time. In 1991, there is almost no production from these wells. However, as a results of the technological innovations in using fracing and horizontal drilling into shale formations, these types of wells have grown dramatically as a share of US hydrocarbon production, rising to more than a quarter of all US hydrocarbon production by 2011. The data come from Drilling Info, Inc (2012).

Figure 2: Shale basins, plays, and prospectivity scores



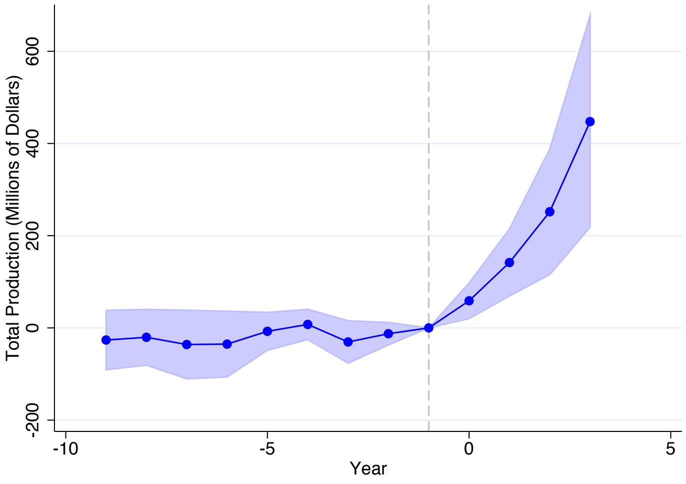
Notes: This figure overlays shale basins, shale plays, and Rystad prspectivity scores over a map of US counties. Shale basins are shown in green, shale plays are shown in blue, and Rystad Prospectivity scores are shown in shades of red, with darker red indicating a higher prospectivity score. Shapefiles for US shale basins and plays comes from the Energy Information Agency (2011), while prospectivity scores were purchased from Rystad Energy (2014).

Figure 3: County prospectivity score classifications



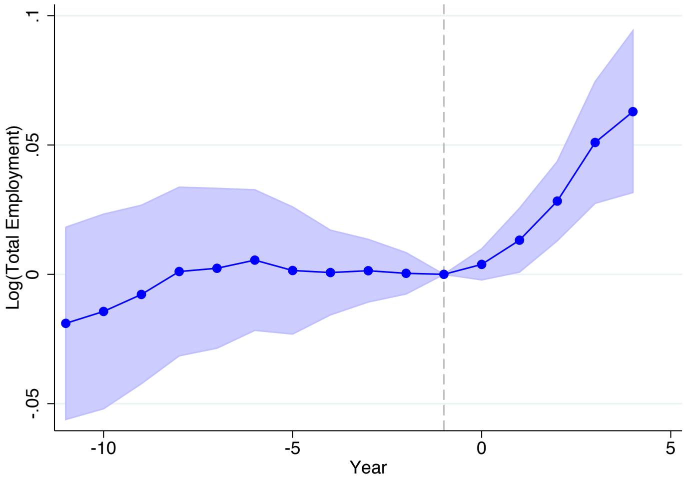
Notes: This figure shows prospectivity score classifications for counties in the contiguous US. Counties in red are in the top quartile of the Rystad prospectivity measure, counties in blue are not in the top quartile of Rystad prospectivity but are within a shale play, and counties in green are not in a shale play, but are in a shale basin.

Figure 4: Event study analysis of county-level value of hydrocarbons



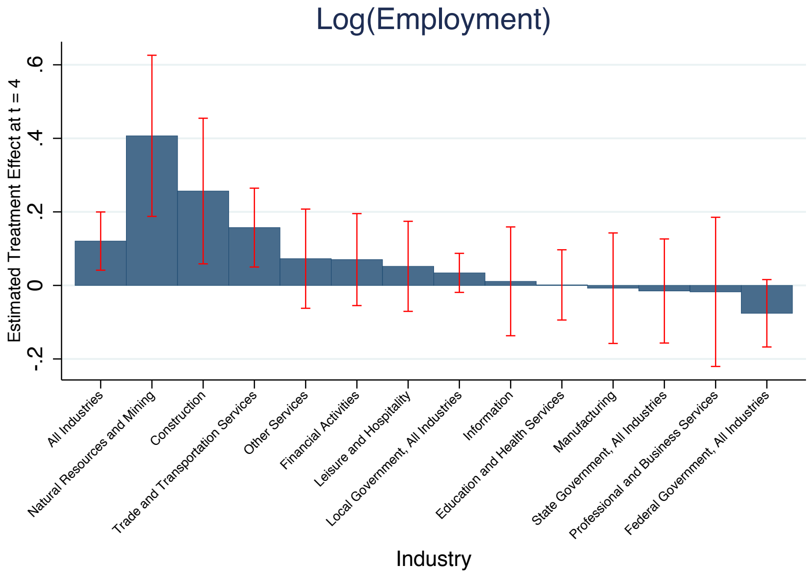
Notes: This figure plots results from an event-study analysis of the difference in the county-level value of hydrocarbon production between high-fracing potential counties and other counties in shale plays before and after fracing began. The reported coefficients come from fitting a modified version of Equation 5.1 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 hydrocarbon production from 1992 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.

Figure 5: Event study analysis of total employment



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. The reported coefficients come from fitting a modified version of Equation 5.1 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 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) (2014). Specifically, we use the the variable CA25-10. The shaded blue region shows 95 percent confidence intervals calculated using standard errors clustered at the county level.

Figure 6: Employment effects by industry



Notes: This figure plots estimates of the effect of fracing on employment by industry five years after the start of fracing. Each bar reports results of fitting Equation 5.2 for the given industry, which corresponds to Column (2) in the tables. Equation 5.2 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.

10 Tables

Table 1: Treatment and control counties by shale basin

Shale Play (1)	Shale Basin (2)	Play First Frac Year (3)	Top Quartile Counties (4)	Outside Top Quartile Counties (5)
Woodford-Anadarko	Anadarko	2008	1	10
Marcellus	Appalachian	2008	28	95
Utica	Appalachian	2012	7	18
Woodford-Ardmore	Ardmore	2007	4	5
Fayetteville	Arkoma	2005	1	13
Woodford-Arkoma	Arkoma	2006	2	7
Niobrara-Denver	Denver	2010	13	4
Barnett	Forth Worth	2001	5	41
Niobrara-Greater Green River	Greater Green River	2012	2	9
Permian All Plays	Permian	2005	11	34
Niobrara-Powder River	Powder River	2010	1	5
Haynesville	TX-LA-MS Salt	2008	5	21
Eagle Ford	Western Gulf	2009	7	21
Bakken	Williston Basin	2007	8	27
Total			95	310

Notes: This table shows the number of counties by shale play and Rystad prospectivity value. Top Quartile = 1 if the county is in the top-quartile of the Rystad max prospectivity measure within its shale-play and 0 otherwise. Different shale plays have different geological features and were developed at different time periods. Column (3) shows the first year the fracing potential of the shale play became public.

Table 2: Comparison of pre-trends and levels across treatment and control counties

	Mean Value in US (1)	Basin vs. Rest of US (2)	Play vs. Basin (3)	Rystad Top Quartile vs. Play (4)	Rystad Top Quartile vs. Pscore Matched Sample (5)	Quartiles 1-3 vs. Pscore Matched Sample (6)
Panel A: Covariate Balance (All Variables measured in 2000 unless noted)						
Panel A1: Non-Crime Variables						
Log(Real Median Home Values)	11.897	-0.402***	-0.071**	0.039	-0.103	-0.149***
		(0.037)	(0.031)	(0.050)	(0.067)	(0.041)
Log(Real Median Home Rental Prices)	6.621	-0.179***	-0.023	0.055	-0.091	-0.095***
		(0.032)	(0.030)	(0.045)	(0.066)	(0.037)
Log(Total Housing Units)	9.427	-0.159***	0.413***	0.082	-0.193	-0.342***
		(0.055)	(0.087)	(0.143)	(0.169)	(0.111)
Log(Total Employment)	9.533	-0.242***	0.402***	0.057	-0.283	-0.397***
		(0.060)	(0.104)	(0.161)	(0.180)	(0.119)
Log(Total Income per capita)	13.594	-0.279***	0.416***	0.032	-0.309	-0.408***
		(0.062)	(0.103)	(0.171)	(0.195)	(0.123)
Share of Population with Bachelor's Degree or more	0.241	-0.041***	0.003	0.042*	-0.001	-0.026**
		(0.010)	(0.016)	(0.025)	(0.027)	(0.013)
Share of Population Ages 18-64	0.619	-0.003	-0.011**	-0.003	-0.001	0.003
		(0.003)	(0.004)	(0.007)	(0.010)	(0.006)
Log(Real Total Government Revenue: 2002 - 1992)	11.512	-0.273***	0.374***	0.050	-0.314*	-0.411***
		(0.059)	(0.101)	(0.159)	(0.178)	(0.115)
Log(Real Total Government Expenditures: 2002 - 1992)	11.515	-0.283***	0.373***	0.063	-0.309*	-0.421***
		(0.060)	(0.102)	(0.162)	(0.181)	(0.117)
Total Value of Hydrocarbon Production: 2000 - 1992	56.238	81.559***	78.570***	108.280*	99.435	-1.201
		(19.990)	(17.698)	(58.527)	(67.217)	(42.595)
F-statistic		23.7	7.6	1.7	3.1	3.3
P-value		0.00	0.00	0.08	0.00	0.00
Counties Exposed	-	715	316	64	64	252
N	2,842	2,842	792	401	1,384	1,599
Panel A2: Crime-Variables						
Log(Violent Crimes)	6.453	-0.405***	0.102	0.163	-0.793***	-0.951***
		(0.096)	(0.185)	(0.229)	(0.252)	(0.198)
Log(Property Crimes)	4.127	-0.223**	0.177	0.113	-0.706***	-0.791***
		(0.097)	(0.172)	(0.216)	(0.256)	(0.200)
F-statistic		12.7	0.6	0.2	3.3	8.7
P-value		0.00	0.64	0.90	0.02	0.00
Counties Exposed		523	266	56	56	210
N	2,071	2,071	586	340	879	1,061
Panel B: Pre-Trends (Change 1990 - 2000 unless noted)						
Panel B1: Non-Crime Variables						
Log(real median home values)	0.110	0.020	-0.022	-0.011	0.043**	0.012
		(0.026)	(0.014)	(0.028)	(0.020)	(0.020)
Log(real median home rental prices)	0.012	0.055***	-0.027***	0.003	-0.013	-0.007
		(0.016)	(0.006)	(0.008)	(0.018)	(0.015)
Log(Total Housing Units)	0.124	-0.035***	-0.054***	0.009	-0.036***	-0.047***
		(0.005)	(0.008)	(0.012)	(0.014)	(0.008)
Log(Total Employment)	0.179	-0.040***	-0.028**	0.028*	-0.013	-0.039***
		(0.007)	(0.012)	(0.016)	(0.018)	(0.012)
Log(Total Income per capita)	0.268	-0.044***	-0.068***	0.034*	-0.022	-0.054***
		(0.007)	(0.014)	(0.018)	(0.021)	(0.014)
Share of Population with Bachelor's Degree or more	0.040	-0.012***	0.002	0.013***	0.011**	-0.003
		(0.003)	(0.003)	(0.005)	(0.005)	(0.003)
Share of Population Ages 18-64	0.001	0.005***	0.000	-0.006	-0.003	0.004
		(0.002)	(0.003)	(0.004)	(0.005)	(0.003)
Log(Real Total Government Revenue: 2002 - 1992)	0.286	-0.063***	-0.113***	0.042	-0.023	-0.064***
		(0.011)	(0.019)	(0.027)	(0.027)	(0.021)
Log(Real Total Government Expenditures: 2002 - 1992)	0.290	-0.029***	-0.124***	0.034	-0.026	-0.059***
		(0.011)	(0.020)	(0.029)	(0.031)	(0.022)
Total Value of Hydrocarbon Production: 2000 - 1992	7.934	6.845*	4.036	28.929	2.638	-27.676
		(4.150)	(7.246)	(18.096)	(22.938)	(19.179)
F-statistic		14.1	8.8	1.4	2.4	4.0
P-value		0.0	0.0	0.2	0.0	0.0
Counties Exposed		715	316	64	64	252
N	2,842	2,842	792	401	1,384	1,599
Panel A2: Crime-Variables (Change 1992 - 2000)						
Log(Violent Crimes)	-0.093	-0.043	-0.125*	0.104	-0.022	-0.130***
		(0.026)	(0.066)	(0.074)	(0.065)	(0.048)
Log(Property Crimes)	-0.020	-0.026	0.132	0.191*	0.187*	-0.055
		(0.039)	(0.089)	(0.108)	(0.104)	(0.074)
F-statistic		0.9	3.0	1.5	1.1	2.5
P-value		0.44	0.03	0.21	0.34	0.06
Counties Exposed		523	266	56	56	210
N	2,071	2,071	586	340	879	1,061

Notes: This table shows coefficients from regressions of baseline outcomes (Panel A) and pre-trends (Panel B) on different measures of exposure to Fracing activity. Column (1) shows the mean value for the entire US. Column (2) shows regressions of covariates and pre-trends on an indicator for being in a shale basin. Column (3) shows regressions of covariates and pre-trends on an indicator for being in a shale-play (restricting the sample to counties in a shale basin). Column (4) shows regressions of covariates and pre-trends on an indicator for being in the top quartile of max prospectivity (restricting the sample to counties in a shale basin). Column (5) shows regressions of covariates and pre-trends on an indicator for being in the top quartile of max prospectivity, but the sample is top quartile counties and the corresponding pscore-matched counties for each shale play. Column (6) shows regressions of covariates and pre-trends on an indicator for being in quartiles one through three of max prospectivity, but the sample is the bottom three quartile counties and the corresponding pscore-matched counties for each shale play. All specifications include both the fracing exposure measure and the fracing exposure measure interacted with an indicator for being in the unbalanced sample (defined as having a first-frac date after 2008). The coefficients reported correspond to the balanced sample. Column (3) includes basin fixed effects and Columns (4), (5), and (6) include play fixed effects. Below Panel A we report the joint F-test that all the coefficients are equal to 0 in the covariate regression. Below Panel B we report the joint F-test that all coefficients are equal to 0 in the pre-trends regression. Estimated outcome variables (such real median home values) are weighted by the sample size for the estimate (such as number of owner occupied homes for real median home values). All monetary figures are shown in 2010 USD. Robust standard errors are reported in parentheses in Columns (2)-(4). Columns (5) and (6) cluster standard errors at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Impact of fracing on the value of hydrocarbon production

	(1)	(2)	(3)
Panel A: Total Value of Oil and Gas Production			
1(Fracing Exposure)*1(Post)	242***	36	36
	(68)	(47)	(23)
t*1(Fracing Exposure)		3	
		(6)	
t*1(Fracing Exposure)*1(Post)		124***	125***
		(37)	(38)
Fracing Exposure Effect at tau=3	242***	409***	410***
	(68)	(123)	(115)
Fracing Exposure Group	Top Quartile	Top Quartile	Top Quartile
Control Group	Quartiles 1-3	Quartiles 1-3	Quartiles 1-3
Fracing Exposure Level Shift	Y	Y	Y
Fracing Exposure Trend	N	Y	Y
Fracing Exposure Trend Break	N	Y	Y
County Fixed Effects	Y	Y	Y
County-Specific Trends	N	N	Y
Year-Play Fixed Effects	Y	Y	Y
Restricted to Balanced Sample	N	N	Y

Notes: This table reports regressions of oil/gas production variables on fracing exposure. Fracing exposure is measured using an indicator for whether the county is in the fourth quartile of the Rystad max prospectivity score among counties within the shale play with a non-missing Rystad value. Oil and gas production data come from HPDI well data aggregated to the county level. 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. 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 Fracing Exposure Effect at tau=3 correspond to the 1(Fracing Exposure)*1(Post) coefficient plus 3 times the t*1(Fracing Exposure)*1(Post) coefficeint. 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 8100 county-year observations from 405 total counties, of which 65 Rystad top quartile and 253 outside top quartile counties are in the balanced sample. Column (3) includes 4,134 observations from 318 total counties, of which 65 Rystad top quartile and 253 outside top quartile counties are in the balanced sample.

Table 4: Impact of fracing on employment and aggregate income: time-series specifications

	(1)	(2)	(3)
Panel A: Log(Total Employment)			
Fracing Exposure Effect at tau=4	0.036**	0.054*	0.049***
	(0.016)	(0.029)	(0.019)
Panel B: Income			
Log(Total Income)			
Fracing Exposure Effect at tau=4	0.056***	0.069**	0.044**
	(0.015)	(0.028)	(0.021)
B1. Log(Total Wage/Salary Income): 56 percent of total personal income			
Fracing Exposure Effect at tau=4	0.076***	0.130***	0.089***
	(0.021)	(0.035)	(0.030)
B2. Log(Total Rents/Dividends): 19 percent of total personal income			
Fracing Exposure Effect at tau=4	0.070***	0.080**	0.068**
	(0.019)	(0.038)	(0.028)
B3. Log(Total Transfers): 10 percent of total personal income			
Fracing Exposure Effect at tau=4	0.012	0.001	-0.005
	(0.012)	(0.020)	(0.008)
B4. Log(Total Proprieter's Income): 18 percent of total personal income			
Fracing Exposure Effect at tau=4	0.036	-0.101	-0.041
	(0.040)	(0.064)	(0.069)
Panel C: Migration			
C1. Log(In Migration)			
Fracing Exposure Effect at tau=4	0.044**	0.073*	0.005
	(0.017)	(0.038)	(0.042)
C2. Log(Out Migration)			
Fracing Exposure Effect at tau=4	-0.001	0.007	-0.047
	(0.013)	(0.031)	(0.035)
Fracing Exposure Group	Top Quartile	Top Quartile	Top Quartile
Control Group	Quartiles 1-3	Quartiles 1-3	Quartiles 1-3
Fracing Exposure Level Shift	Y	Y	Y
Fracing Exposure Trend	N	Y	Y
Fracing Exposure Trend Break	N	Y	Y
County Fixed Effects	Y	Y	Y
County-Specific Trends	N	N	Y
Year-Play Fixed Effects	Y	Y	Y
Restricted to Balanced Sample	N	N	Y

Notes: This table reports regressions of aggregate economic outcomes on fracing exposure measured using an indicator for whether the county is in the fourth quartile of the Rystad max prospectivity score among counties within the shale play with a non-missing Rystad value. 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. Column (1) allows for a level shift in fracing exposed counties. Columns (2) and (3) allow for pre-trends, a post-fracing level shift, and a post-fracing trend break in counties exposed to fracing. In Columns (1) and (2), all fracing exposure 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 + 4 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: Includes all counties in any shale play with non-missing data in all years from 1990 to 2012. Panels A, B, B1, B2, and B3, Columns (1) and (2) include 9246 observations from 402 total counties, of which 65 Rystad top quartile counties and 252 outside top quartile counties are in the balanced sample. Panels A, B, B1, B2, and B3, Column (3) include 5,072 observations from 317 total counties, of which 65 Rystad top quartile and 252 outside top quartile counties are in the balanced sample.

Panel B4, Columns (1) and (2) include 8,740 observations from 380 total counties, of which 60 Rystad top quartile and 237 outside top quartile counties are in the balanced sample. Panel B4, Column (3) includes 4,752 observations from 297 total counties, of which 60 Rystad top quartile and 237 outside top quartile counties are in the balanced sample.

Panel C, Columns (1) and (2) include 7,900 observations from 395 total counties, of which 63 Rystad top quartile and 248 outside top quartile counties are in the balanced sample. Panel C, Column (3) includes 4,043 observations from 311 total counties, of which 63 Rystad top quartile and 248 outside top quartile counties are in the balanced sample.

Table 5: Impact of fracing on employment and aggregate income: long-difference specifications

	(1)
Panel A: Employment Outcomes:	
A1. Log(Total Employment)	
	0.048***
	(0.017)
A2. Employment-to-Population Ratio	
	0.026***
	(0.009)
A3. Unemployment Rate	
	-0.006*
	(0.003)
Panel B: Household Income:	
B1. Log(Mean Real Household Income)	
	0.058***
	(0.012)
B2. Log(Mean Real Household Wage and Salary Income)	
	0.075***
	(0.017)
B3. Log(Mean Real Rent and Dividend Income)	
	0.093**
	(0.037)
Panel C: Population:	
C1. Log(Population)	0.027*
	(0.016)
Fracing Exposure Group	Top Quartile
Control Group	Quartiles 1-3
Play Fixed Effects	Y

Notes: This table reports long-difference regressions of the change in county aggregate economic outcomes between 2000 and 2009/2013 on a measure of fracing exposure. Fracing exposure is measured using 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, and the control group are quartiles one through three. The fracing exposure measure 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. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Panels A1, B, and C include observations from 404 total counties, of which 65 Rystad top quartile and 253 outside top quartile counties are in the balanced sample.

Panels A2 and A3 include observations from 403 total counties, of which 64 Rystad top quartile and 253 outside top quartile counties are in the balanced sample.

Table 6: Impact of fracing on crime

	(1)	(2)	(3)
Panel A. Log(Total Crime)			
Top Quartile Effect at tau=5	0.072	-0.042	-0.004
	(0.056)	(0.082)	(0.101)
Panel B: Log(Violent Crime)			
Top Quartile Effect at tau=5	0.116*	0.208*	0.109
	(0.068)	(0.124)	(0.142)
Panel C: Log(Property Crime)			
Top Quartile Effect at tau=5	0.065	-0.057	0.000
	(0.057)	(0.087)	(0.106)
Fracing Exposure Group	Top Quartile	Top Quartile	Top Quartile
Control Group	Quartiles 1-3	Quartiles 1-3	Quartiles 1-3
Fracing Exposure Level Shift	Y	Y	Y
Fracing Exposure Trend	N	Y	Y
Fracing Exposure Trend Break	N	Y	Y
County Fixed Effects	Y	Y	Y
County-Specific Trends	N	N	Y
Year-Play Fixed Effects	Y	Y	Y
Restricted to Balanced Sample	N	N	Y

Notes: This table reports regressions of crime rates on fracing exposure. Fracing exposure is measured using an indicator for being in the Top Quartile of max prospectivity among the counties with Rystad data within the shale play. The fracing exposure 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 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. 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)-(3) include all counties in any shale play with non-missing data in all years from 1992 to 2013. Columns (1) and (2) include 7480 observations from 340 total counties, of which 56 Rystad top quartile and 210 outside top quartile counties are in the balanced sample. Column (3) includes 3,990 observations from 266 total counties, of which 56 Rystad top quartile and 210 outside top quartile counties are in the balanced sample.

Table 7: Impact of fracing on local government revenues and expenditures

		(1)
Panel A: Log(Total Expenditures): 2012 - 2002		
		0.129***
		(0.034)
A. Log(Direct Expenditures)		
		0.123***
		(0.033)
A1. Direct Expenditures by Type		
A1a. Log(Current Operating Expenditure):		
	[84%]	0.107***
		(0.028)
A1b. Log(Capital Outlays):		
	[12%]	0.181
		(0.135)
A2. Direct Expenditures by Purpose		
A2a. Log(Education Expenditures):		
	[48%]	0.025
		(0.032)
A2b. Log(Public Safety Expenditures):		
	[8%]	0.195***
		(0.063)
A2c. Log(Welfare and Hospital Expenditures):		
	[10%]	0.240
		(0.154)
A2d. Log(Infrastructure and Utility Expenditures):		
	[18%]	0.242***
		(0.071)
A2e. Log(Other Expenditures):		
	[16%]	0.122*
		(0.063)
Panel B: Log(Total Revenues): 2012 - 2002		
		0.155***
		(0.032)
B1. Revenues by Type		
B1a. Log(Property Tax Revenues):		
	[24%]	0.133***
		(0.042)
B1b. Log(Sales Tax Revenues):		
	[4%]	0.594***
		(0.120)
B1c. Log(Other Tax Revenues):		
	[2%]	0.038
		(0.155)
B1d. Log(Intergovernmental Revenues):		
	[42%]	0.100
		(0.081)
B1e. Log(Charges Revenues):		
	[14%]	0.095
		(0.079)
B1f. Log(Other Revenues):		
	[14%]	0.261***
		(0.066)
Panel C: Government Balance Sheets		
C. Net Financial Position as Share of Revenues		
		-0.020
		(0.067)
Panel D: Log(Elem/Sec Education Spending per Pupil)		
		0.008
		(0.034)
Fracing Exposure Group		Top Quartile
Control Group		Quartiles 1-3
Play Fixed Effects		Y

Notes: This table shows regressions on the change in government spending and revenues between 2002 and 2012 on fracing exposure measured using 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, and the control group are quartiles one through three. The fracing exposure measure 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. Panels A1 and B1 show the share of total government revenues or expenditures represented by the given category in brackets below the category name. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Panels A, B, and C, include all counties in any shale play, 405, of which 65 Rystad top quartile and 253 outside top quartile counties are in the balanced sample. Panel D includes all 385 counties in shale plays with non-missing school enrollment data for all districts in 1997, 2002, and 2012, of which 61 Rystad top quartile and 244 outside top-quartile counties are in the balanced sample.

Table 8: Impact of fracing on housing outcomes

	(1)
Panel A: House Values	
A1. Log(Median House Value)	
	0.057***
	(0.018)
A2. Log(Mean Housing Value)	
	0.057***
	(0.018)
A3. Log(Mobile Housing Units: Median Housing Value)	
	0.079**
	(0.037)
Panel B: Rental Prices	
B1. Log(Median Rental Price)	
	0.020*
	(0.010)
B2. Log(Mean Rental Price)	
	0.029***
	(0.011)
Panel C: Housing Quantities	
C1. Log(Total Housing Units)	
	0.011
	(0.012)
C2. Log(Total Mobile Homes)	
	0.022
	(0.028)
C3. Share of Housing Units Vacant	
	-0.010**
	(0.005)
C4. Log(Acres of Agricultural Land)	
	-0.099
	(0.144)
Fracing Exposure Group	Top Quartile
Control Group	Quartiles 1-3
Play Fixed Effects	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 a measure of fracing exposure. Fracing exposure is measured using 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, and the control group are quartiles one through three. The fracing exposure measure 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. 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. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Sample: Includes all counties in any shale play. Panels A1, A3, B1, B2, C1, C2, and C3 contain observations from 404 total counties, of which 65 Rystad top quartile and 253 outside top quartile counties are in the balanced sample. Panel C4 contains observations from 345 total counties, of which 53 Rystad top quartile and 211 outside top quartile counties are in the balanced sample.

Table 9: Play specific Estimates

	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													
	409***	972**	322*	69	1,730*	185***	-452***	123*	199	169	F-stat	11.4	1,412***
	(123)	(414)	(183)	(78)	(903)	(70)	(65)	(70)	(158)	(134)	p-value	0.00	(270)
Panel C: Labor Markets													
C1. Log(Mean household total income)													
	0.058***	0.293***	0.045*	0.099	0.080	0.049***	0.069	-0.013	0.000	0.170***		5.4	-0.015
	(0.012)	(0.083)	(0.025)	(0.110)	(0.053)	(0.012)	(0.084)	(0.079)	(0.134)	(0.049)		0.00	(0.046)
C2. Log(Mean household wage and salary income)													
	0.075***	0.286***	0.031	-0.014	0.078	0.078***	0.079	-0.028	0.075	0.177***		5.5	-0.003
	(0.012)	(0.100)	(0.030)	(0.133)	(0.064)	(0.014)	(0.102)	(0.095)	(0.161)	(0.059)		0.00	(0.056)
C3. Log(Mean household rent, dividend, and interest income)													
	0.093**	0.833***	0.061	0.671	0.078	0.086*	-0.171	0.116	0.495	-0.006		1.7	0.173
	(0.038)	(0.313)	(0.095)	(0.417)	(0.201)	(0.045)	(0.319)	(0.297)	(0.505)	(0.183)		0.09	(0.174)
C4. Log(Population)													
	0.027*	0.130***	0.071	-0.014	-0.045	0.018	0.060	0.042	-0.038	-0.007		1.4	-0.090*
	(0.016)	(0.045)	(0.053)	(0.115)	(0.055)	(0.024)	(0.117)	(0.075)	(0.089)	(0.039)		0.20	(0.048)
Panel D: Housing Prices													
D1. Log(Median home values)													
	0.057***	0.228***	-0.046	0.018	-0.071	0.089***	-0.074	-0.032	0.051	0.029	F-stat	6.0	-0.021
	(0.012)	(0.086)	(0.030)	(0.111)	(0.057)	(0.014)	(0.091)	(0.082)	(0.138)	(0.051)	p-value	0.00	(0.055)
Panel E: Annual Change in WTP for Amenities and Welfare per Household, Using Change in Mean Home Values (dollars)													
E1. Change in amenities	-$964	-$2,395	-$1,518	$631	-$4,455	-$484	-$3,882	$729	-$1,466	-$5,409			-$1,543
E2. Change in welfare	$1,931	$9,068	$157	$2,884	-$1,784	$2,583	-$1,352	$197	$1,182	$533			-$873
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 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. Panel B allows 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 Panel B 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 and 4 years after the start of fracing for Panel B. Panels C and D 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. Panel D also includes controls for changes in average county owner(renter) occupied housing characteristics. 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 C and D data come from the 2000 Decennial Census and the 2009-2013 American Community Survey. In Panel B standard errors clustered at the county level are reported in parentheses. In Panels C and D, robust standard errors are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Panel E 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 (β) and the standard deviation of idiosynchratic preferences for location (s) of $\beta = .65$ and $s=.4$ respectively. Panel E shows estimates where the change in housings costs is measured using the estimated percentage change in median home prices. We report both the estimated change in amenities and the estimated change in total welfare. The calculations are 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 percent, 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.

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

Table 10: Welfare estimates

	Δ in housing costs = 2.9%		Δ in housing costs = 5.7%	
	WTP for change in:			
	Amenities (1)	Welfare (2)	Amenities (3)	Welfare (4)
Panel A: Annual Impacts per household				
s = 0.4 and β = 0.65	-$1,582	$1,313	-$964	$1,931
s = 0.2 and β = 0.33	-$2,084	$812	-$1,770	$1,125
s = 0.4 and β = 0.33	-$1,901	$995	-$1,587	$1,308
s = 0.6 and β = 0.33	-$1,718	$1,178	-$1,404	$1,491
s = 0.2 and β = 0.65	-$1,765	$1,130	-$1,147	$1,748
s = 0.4 and β = 0.65	-$1,582	$1,313	-$964	$1,931
s = 0.6 and β = 0.65	-$1,399	$1,496	-$781	$2,114
Panel B: Total Aggregate Impacts for Top Quartile Counties (in billions)				
s = 0.4 and β = 0.65	-$53	$44	-$32	$64
s = 0.2 and β = 0.33	-$69	$27	-$59	$38
s = 0.4 and β = 0.33	-$63	$33	-$53	$44
s = 0.6 and β = 0.33	-$57	$39	-$47	$50
s = 0.2 and β = 0.65	-$59	$38	-$38	$58
s = 0.4 and β = 0.65	-$53	$44	-$32	$64
s = 0.6 and β = 0.65	-$47	$50	-$26	$70

Notes: This table reports estimates of the effect of fracing on amenities and welfare in dollars under different assumptions regarding the share of wage and salary income spent on housing (β) and the standard deviation of idiosynchratic preferences for location (s). 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. Columns (1) and (2) report results where the change in housing costs is measured using the estimated percent change in median rents (.029), while Columns (3) and (4) show 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 (Columns (1) and (3)) and the estimated change in total welfare (Columns (2) and (4)). Our preferred parameter values are s=.4 and β = .65. The calculations are converted to dollars using the mean household wage and salary income in top quartile counties of of $34,382 and mean household interest and dividend income in top quartile counties of $3,236. 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.

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

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

12/10/16

---

*Bartik: Department of Economics, MIT, abartik@mit.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 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.

A 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:

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

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] Solving for the consumer's problem and taking logs yields the indirect utility function:

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

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

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

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

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 A.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:

---

[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.

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

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

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

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 ξ is time-invariant. Imposing this assumption, as well as the assumption that each household consumes one housing unit,[3] combining Equations A.3 and A.4, and re-arranging yields the following expression for 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

---

[3] In light of this assumption that each household consumes one-unit of housing, the flexible H_{ijt} in Equation A.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 A.6: $\Delta \ln r_{at} = \frac{\kappa_a}{s+(\beta+s)\kappa_a}\Big((1+s)\Delta\ln w_{at} + \alpha\Delta\ln A_{at}\Big)$. 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 do not translate into changes in rents.

Δw_{at}[4] and changes amenities by some value ΔA_{at}. The change in relative rents is then given by:

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

If we further suppose that fracing does not change amenities or wages in location b, 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:

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

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 homogeneous 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[5], 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 A.4 and re-arrange, yielding:[6]

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

This is a remarkably useful expression because it provides an estimate of willingness-to-pay for

---

[4] 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 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.

[5] 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. It would be interesting to complicate this welfare calculation by estimating separate welfare changes for owners and renters.

[6] 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.

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.[7] 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 before the change in amenities is:

$$\begin{aligned}
\Delta\widehat{V_{at}} \approx \quad & N_{at}\bar{W}_a \times \left(\widehat{\Delta \ln w_{at}} + \alpha\widehat{\Delta \ln A_{at}} - \beta\widehat{\Delta \ln r_{at}}\right) + N_{at} \times \left(\bar{R}_a\widehat{\Delta \ln r_{at}}\right) && \text{(A.8)} \\
& + N_{at}^{\text{own}} \times \bar{Y}_a^{\text{own}} \times \Delta \widehat{\ln y_{at}^{\text{own}}} \\
= \quad & N_{at}\bar{W}_a \times \left(\widehat{\Delta \ln w_{at}} + \alpha\widehat{\Delta \ln A_{at}} - \beta\widehat{\Delta \ln r_{at}}\right) + N_{at} \times \left(\bar{W}_a\beta\widehat{\Delta \ln r_{at}}\right) && \text{(A.9)} \\
& + N_{at} \times \bar{Y}_a \times \widehat{\Delta \ln y_{at}} \\
= \quad & N_{at} \times \left(\bar{W}_a\widehat{\Delta \ln w_{at}} + \bar{W}_a\alpha\widehat{\Delta \ln A_{at}} + \bar{Y}_a \times \widehat{\Delta \ln y_{at}}\right) && \text{(A.10)}
\end{aligned}$$

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.[8]

B Fracing Background Appendix

B.0.1 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 hy-

---

[7] 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 immigrants. Finally, the expression omits profits of landowners who develop new housing units or rent previously vacant housing units.

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

draulic 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)). 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 rapidly improved, and 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)).[9] 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 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, many industry members initially thought that fracing techniques would not work 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.[10] 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.

C 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.[11] Second,

---

[9] 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)).

[10] 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 et al. (2012).

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

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 transported prior to injection.[12] 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)).[13]

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 et al. (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[14] into deep injection wells can cause earthquakes.[15] 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 et al. (2015)). Additionally, it is currently unknown whether injection wells will cause earthquakes large enough to cause significant damage.[16] 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 in-

---

[12] 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

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

[14] 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)).

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

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

creases in visual dis-amenities. The typical production well requires the delivery of between 550 and 1,400 truckloads of water.[17] 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 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.[18]

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 (2015) 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.[19]

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.

D Research Design

D.1 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

---

[17] 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`).

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

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

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, as well as the thermal maturity, porosity, permeability, clay content, and total organic content of the local shale rock (Zagorski et al. (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.[20][21]

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)).[22] 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. In practice, 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.

Appendix Figures G.1 through G.3 illustrate how the Rystad prospectivity scores are used

---

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

[21] See Budzik (2013) for a general discussion of the role played by different geological characteristics in determining the effectiveness of fracking. Zagorski et al. (2012) describes the geological features of the Marcellus and their role in drilling productivity, Covert (2014) 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.

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

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

Appendix Figure G.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.

D.2 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 E.9 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 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.[23] 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 all available information about the prospects for fracing in their respective communities in each of these years.

D.3 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 counties. Second, we follow Imbens and Rubin (2015) and select covariates and second-order terms to include in the propensity score

---

[23] 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

using an iterative procedure[24], resulting in a list of 17 first-order and second-order terms[25]. 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[26]. Fifth, we construct a control sample for each shale play using the unique matched counties for that play[27]. 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 [28]. Note that 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.

D.3.1 Propensity Score Matching Results

Tables 14 through 19 report estimates of Equation 5.2 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 credible empirical design, we report the estimates and

---

[24] 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

[25] 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

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

[27] 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

[28] 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 and ϕ_p^{pscore} is the value of ϕ when the control group is the pscore matched sample.

discuss them below for completeness.

Starting with 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, Table 16 providse 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 9.8 percent, roughly twice the estimates in Table 5. 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.

Table 15 the matching estimator in column (1) again suggests that our main estimates in Table 4 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 4. 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 Figure G.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 8 were imprecise. The matching estimates suggest one potential reason for this imprecision: the possible presence of spillovers to nearby counties. In Table 19 we find a positive and statistically significant increase in crime for both the top-quartile and non-top quartile counties. The signs of the estimated effect of fracing on total crime and property crime are sensitive to the specification, and imprecisely estimated.

The matching estimator for top quartile counties reported in Table 17 in general tell a similar story to our main estimates in Table 6 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 6, 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 larger increase in "other tax revenue" and charges revenue in column (1) compared to Table 6.

Finally, turning to housing price estimates using the matching estimator in Table 18, 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.

E Data Appendix

E.1 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.[29] Prospectivity values for each shale play the county intersects. We then compute indicators for being in the top quartile and top octile 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.[30]

---

[29] 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.

[30] 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

E.2 Hydrocarbon Production

Oil and gas production data for 1992 through 2011 come from data purchased from Drilling Info, Inc (2012) and used with permission. The HPDI 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.[31] 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.[32] 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.[33] Finally, using EIA data on the heat rates of different types of power plants, we also convert oil and gas production into GWh.[34]

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.

E.3 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[35] 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

---

[31] 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.

[32] 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.

[33] 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

[34] 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.

[35] 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

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 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 7 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 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.[36] 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[37] 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. Similarly, Column (3) reports the difference in the percentage change in the given housing outcome between 2000 and 2007/11[38] when computed using the universe of specified

---

[36] 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.

[37] 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.

[38] 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

housing units for both years compared to using the full universe for both years.

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.

E.4 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), US Department of Agriculture (2014)). 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.5 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) (2014)). 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

---

definitions.

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).

E.6 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 in(out)-migration $\text{rate}_{it} = \text{in(out)-migration}_{it}/\text{pop}_{i,t-1}$. We can also compute net-migration as $\text{in-migration}_{it} - \text{out-migration}_{it}$.

E.7 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 dicenniel censii and the pooled 2009-2013 ACS for % of the county that is urban, employment-to-population ratio, % of people with less than a high school diploma, fraction with a college degree or more, % manufacturing employment, and % mining employment.

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.

E.8 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[39],

---

[39] 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.

robbery, burglary, larceny, and motor-vehicle theft, as well as the population that the law enforcement agency covers. 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.[40] 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 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.[41] 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 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 a clean, agency-year dataset.

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

---

[40] 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

[41] 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.

crime value for all[42] index crimes in every year from 1992[43] to 2013.[44] 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.

E.9 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.[45] We determined first-frac dates using two different definitions of the start of fracing. One definition is the first date that a well using modern hydraulic fracturing techniques was completed and showed promising production in the given shale formation. Our second definition 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.

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

---

[42] 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.

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

[44] 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.

[45] 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.

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 worked really well 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. 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 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.

E.9.1 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.[46]

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 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

---

[46] 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."

success (Nordeng (2010) and Williams (2004a))[47]. 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 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)).[48] 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 25 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.

E.9.2 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

---

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

[48] 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)).

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)).[49] 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 slickwater frac techniques with horizontal wells that all of the key elements for producing from the Barnett were put together.[50] 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.

E.9.3 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[51]. 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

---

[49] 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)).

[50] 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.

[51] 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 (2011a)). 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 200s that no one knew about. We and EOG (Resources) were pushing this formation way before these other guys even thought about it." (Toon (2011a))

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.

E.9.4 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 would be expected from its reservoir properties" (Williams (2004b)) and hypothesized that the Wedington sand wells were drawing gas from the surrounding Fayetteville shale formation (Williams (2004b)). 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.

E.9.5 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.

E.9.6 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.

E.9.7 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 facing 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 (2010c) 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.[52] 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.

E.9.8 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 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.

E.9.9 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 et al. (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 (2011b)).

---

[52] 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)).

E.9.10 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.[53,54] (Dutton et al. (2005) and Gold (2014))

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. 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)).[55]

Around the same time, in the eastern part of the Permian in the Midland sub-basin, firms

---

[53] 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).

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

[55] 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.

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 formations 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 (2010a), Toon (2011b)). 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)),[56] 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.

E.9.11 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[57] part of the Utica (Harper (2011)).[58] 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 (2011b)). Later in 2011, Chesapeake announced good results from its initial horizontal wells in the Utica, particularly the Buell well (Chesapeake Energy Corporation (2011a)). 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

---

[56] 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.

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

[58] 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)).

March 22, 2010 as the completion date for the first modern well and Chesapeake's July 28, 2011 announcement as the first publicity date.

E.9.12 Woodford: Anadarko

The Woodford Anadarko[59] lies in central and western Oklahoma, an area with an extensive history of conventional drilling. Devon Energy started exploring 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.

E.9.13 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 (2007b)). 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 (2007a)). 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 (2010b)

---

[59] 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 (2010b) or Cardott (2013)). 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.

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.

E.9.14 Woodford: Arkoma

The Woodford Arkoma is the portion of the Woodford shale in the Arkoma basin in eastern Oklahoma.[60] The Woodford Arkoma is particularly attractive 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.

E.10 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 Urban Consumers (CPI-U) price index for all items.

---

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

F Appendix Tables

Appendix Table 1: First-frac dates

Play (1)	Basin (2)	First Production Date (3)	First Publicity Date (4)	Rystad Data Available (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 rable 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 tau=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)	267**	242**	314***
	(107)	(94)	(99)
Top Tercile(Mean Rystad)	233***	209***	250***
	(82)	(74)	(77)
Top Octile(Mean Rystad)	499***	453***	535***
	(162)	(143)	(146)
Panel B. Log(Total Income)			
Top Quartile(Max Rystad)	0.069**	0.091***	0.073***
	(0.028)	(0.030)	(0.027)
	[0.036]	[0.035]	[0.031]
Top Tercile(Max Rystad)	0.068***	0.096***	0.066***
	(0.025)	(0.028)	(0.024)
Top Octile(Max Rystad)	0.067	0.099**	0.061
	(0.042)	(0.042)	(0.040)
Top Quartile(Mean Rystad)	0.018	0.023	0.015
	(0.027)	(0.026)	(0.026)
Top Tercile(Mean Rystad)	0.037	0.044**	0.028
	(0.023)	(0.022)	(0.022)
Top Octile(Mean Rystad)	0.036	0.046	0.037
	(0.032)	(0.033)	(0.029)
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)	0.048***	0.023**	0.047***
	(0.011)	(0.011)	(0.011)
Top Tercile(Mean Rystad)	0.046***	0.023**	0.045***
	(0.010)	(0.010)	(0.010)
Top Octile(Mean Rystad)	0.061***	0.023	0.061***
	(0.012)	(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 102(102)[102], 75(75)[75], and 39(39)[39] balanced sample counties in Columns (1) and (2) and 111(111)[111], 82(82)[82], 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 216(215)[302], 243(242)[329], and 279(278)[365] balanced sample counties in Columns (1) and (2) and 235(234)[293], 264(263)[322], and 304(303)[362] balanced sample counties in Column (3).

Appendix Table 5: Play specific estimates: additional outcomes

	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. 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													
	70	275***	390**	6	214	409***	-219***	131	-32	104	F-stat	10.5	210***
	(225)	(90)	(198)	(52)	(214)	(131)	(29)	(126)	(93)	(86)	p-value	0.00	(54)
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)													
	0.054*	0.294***	0.031*	-0.063***	0.109	-0.023	-0.111***	-0.045	0.019	0.097**		6.7	0.297**
	(0.029)	(0.085)	(0.018)	(0.014)	(0.072)	(0.018)	(0.026)	(0.099)	(0.076)	(0.045)		0.00	(0.124)
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)													
	0.080**	0.360***	-0.027	-0.019	0.339***	0.032	0.076	0.115**	0.006	-0.074	F-stat	4.9	0.429***
	(0.038)	(0.071)	(0.058)	(0.036)	(0.109)	(0.023)	(0.069)	(0.051)	(0.102)	(0.092)	p-value	0.00	(0.136)
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 (1)	Bakken (2)	Barnett (3)	Fayetteville (4)	Haynesville (5)	Marcellus (6)	Woodford, Anadarko (7)	Woodford, Ardmore (8)	Woodford, Arkoma (9)	Permian Plays (10)	Eagle Ford (12)
Panel A: Δ in housing costs measured using change in mean rents											
A1. Change in Amenities and Welfare per Household (dollars)											
Change in amenities	-$1,582	-$3,260	-$684	$2,638	-$2,382	-$1,788	-$4,460	$2,147	-$684	-$3,829	-$775
Change in welfare	$1,313	$8,203	$992	$4,892	$289	$1,279	-$1,930	$1,615	$1,965	$2,113	-$104
A2: Total Change in Amenities and Welfare (billions of dollars)											
Change in amenities	-$52.8	-$1.3	-$2.7	$0.5	-$2.2	-$45.6	-$1.5	$1.3	-$0.1	-$4.5	-$1.1
Change in welfare	$43.8	$3.3	$4.0	$1.0	$0.3	$32.6	-$0.7	$1.0	$0.3	$2.5	-$0.2
Panel B: Δ in housing costs measured using change in mean home values											
B1. Change in Amenities and Welfare per Household (dollars)											
Change in amenities	-$964	-$2,395	-$1,518	$631	-$4,455	-$484	-$3,882	$729	-$1,466	-$5,409	-$1,543
Change in welfare	$1,931	$9,068	$157	$2,884	-$1,784	$2,583	-$1,352	$197	$1,182	$533	-$873
B2: Total Change in Amenities and Welfare (billions of dollars)											
Change in amenities	-$32.2	-$1.0	-$6.1	$0.1	-$4.1	-$12.3	-$1.3	$0.4	-$0.2	-$6.4	-$2.2
Change in welfare	$64.4	$3.7	$0.6	$0.6	-$1.6	$65.9	-$0.5	$0.1	$0.2	$0.6	-$1.3
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 (β) and the standard deviation of idiosynchratic preferences for location (s) of β = .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, while 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 are 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 percent, 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)
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)
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 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 and state-year fixed effects. The reported estimates in Panel B 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 and 4 years after the start of fracing for Panel B. Panels C and D 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, and also include play and state fixed-effects. Panel D also includes controls for changes in average county owner(renter) occupied housing characteristics. 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 C and D data come from the 2000 Decennial Census and the 2009-2013 American Community Survey. In Panel B standard errors clustered at the county level are reported in parentheses. In Panels C and D, robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

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)	0.130***	0.152***	0.133***
	(0.035)	(0.037)	(0.034)
	[0.051]	[0.051]	[0.042]
Top Tercile(Max Rystad)	0.127***	0.159***	0.127***
	(0.030)	(0.034)	(0.029)
Top Octile(Max Rystad)	0.121**	0.148***	0.126***
	(0.053)	(0.054)	(0.047)
Top Quartile(Mean Rystad)	0.055*	0.057*	0.053*
	(0.033)	(0.032)	(0.032)
Top Tercile(Mean Rystad)	0.078***	0.082***	0.071***
	(0.028)	(0.027)	(0.027)
Top Octile(Mean Rystad)	0.071*	0.083**	0.077**
	(0.040)	(0.041)	(0.038)
Panel B. Log(Employment)			
Top Quartile(Max Rystad)	0.054*	0.068**	0.068**
	(0.029)	(0.030)	(0.029)
	[0.030]	[0.031]	[0.029]
Top Tercile(Max Rystad)	0.072***	0.092***	0.079***
	(0.027)	(0.029)	(0.027)
Top Octile(Max Rystad)	0.044	0.066	0.061
	(0.042)	(0.041)	(0.042)
Top Quartile(Mean Rystad)	0.031	0.038	0.035
	(0.030)	(0.029)	(0.030)
Top Tercile(Mean Rystad)	0.043*	0.051**	0.039
	(0.025)	(0.025)	(0.025)
Top Octile(Mean Rystad)	0.068**	0.075**	0.079***
	(0.027)	(0.030)	(0.028)
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)	0.200	0.115	0.146
	(0.163)	(0.172)	(0.185)
Top Tercile(Mean Rystad)	0.241*	0.169	0.205
	(0.140)	(0.149)	(0.161)
Top Octile(Mean Rystad)	0.099	0.008	-0.037
	(0.210)	(0.224)	(0.248)
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 102[95], 75[68], and 39[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 7, 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 (1)	Change Computed Using Specified Units as Base Year Compared to Full Universe (2)	Change for Specified Units Compared to Change in Full Universe (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 1997 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 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 quartile and 253 outside top quartile counties are in the balanced sample. Panel D includes observations from all 385 counties over shale plays with non-missing school enrollment data from all districts in 1997, 2002, and 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)	54**	21***
	(22)	(8)
t*1(Exposed)		
t*1(Exposed)*1(Post)	135***	12*
	(36)	(7)
Fracing Exposure Effect at tau=3	460***	58***
	(108)	(22)
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 tau=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 12,961 observations from 997 county-play combinations, of which 65 Rystad top quartile and 553 unique pscore-matched counties are in the balanced sample. Column (2) includes 15,405 observations from 1,185 total county-play combinations, of which 253 Rystad quartiles one through three and 553 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 tau=4	0.098***	0.054***
	(0.022)	(0.010)
Panel B: Income		
Log(Total Income)		
Fracing Exposure Effect at tau=4	0.097***	0.056***
	(0.027)	(0.011)
B1. Log(Total Wage/Salary Income): 56 percent of total personal income		
Fracing Exposure Effect at tau=4	0.166***	0.081***
	(0.035)	(0.011)
B2. Log(Total Rents/Dividends): 19 percent of total personal income		
Fracing Exposure Effect at tau=4	0.166***	0.103***
	(0.030)	(0.014)
B3. Log(Total Transfers): 10 percent of total personal income		
Fracing Exposure Effect at tau=4	0.010	0.015***
	(0.008)	(0.006)
B4. Log(Total Proprieter's Income): 18 percent of total personal income		
Fracing Exposure Effect at tau=4	-0.039	0.001
	(0.065)	(0.044)
Panel C: Migration		
C1. Log(In-Migration)		
Fracing Exposure Effect at tau=4	0.119***	0.121***
	(0.043)	(0.018)
C2. Log(Out-Migration)		
Fracing Exposure Effect at tau=4	0.014	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 15,952 observations from 997 total county-play combinations, of which 65 Rystad top quartile and 553 unique pscore-matched counties are in the balanced sample. Column (2) include 18,944 observations from 1,184 total county-play combinations, of which 252 Rystad quartiles one through three counties and 553 unique pscore-matched counties are in the balanced sample.

Panel B4, Column (1) includes 15,664 observations from 979 county-play combinations, of which 60 Rystad top quartile and 546 unique pscore matched counties are in the balanced sample. Panel B4, Column (2) includes 18,496 observations from 1,156 county-play combinations, of which 237 Rystad quartile one through three counties and 546 unique pscore matched counties are in the balanced sample.

Panel C, Column (1) includes 12,870 observations from 990 county-play combinations, of which 63 Rystad top quartile and 551 unique pscore-matched counties are in the balanced sample. Panel C, Column (2) includes 15,275 observations from 1,175 county-play combinations, of which 248 Rystad quartile one through three and 551 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)	0.058***	0.010
	(0.019)	(0.011)
A2. Employment-to-Population Ratio	0.011	-0.008
	(0.008)	(0.005)
A3. Unemployment Rate	-0.014***	-0.010***
	(0.004)	(0.004)
Panel B: Household Income:		
B1. Log(Median Real Household Income)	0.063***	0.013
	(0.013)	(0.012)
B2. Log(Mean Real Household Wage and Salary Income)	0.074***	0.010
	(0.017)	(0.014)
B3. Log(Mean Real Rent and Dividend Income)	0.105***	0.065**
	(0.035)	(0.027)
Panel C: Population:		
C1. Log(Population)	0.007	-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 1,397 total counties, of which 65 Rystad top quartile and 552 pscore-matched sample counties are in the balanced sample. Panels A1, B, and C, Column (2) include observations from 1,613 total counties, of which 253 Rystad quartiles one through three and 552 pscore-matched sample counties are in the balanced sample.

Panels A2 and A3, Column (1) include observations from 1,396 county-play combinations, of which 64 Rystad top quartile and 552 pscore-matched counties are in the balanced sample. Panels A2 and A3, Column (2) include observations from 1,613 county-play combinations, of which 253 Rystad quartiles one through three and 552 unique pscore-matched counties are in the balanced sample.

Appendix Table 17: Impact of fracing on local government revenues and expenditures: Propensity-score estimates

	(1)	(2)
Panel A: Log(Total Expenditures): 2012 - 2002		
	0.149***	0.021
	(0.036)	(0.021)
A. Log(Direct Expenditures)		
	0.131***	0.008
	(0.035)	(0.021)
A1. Direct Expenditures by Type		
A1a. Log(Current Operating Expenditure):		
[84%]	0.098***	-0.009
	(0.028)	(0.017)
A1b. Log(Capital Outlays):		
[12%]	0.325**	0.144*
	(0.143)	(0.082)
A2. Direct Expenditures by Purpose		
A2a. Log(Education Expenditures):		
[48%]	0.053*	0.029
	(0.032)	(0.020)
A2b. Log(Public Safety Expenditures):		
[8%]	0.253***	0.058*
	(0.061)	(0.032)
A2c. Log(Welfare and Hospital Expenditures):		
[10%]	0.298**	0.056
	(0.134)	(0.108)
A2d. Log(Infrastructure and Utility Expenditures):		
[18%]	0.276***	0.034
	(0.071)	(0.039)
A2e. Log(Other Expenditures):		
[16%]	0.106*	-0.016
	(0.062)	(0.038)
Panel B: Log(Total Revenues): 2012 - 2002		
	0.164***	0.009
	(0.035)	(0.018)
B1. Revenues by Type		
B1a. Log(Property Tax Revenues):		
[24%]	0.168***	0.034*
	(0.045)	(0.021)
B1b. Log(Sales Tax Revenues):		
[4%]	0.467***	-0.132*
	(0.116)	(0.073)
B1c. Log(Other Tax Revenues):		
[2%]	0.271*	0.233***
	(0.146)	(0.090)
B1d. Log(Intergovernmental Revenues):		
[42%]	0.142*	0.041
	(0.080)	(0.034)
B1e. Log(Charges Revenues):		
[14%]	0.173**	0.078
	(0.077)	(0.052)
B1f. Log(Other Revenues):		
[14%]	0.219***	-0.042
	(0.067)	(0.041)
Panel C: Government Balance Sheets		
C. Net Financial Position as Share of Revenues		
	-0.047	-0.026
	(0.070)	(0.040)
Panel D: Log(Elem/Sec Education Spending per Pupil)		
	0.044	0.043**
	(0.034)	(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 1,386 county-play combinations, of which 65 Rystad top quartile and 553 unique pscore-matched counties are in the balanced sample. Panel D, Column (1) includes 1,220 county-play combinations with non-missing school enrollment data for all districts in 1997, 2002, and 2012, of which 61 Rystad top quartile and 488 unique pscore-matched counties are in the balanced sample.

Column (2), Panels A, B, and C, include 253 county-play combinations, of which 553 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 488 unique pscore-matched counties are in the balanced sample.

Appendix Table 18: Impact of fracing on housing outcomes: Propensity-score estimates

	(1)	(2)
Panel A: House Values		
A1. Log(Median House Value)		
	-0.019	-0.058***
	(0.013)	(0.015)
A2. Log(Mean Housing Value)		
	-0.001	-0.036***
	(0.014)	(0.013)
A3. Log(Mobile Housing Units: Median Housing Value)		
	0.057	-0.001
	(0.043)	(0.029)
Panel B: Rental Prices		
B1. Log(Median Rental Price)		
	-0.001	-0.020**
	(0.013)	(0.009)
B2. Log(Mean Rental Price)		
	0.006	-0.021**
	(0.014)	(0.008)
Panel C: Housing Quantities		
C1. Log(Total Housing Units)		
	-0.017	-0.027***
	(0.013)	(0.008)
C2. Log(Total Mobile Homes)		
	0.041	0.019
	(0.028)	(0.016)
C3. Share of Housing Units Vacant		
	-0.015***	-0.004
	(0.005)	(0.003)
C4. Log(Acres of Agricultural Land)		
	-0.131	-0.012
	(0.145)	(0.079)
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 1,385 county-play combinations, of which 65 Rystad top quartile and 553 unique pscore-matched sample counties are in the balanced sample. Column (1), Panel C4 contains observations from 1,178 county-play combinations, of which 53 Rystad top quartile and 467 unique pscore-matched sample counties are in the balanced sample.

Column (2), Panels A1, A3, B1, B2, C1, C2, and C3 contain observations from 1,601 county-play combinations, of which 253 quartiles one through three and 553 unique pscore-matched sample counties are in the balanced sample. Column (2), Panel C4 contains observations from 1,361 county-play combinations, of which 211 quartiles one through three and 467 unique pscore-matched sample counties are in the balanced sample.

Appendix Table 19: Impact of fracing on crime

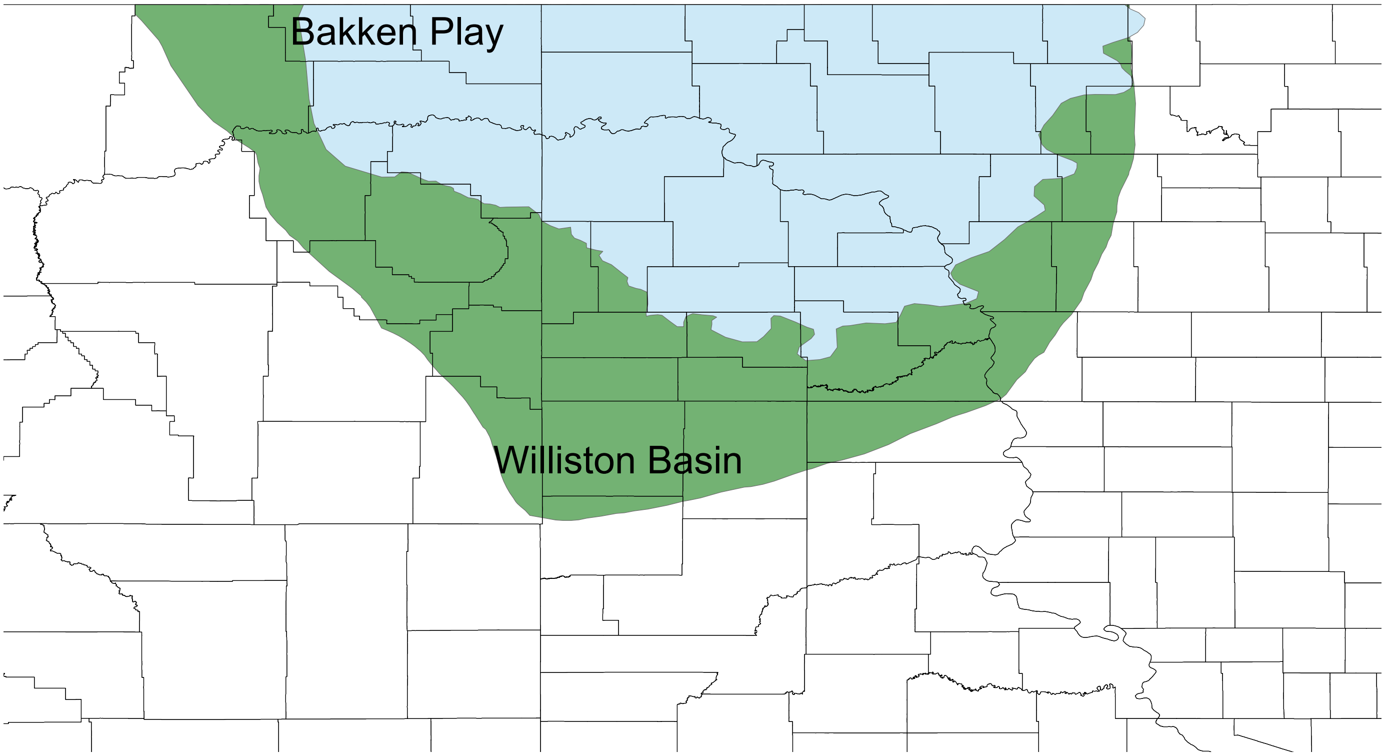
	(1)	(2)
Panel A. Log(Total Crime)		
Top Quartile Effect at tau=5	0.120	0.125**
	(0.086)	(0.064)
Panel B: Log(Violent Crime)		
Top Quartile Effect at tau=5	0.307**	0.159**
	(0.131)	(0.076)
Panel C: Log(Property Crime)		
Top Quartile Effect at tau=5	0.097	0.104
	(0.090)	(0.067)
Fracing Exposure Group	Top 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 9,330 observations from 622 county-play combinations, of which 56 Rystad top quartile and 358 unique pscore-matched sample counties are in the balanced sample. Column (2) includes 11,640 observations from 776 county-play combinations, of which 210 Rystad quartile one through three and 358 unique pscore-matched sample counties are in the balanced sample.

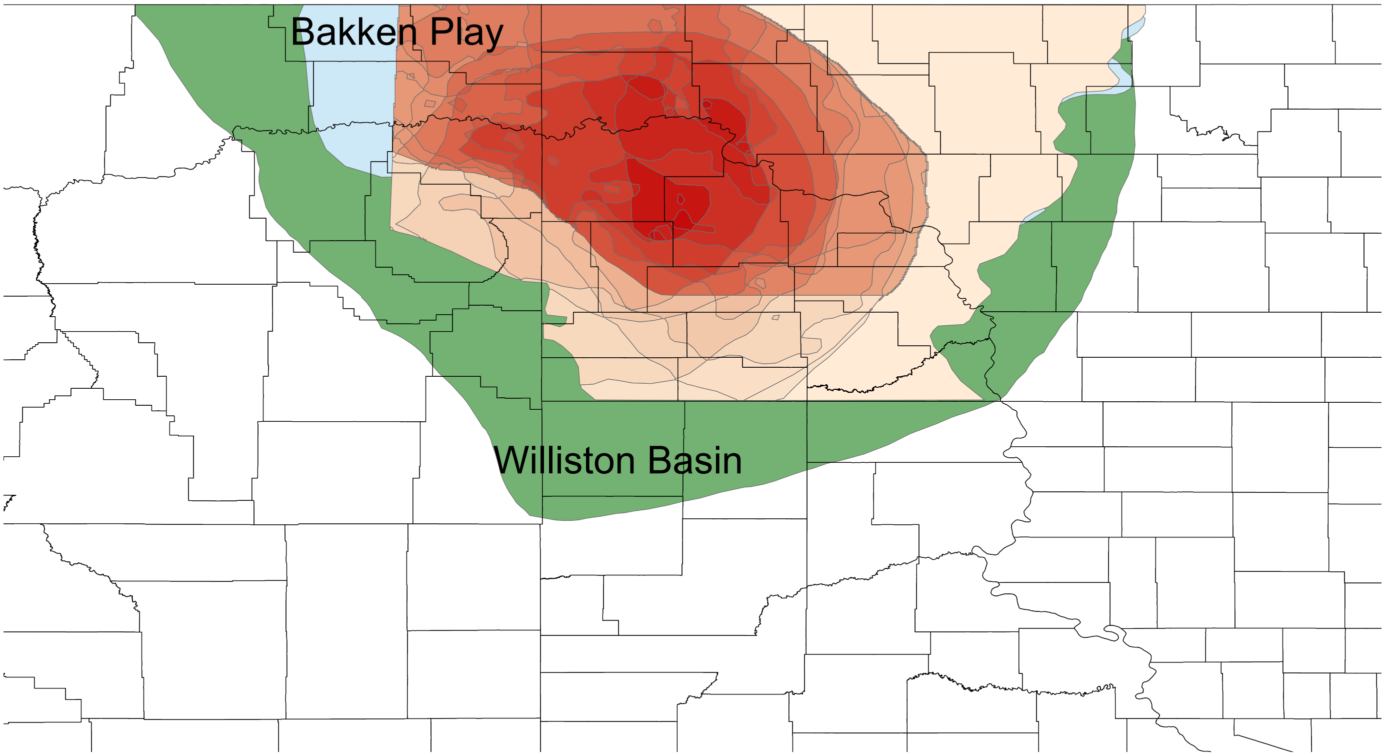
G Appendix Figures

Appendix Figure G.1: Bakken play and Williston Basin



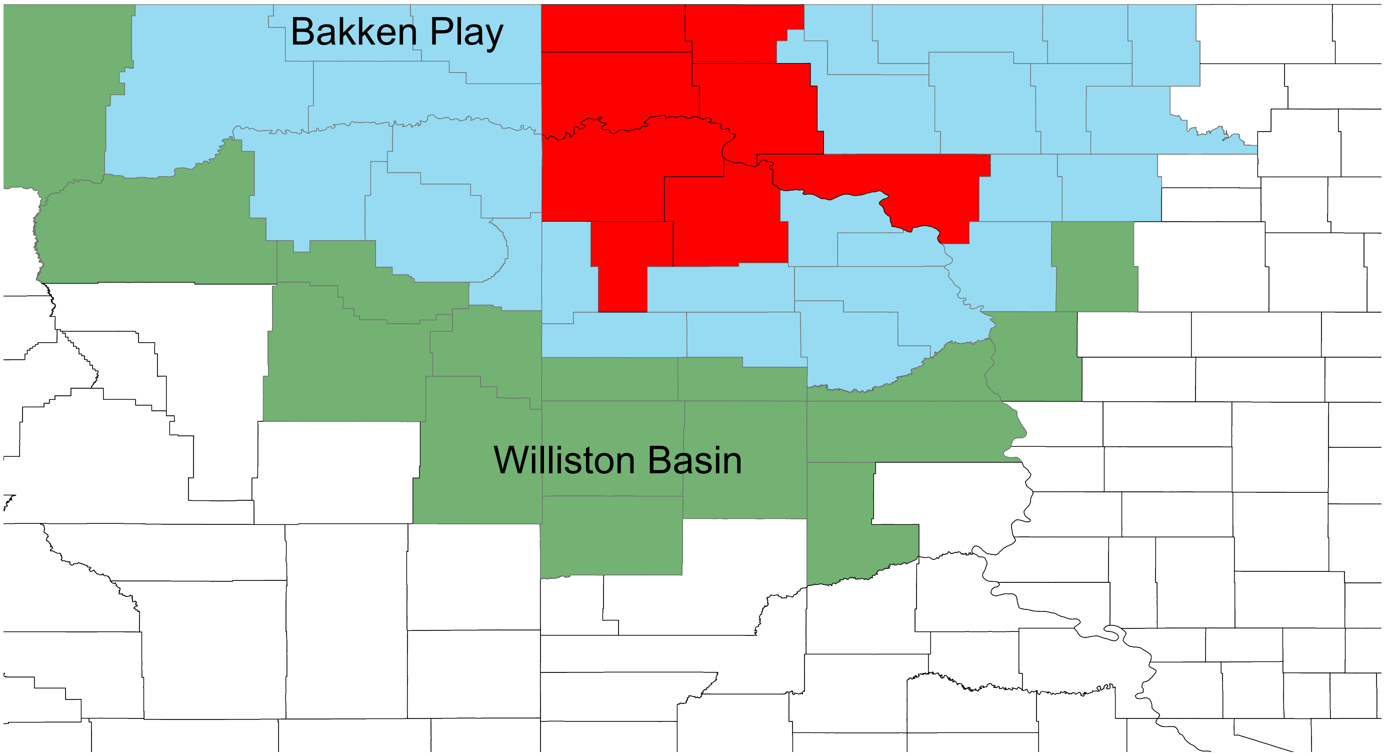
Notes: This figure overlays the Williston Basin and Bakken shale play over nearby counties in North Dakota, Montana, Wyoming, and South Dakota. The Williston Basin is shown in green and the Bakken shale play is shown in blue. Shapefiles for US shale basins and plays comes from the Energy Information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

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



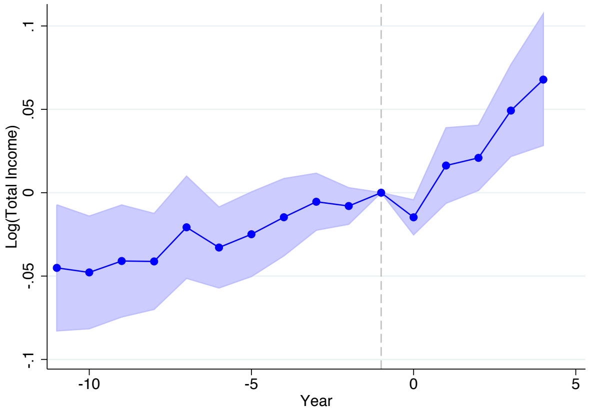
Notes: This figure overlays the Williston Basin, Bakken shale play, and Rystad Prospectivity scores over nearby counties in North Dakota, Montana, Wyoming, and South Dakota. The Williston Basin is shown in green and the Bakken shale play is shown in blue. Darker shades of red correspond to higher Rystad 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 G.3: Bakken play with county assignments



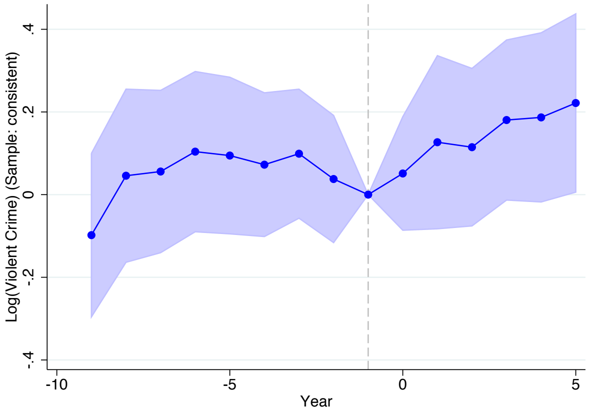
Notes: This figure shows county-assignments for counties in the area around the Williston Basin. Counties shaded red are in the top quartile of Rystad Prospectivity for the Bakken play. Counties shaded blue are outside of the Rystad top quartile, but within the Bakken shale play boundaries. Counties 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 G.4: Event study analysis of county-level total income



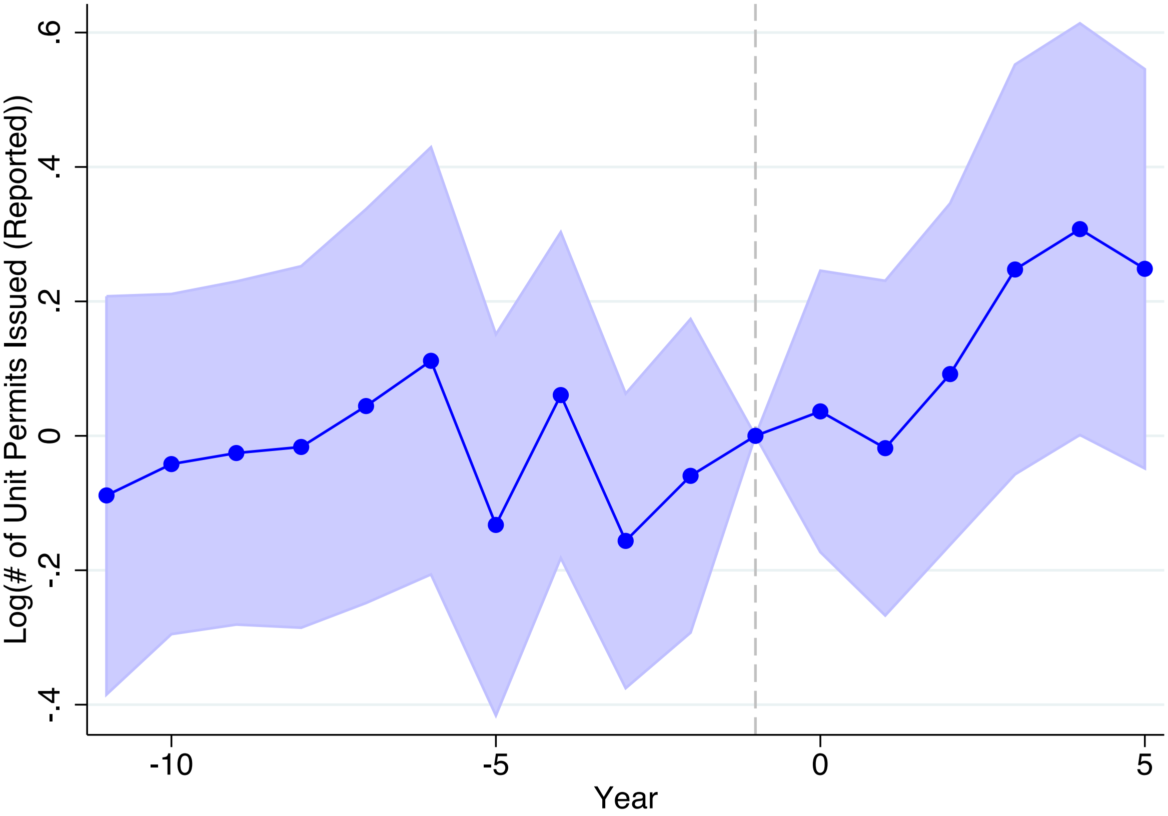
Notes: This figure plots results from an event-study analysis of the difference in log(total income) between high-fracing potential counties and other counties in shale plays before and after fracing began. The reported coefficients come from fitting a modified version of Equation 5.1 where we interact 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) (2014). 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 G.5: Event study analysis of county-level violent crime



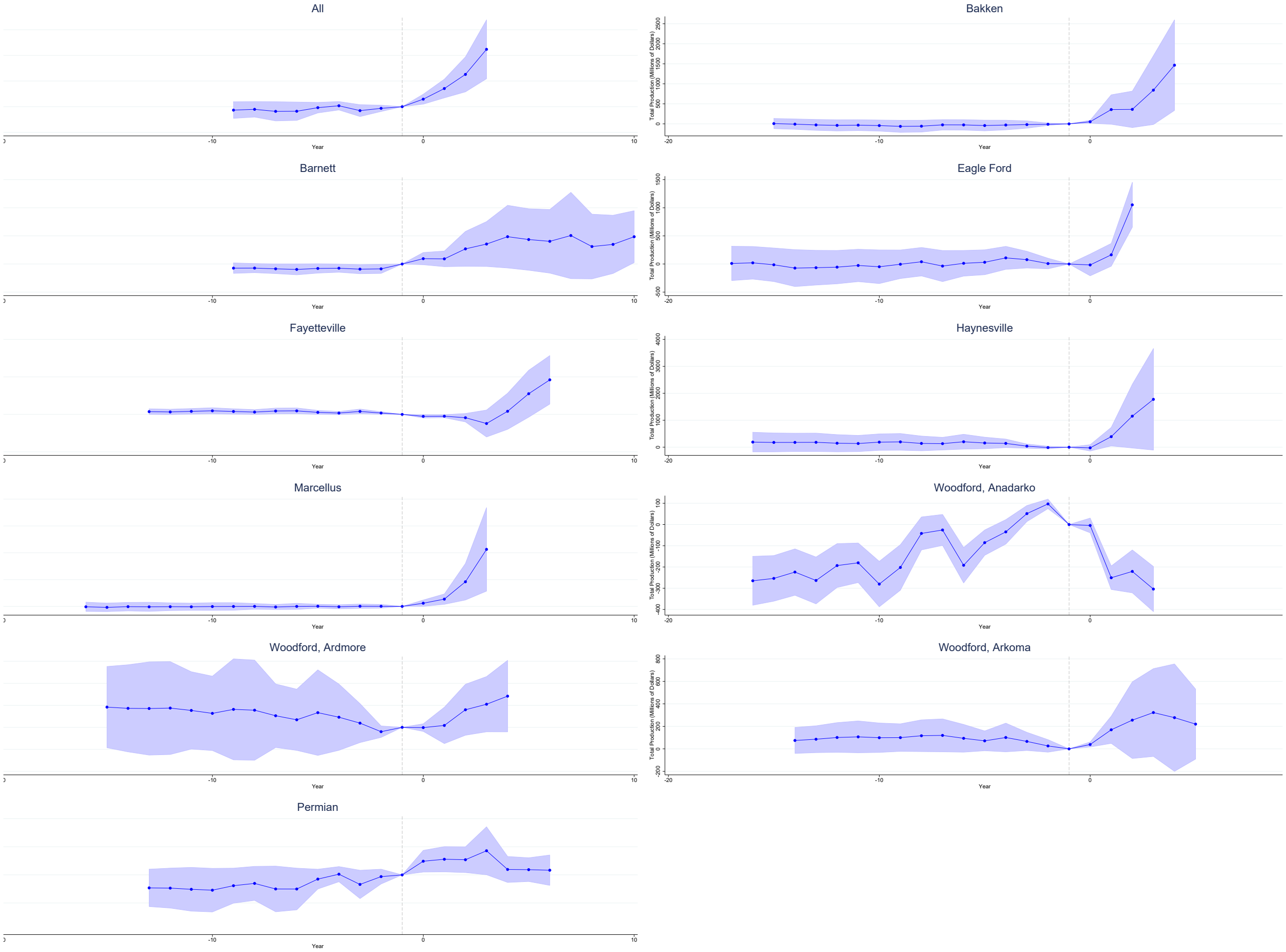
Notes: This figure plots results from an event-study analysis of the difference in the county-level log(violent crimes) between high-fracing potential counties and other counties in shale plays before and after fracing began. The reported coefficients come from fitting a modified version of Equation 5.1 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 G.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 5.1 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 G.7: Event study analysis of county-level hydrocarbon production by play

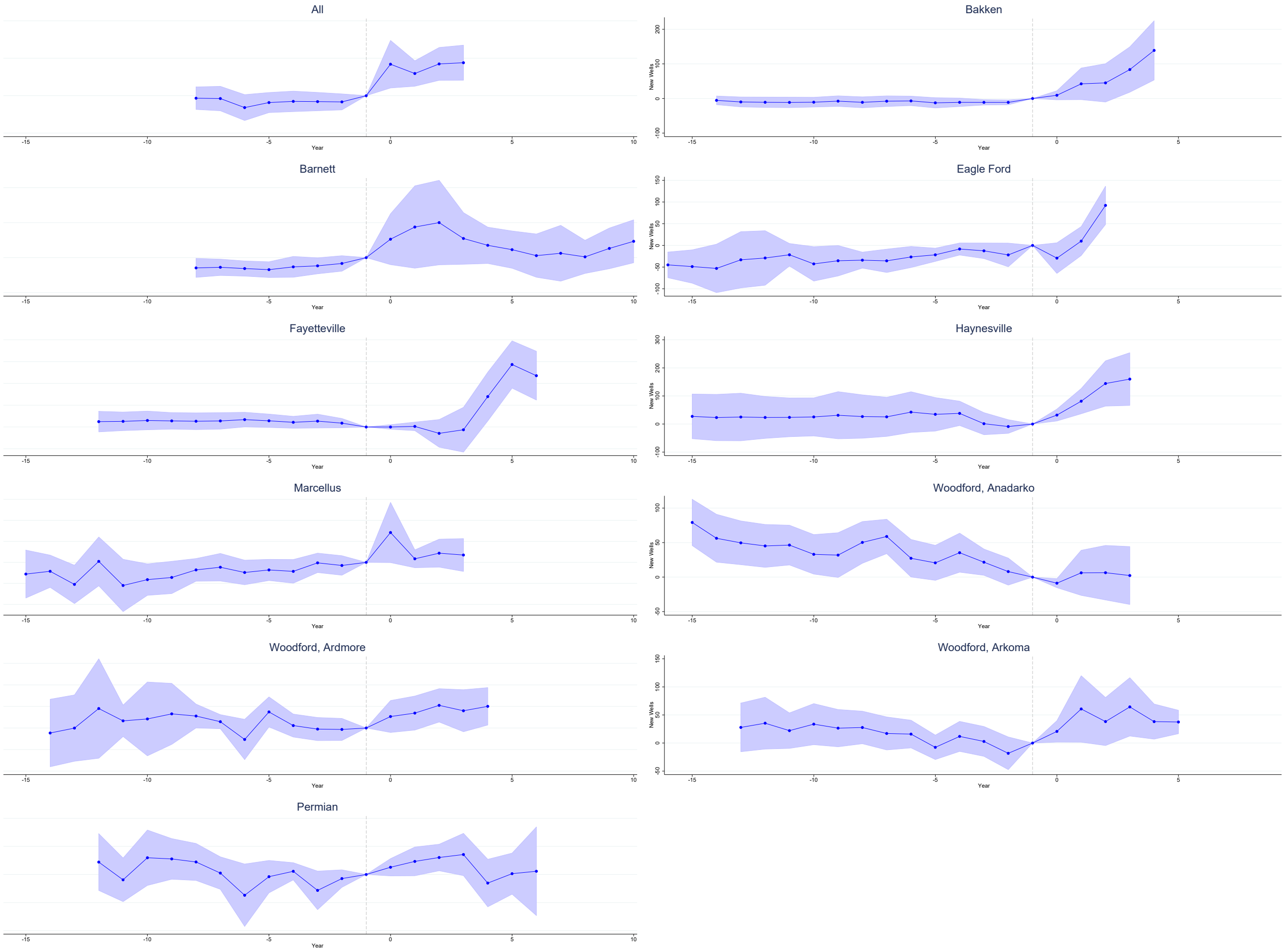


Notes: This figure plots results from an event-study analysis of the difference in the county-level value of hydrocarbon production between high-fracing potential counties and other counties in shale plays before and after fracing began separately by shale play. The top-left figure reports results for 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 5.1 where we interact $1[\text{Rystad Top Quartile}]_c$ 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 G.8: Event study analysis of new producing well by play

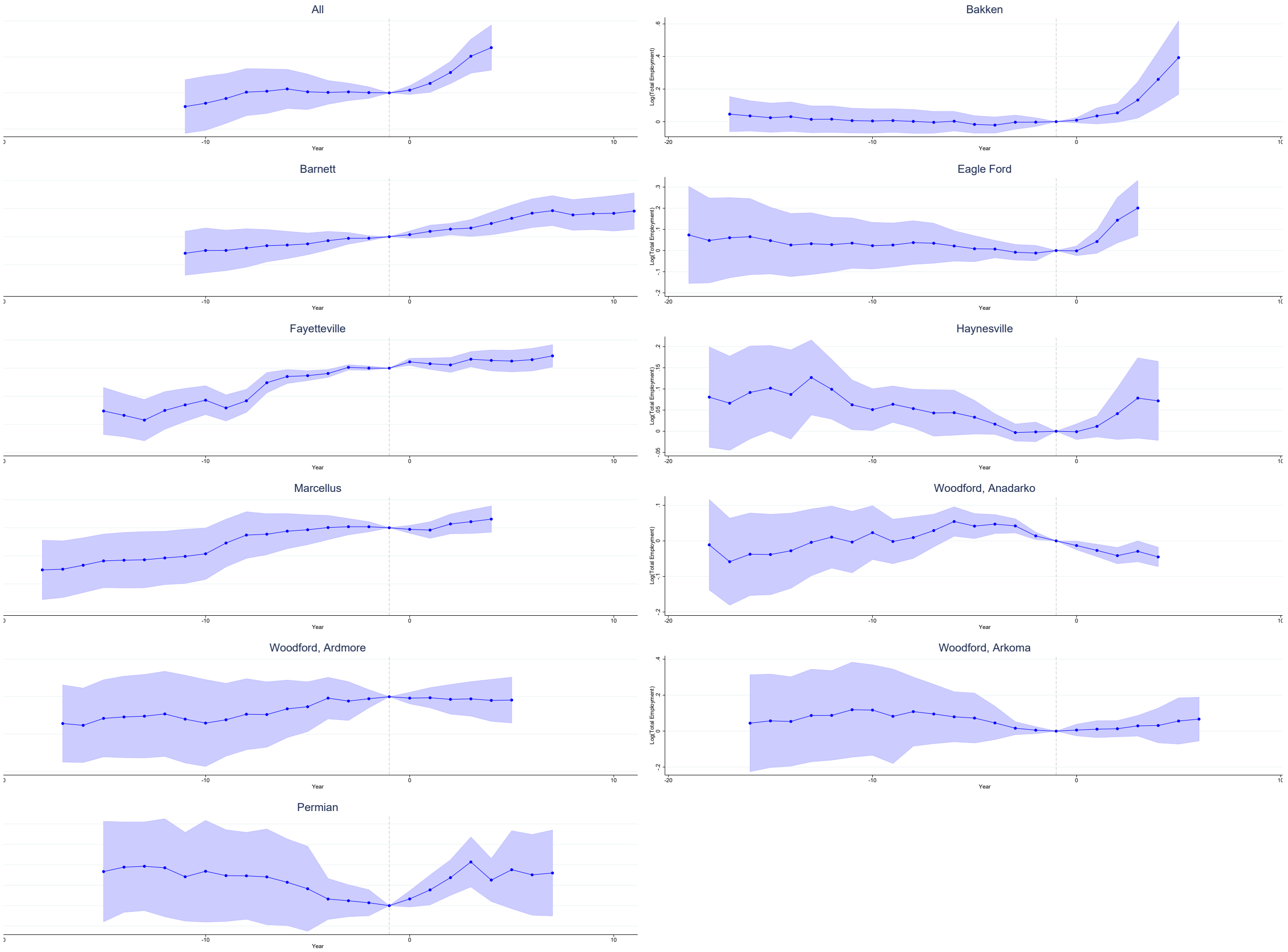


Notes: This figure plots results from an event-study analysis of the difference in the county-level total number of new wells drilled between high-fracing potential counties and other counties in shale plays before and after fracing began separately by shale play. The top-left figure reports results 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 5.1 where we interact $1[\text{Rystad Top Quartile}]_c$ 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 G.9: Event study analysis of county-level employment by play


All

Bakken

Barnett

Eagle Ford

Fayetteville

Haynesville

Marcellus

Woodford, Anadarko

Woodford, Ardmore

Woodford, Arkoma

Permian

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 5.1 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 ?. 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.

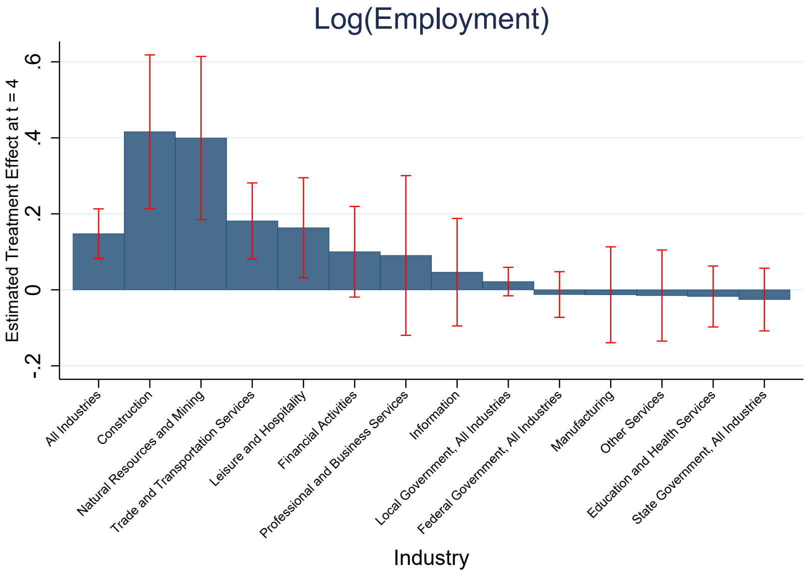
Appendix Figure G.10: Event study analysis of county-level log(violent crime) by play

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 5.1 where we interact $1[\text{Rystad Top Quartile}]_c$ 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 G.11: Employment effects by industry: pscore-matching



Notes: This figure plots estimates of the effect of fracing on employment by industry five years after the start of fracing. Each bar reports results of fitting Equation 5.2 for the given industry, which corresponds to Column (4) in the tables. Equation 5.2 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.

References

Albouy, D. (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, 38.

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

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

——— (2007b): *Bankers Petroleum Discovers Natural Gas in Oklahoma and Provides Update on US Exploration Program*, Calgary.

Boyd, D. T. (2008): "Oklahoma 2007 Drilling Highlights," *Shale Shaker: The Journal of the Oklahoma Geological Society*, 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," Tech. rep., 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," Tech. rep.

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

Carter, K. M., J. a. Harper, K. W. Schmid, and J. Kostelnik (2011): "Unconventional natural gas resources in Pennsylvania: The backstory of the modern Marcellus Shale play," *Environmental Geosciences*, 18, 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.

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

——— (2011b): *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, 1–12.

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

Connelly, K., D. Barer, and Y. Skorobogatov (2015): *How Oil and Gas Disposal Wells Can Cause Earthquakes.*

Covert, T. R. (2014): "Experiential and Social Learning in Firms : The Case of Hydraulic Fracturing in the Bakken Shale JOB MARKET PAPER," 1–43.

Crooks, E. (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.

Davis, L. W. and L. Kilian (2011): "The Allocative Cost of Price Ceilings in the US Residential Market for Natural Gas," *The Journal of Political Economy*, 119, 212–241.

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, R. (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, L. (2008): "Louisana Play a 'Company Maker'?" *American Association of Petroleum Geosciences: Explorer.*

Dutton, S. P., E. M. Kim, R. F. Broadhead, W. D. Raatz, C. L. Breton, S. C. Ruppel, and C. Kerans (2005): "Play analysis and leading-edge oil-reservoir development methods in the Permian basin: Increased recovery through advanced technologies," *AAPG Bulletin*, 89, 553–576.

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

Ellsworth, W. L., A. L. Llenos, A. F. Mcgarr, A. J. Michael, J. L. Rubinstein, C. S. Mueller, M. D. Petersen, and E. Calais (2015): "Increasing seismicity in the US midcontinent: Implications for earthquake hazard," *The Leading Edge*, 34, 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, T. and G. G. Lash (2008): "Unconventional Natural Gas Reservoir Could Boost US Supply," Tech. rep., 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," Tech. Rep. June.

EOG Resources (2007): *EOG Resources Reports 2006 Results and Increases Dividend.*

——— (2010): *EOG Resources Announces 2010 Analyst Conference Highlights*, Houston.

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

Fetzer, T. (2015): "Fracking Growth," Working paper, London School of Economics.

Finn, T. and R. 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*, Denver, CO: United States Geological Survey, chap. 6.

Fletcher, S. (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, 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, 492–499.

Gold, R. (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," 116.

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

Harper, J. and J. Kostelnik (NA): "The Marcellus Shale Play in Pennsylvania: Part 3: The Modern Marcellus Shale Play," in *Pennsylvania Geological Survey*, Department of Conservation and Natural Resources.

Harper, J. A. (2011): *Activity and Potential of the Utica Shale in Pennsylvania.*

HORNBECK, R. AND E. MORETTI (2015): "Who Benefits From Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality," Tech. rep.

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

IMBENS, G. AND D. RUBIN (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, New York, NY: Cambridge University Press, first ed.

INTERNAL REVENUE SERVICE (2015): *Statistics on Income Tax Stats - US Population Migration Data*, Washington, DC: IRS.

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

KLANN, S. (2012): "Niobrara Challenge," *Oil and Gas Investor*, 32, NA.

KRAUSS, C. (2009): *Natural Gas Prices Plummet to a Seven-Year Low.*

LANGFORD, S. (2008): "The Woodford Shale: Newfield Investor Conference 2008," Tech. rep., Newfield Investor Conference 2008.

LEWIS ENERGY GROUP (2015): *South Texas is Lewis Energy Territory.*

MARCIL, J.-S., P. DORRINS, J. LAVOIE, N. MECHTI, AND J.-Y. 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, D. F. (2007): "History of the Newark East field and the Barnett Shale as a gas reservoir," *AAPG Bulletin*, 91, 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, 32–43.

MONTANA BOARD OF OIL AND GAS CONSERVATION (2015): *Wells Data.*

MONTGOMERY, C. AND M. SMITH (2010): "Hydraulic Fracturing: A History of An Enduring Technology," *Journal of Petroleum Technology*, 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.

NEW YORK STATE DEPARTMENT OF ENVIRONMENTAL CONSERVATION (2015): *New York's Oil and Gas Database.*

NEWFIELD EXPLORATION (2006): *Newfield Provides Update on Woodford Shale Drilling Program*, Houston.

NEWFIELD EXPLORATION COMPANY (2006): "Newfield Exploration Company: 2005 Annual Report," Tech. rep.

NORDENG, S. (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.

OKLAHOMA GEOLOGICAL SURVEY (2015): *Oklahoma Gas Shales Well Database.*

OLDHAM, J. (2015): *The Real Estate Crisis in North Dakota's Man Camps.*

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

OSBORN, S. G., A. VENGOSH, N. R. WARNER, AND R. 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, 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, S. (2014): *Pennsylvania Frack Ponds Now Number More than 500.*

PHISH, T., T. MCDERMOTT, AND S. WATEROUS (2010): "Niobrara Oil: Regional Spotlight," *Oil and Gas Investor*, 30, 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, C. (2013): "The Cline Shale: Another Texas Giant?" *Oil and Gas Investor*, 33, NA.

RUBINSTEIN, J. L. AND A. B. MAHANI (2015): "Myths and Facts on Wastewater Injection, Hydraulic Fracturing, Enhanced Oil Recovery, and Induced Seismicity," *Seismological Research Letters*, 86, 1060–1067.

RUGGLES, S., K. GENADEK, R. GOEKEN, J. GROVER, AND M. SOBEK (2015): *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database].*, Minneapolis: University of Minnesota.

RYSTAD ENERGY (2014): *North American Shale Plays Maps: Q3, 2014*, Used with permission: http://www.rystadenergy.com/.

SHEEHAN, C. (2013): "The Niobrara Extends its Reach," *Oil and Gas Investor*, 33, NA.

SHELBY, P. (2008): "The Fayetteville Shale Play of North Central Arkansas A Project Update," in *Search and Discovery*, San Antonio: American Association of Petroleum Geosciences: Annual Convention.

SIDEL, R. AND C. CUMMINS (2001): *Devon Energy to Acquire Mitchell in $3.1 Billion Cash-and-Stock Deal.*

SILVER, J. (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, J. (2009): "Bakken Breakout," *Oil and Gas Investor*, 29, NA.

SUAREZ SERRATO, J. C. AND O. ZIDAR (2016): "Who Benefits From State Corporate Tax Cuts ? a Local Labor Markets Approach with Heterogeneous Firms," *American Economic Review*, Forthcoming.

SUTTON, L. (2015): *Permian Basin Production - Midland vs. Delaware Basins.*

TAKERSLEY, J. (2015): *After plunge in oil prices, hope fades for group of long-beleaguered workers.*

TAYLOR, L. (2015): *Fayetteville Shale.*

TOAL, B. A. (2007): "Fayetteville and Woodford Shales," *Oil and Gas Investor*, 27, 38–48.

TOON, S. (2011a): "Eagle Ford Rising," *Oil and Gas Investor*, 31.

——— (2011b): "Pioneer, EOG Lead Foray Into Emerging Horizontal Wolfcamp," *Oil and Gas Investor*, 31, NA.

Trembath, A., J. Jenkins, T. Nordhaus, and M. Schellenberger (2012): “Where the Shale Gas Revolution Came From: Government’s Role in the Development of Hydraulic Fracturing in Shale,” Tech. Rep. 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).*

——— (2013): “American Community Survey Puerto Rico Community Survey 2013 Subject Definitions,” Tech. rep., US Census Bureau.

——— (2014): *Building Permits Survey.*

Vaughn, V. (2012): *Shale play built on expertise, risk, secrecy - and a simple handshake.*

Wang, Z. and A. Krupnick (2013): “A Retrospective Review of Shale as Development in the United States: What Led to the Boom?” Tech. rep.

Warlick, D. (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, 1336–1340.

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

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

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

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

——— (2009): “Fly Like an Eagle Ford,” *Oil and Gas Investor*, 29, 51–54.

——— (2010a): “Bone Spring Springs Forth,” *Oil and Gas Investor*, 30, NA.

——— (2010b): “Oklahoma’s Woodford,” *Oil and Gas Investor*, 30, NA.

——— (2010c): “The Niobrara,” *Oil and Gas Investor*, 30, NA.

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

ZAGORSKI, W. A., G. WRIGHTSTONE, AND D. 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*, ed. by J. A. Breyer, no. 97 in AAPG Memoir, 172 – 200.

ZUCKERMAN, G. (2013): *The Frackers: The Outrageous Inside Story of the New Billionare Wildcatters*, New York, NY: Penguin.