American Economic Journal: Applied Economics 2019, 11(4): 105–155
https://doi.org/10.1257/app.20170487

The Local Economic and Welfare Consequences of Hydraulic Fracturing[†]

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

Exploiting geological variation and timing in the initiation of hydraulic fracturing, we find that fracking leads to sharp increases in oil and gas recovery and improvements in a wide set of economic indicators. There is also evidence of deterioration in local amenities, which may include increases in crime, noise, and traffic and declines in health. Using a Rosen-Roback-style spatial equilibrium model to infer the net welfare impacts, we estimate that willingness-to-pay (WTP) for allowing fracking equals about $2,500 per household annually (4.9 percent of household income), although WTP is heterogeneous, ranging from more than $10,000 to roughly 0 across 10 shale regions. (*JEL* D12, K42, L71, Q35, Q51, Q53, R41)

The discovery of hydraulic fracturing is considered the most important change in the energy sector since the commercialization of nuclear energy in the 1950s. Fracking 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 has been sharp increases in US production of oil and natural gas to levels unimaginable even a decade 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, there are extensive shale deposits of both

* Bartik: Department of Economics, University of Illinois at Urbana-Champaign, 214 David Kinley Hall, 1407 West Gregory Drive, Urbana, IL 61801 (email: abartik@illinois.edu); Currie: Department of Economics and Public Affairs, Princeton University, 185A Julis Romo Rabinowitz Building, Princeton, NJ 08544, Center for Health and Well-Being, and NBER (email: jcurrie@princeton.edu); Greenstone: Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637, the Energy Policy Institute at Chicago, and NBER (email: mgreenst@uchicago.edu); Knittel: Department of Applied Economics, Sloan School of Management, Massachusetts Institute of Technology, 100 Main Street, Cambridge, MA 02142, the Center for Energy and Environmental Policy Research, and NBER (email: knittel@mit.edu). Benjamin Olken was coeditor for this article. This paper has benefited from conversations with John Deutch, Lucas Davis, Manasi Deshpande, Joseph Doyle, Esther Duflo, Edson Severnini, and seminar participants at the ASSA 2014 Meetings, the Center for Energy and Environmental Policy Research Spring Workshop, the Energy Institute at Haas Summer Energy Camp, MIT, Tufts University, University of Chicago, and the NBER Summer Institute. We thank Mateo Arbelaez, Brian Goggin, Faraz Hayat, Colin McDonnell, Ishan Nath, Lucy Page, Jonathan Petkun, Harshil Sahai, and Daniel Stuart for excellent research assistance. Greenstone gratefully acknowledges financial support from MIT's Provost Office. Bartik gratefully acknowledges funding from the MIT Center for Energy and Environmental Policy Research (CEEPR). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.
[†] Go to https://doi.org/10.1257/app.20170487 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

natural gas and oil around the world that greatly increase the potential supply of inexpensive fossil fuels. These deposits offer immediate economic benefits in terms of lower energy prices, but also pose a challenge for reducing the rate of climate change.

Ultimately, access to these energy resources rests on the willingness of local communities to allow fracking within their jurisdictions. Drilling brings royalty payments and economic activity, but there are substantial concerns about potential impacts on the quality of life, including pollution, traffic congestion, and crime.[1] There has been substantial heterogeneity in communities' reactions: Pennsylvania, Texas, and North Dakota have embraced fracking, while other localities, such as New York, Vermont, and some countries, including Germany and France, have banned it. In making these decisions about allowing fracking, policymakers and their communities have not had systematic evidence of its benefits or costs.

This paper's aim is to develop a measure of the net welfare consequences of fracking on local communities that accounts for both its benefits and costs. This task requires developing a counterfactual for what would have happened in fracking communities in the absence of fracking and a theoretically grounded measure of welfare. With respect to the empirical challenge, the task is complicated by the fact that fracking communities are not randomly assigned. Empirically, these communities differ from other parts of the United States in both levels and trends of economic variables. Consequently, we develop an identification strategy that is based on geological variation within shale formations and local variation in the initiation of fracking. Specifically, we exploit Rystad Energy's index of fracking suitability. Rystad is an international oil and gas consulting company, and their index is based on several factors, including thickness, depth, and thermal maturity of the shale deposit. Thus, our identification strategy focuses on counties in the same shale play or formation. The second source of variation is the difference in fracking's initiation timing 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, comparing the change in local economic outcomes in areas with high geological potential for fracking to that in areas with lower geological potential for fracking within the same shale formation.

Our estimates of the welfare consequences are based on a Rosen-Roback-style model of locational equilibrium. Building on the work in Moretti (2011) and Hornbeck and Moretti (2019), which adds moving costs and elastic housing supply to Roback's (1982) canonical model, we allow fracking to shift both local productivity and local amenities. We derive an expression for willingness-to-pay (WTP) for fracking that is equal to the product of a locality's total population and the change in real income (i.e., accounting for changes in income and local prices, measured by housing prices) and the WTP for amenity changes. We also derive an expression for WTP for amenity changes. Importantly, these expressions are functions of the reduced-form estimates delivered by the paper's identification strategy.

There are three primary findings. First, counties with high fracking potential experience a natural resources boom. They produce roughly an additional

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

\$400 million of oil and natural gas annually three years after the discovery of successful fracking techniques, relative to other counties in the same shale play. Furthermore, they experience marked increases in economic activity, with gains in total income (3.3–6.1 percent), employment (3.7–5.5 percent), and salaries (5.4–11.0 percent). There are also increases in housing prices (5.7 percent) and rental rates (2.9 percent). Additionally, local governments see substantial gains in revenues (15.5 percent) that are larger than the average increases in expenditures (12.9 percent).

Second, there is evidence of deterioration in the noneconomic quality of life or total amenities. One advantage of the model is that it allows for an estimate of the WTP for the total change in amenities, even when the full vector of amenities is unobservable. Using the model's results, we estimate that annual WTP for fracking-induced changes in local amenities is roughly equal to −\$1,400 per household annually, or −2.7 percent of mean annual household income. Direct empirical estimates of local amenities are more difficult given limited data on local amenities. However, we find some, albeit noisy, evidence of higher violent crime rates, despite a 20 percent increase in public safety expenditures.

Third, we use the model to develop a measure of the net change in welfare among households that lived in these communities prior to fracking's initiation that accounts for the economic benefits *and* costs of declining amenities. We estimate that across all US shale plays, the WTP for allowing fracking equals about \$2,500 per household annually, or about 4.9 percent of mean household income. Importantly, we uncover substantial heterogeneity in WTP across shale plays. The largest estimates come from the Bakken's (primarily in North Dakota and Montana) annual WTP of \$11,700 and the Woodford-Arkoma's (Oklahoma) WTP of \$3,700, although there are also large net gains in the Fayetteville (in Arkansas and Oklahoma), the Marcellus (largely in Pennsylvania, West Virginia, and Ohio), the Woodford-Ardmore, and the Permian Basin (West Texas) plays. The estimates of WTP are roughly zero and statistically insignificant for several of the plays.

This paper makes several contributions. First, the focus on net welfare consequences provides a broad picture of fracking's overall impacts.[2] Previous work has largely focused on estimating either the local labor market benefits of fracking or the environmental and social costs, and we bridge the gap between these two literatures by using an equilibrium model of local decisions to estimate the value of both the benefits and costs to local communities. Of course, these estimates are only as good as the information on impacts of fracking that households have at their disposal, and as new information emerges about potential health consequences and other impacts, this information may change.[3]

[2] 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 fracking on the housing market. An important paper by Muehlenbachs, Spiller, and Timmins (2015) finds that in a sample of roughly 1,000 Marcellus-region houses, proximity to a fracking 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 fracking affects global greenhouse gas emissions and geopolitics.

[3] The EPA released a preliminary report on a wide-ranging study of the health and environmental risks of fracking (EPA 2015). Regulations also continue to evolve.

Second, the examination of nine different shale plays provides a more comprehensive measure of the impacts of fracking across the United States,[4] building on important previous work that has focused largely on single shale plays, especially the Marcellus in Pennsylvania (Gopalakrishnan and Klaiber 2013; Muehlenbachs, Spiller, and Timmins 2015). Third, the paper demonstrates that areas of the country with abundant opportunities for fracking differ from the rest of the country in important ways. As a solution to this identification problem, this paper offers a new and credible identification strategy based on the geological characteristics of shale deposits and the timing of when new technologies became available. 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, provide a fuller picture of fracking's impacts than has been available previously. In this respect, the paper expands our understanding of resource booms (see, e.g., Wynveen 2011).[5] In related work, Jacobsen (2019) also finds that fracking has benefited local communities economically as measured by wages and housing rental rates.

Our estimates are likely to be relevant for communities making decisions about whether to allow fracking. There are vast shale deposits around the globe that have not yet been accessed due to a mix of legal, institutional, technical, and economic reasons. As some of these barriers are removed, more jurisdictions will be confronted with decisions about whether to allow fracking.[6]

The paper proceeds as follows. Section I outlines our conceptual framework. Section II discusses hydraulic fracturing and how it differs from conventional oil and natural gas recovery. Section III discusses the data used in the analysis, while Section IV describes our identification strategy. Section V provides preliminary evidence, our econometric estimates, and evidence about the robustness of those results. Section VI presents evidence of local welfare implications of our estimates. Finally, Section VII concludes.

I. Conceptual Framework

This section extends a stylized model that builds upon the canonical Roback (1982) model, following Chay and Greenstone (2005); Greenstone, Hornbeck, and Moretti (2010); and Kline and Moretti (2014). The model is a slightly modified version of Moretti (2011) and Hornbeck and Moretti (2019), which incorporate the possibility of moving costs and elastic housing supply into a Roback-style model (Roback 1982). The only difference between the model we present here and Hornbeck and Moretti (2019) is that the latter is focused on the effects of a pure productivity shock, whereas we allow the introduction of fracking to shift both local

[4] We restrict the sample to nine plays to ensure enough post-fracking data to identify the effects.

[5] Our work does not shed light on the potential for the "Dutch disease" (see, e.g., Allcott and Keniston 2018 and Fetzer 2014 for work on this topic) or our understanding of how these effects propagate (see, e.g., Feyrer, Mansur, and Sacerdote 2017).

[6] See Covert, Greenstone, and Knittel (2016) for a discussion of these issues and http://www.eia.gov/todayinenergy/detail.cfm?id=14431 for a map of world resources.

productivity and amenities. The ultimate aim is to develop expressions for WTP even though not all amenity changes are observed.

We assume that household i in location j at time t obtains utility from the consumption of a numéraire good sold on a global market, C_{ijt} (with price normalized to 1); housing, H_{ijt}; location amenities, A_{jt}; and idiosyncratic place-based preferences and moving costs, μ_{ijt}. Assuming Cobb-Douglas utility yields

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

where β is the share of household income spent on housing and the exponent s measures the size of moving costs or variance of idiosyncratic preferences; in the canonical Roback model, these idiosyncratic preferences and moving costs do not exist, which is equivalent to assuming $s = 0$. Additionally, 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.[7] Solving for the consumer's problem and taking logs yields the indirect utility function:

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

where $\epsilon_{ijt} = \ln \mu_{ijt}$. A key feature of the model is that housing supply is elastic, where inverse housing supply (i.e., X_{jt}) is given by

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

For intuition on how prices allocate individuals across locations, consider the case where there are only two locations: a and b. Assuming that μ_{ijt} are independently drawn every period so that future shocks do not affect current decisions, the household's problem simplifies to choosing the location that maximizes current-period utility.[8] Consequently, a household chooses to live in location a in period t if and only if $u_{iat} - u_{ibt} > 0$. Defining $\tilde{x} = x_a - x_b$ and using our expression for indirect utility in equation (2), we can write the household's decision rule as

$$\widetilde{\ln w}_t - \beta \widetilde{\ln r}_t + \alpha \widetilde{\ln A}_t + s \times \tilde{\epsilon}_{it} > 0.$$

[7] We abstract away from differences in housing rents and housing prices. In the simplest model with competitive housing markets, the housing price will equal $(1/(1-\rho))\bar{r}$, 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.

[8] This assumption rules out most dynamic elements of a household's decision.

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

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

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

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

Taking differences of equation (4), assuming that location a is "small" relative to location b, and rearranging yields an expression for household WTP for the amenity changes caused by fracking:[9]

$$\Delta WTP\ for\ amenities = \alpha \Delta \ln A_{at} = s \Delta \ln N_{at} - (\Delta \ln w_{at} - \beta \Delta \ln r_{at}). \tag{5}$$

Equation (5) is of tremendous practical value because it provides an expression for the full set of amenity changes,[10] even though a dataset with the complete vector of amenities and information on WTP for these amenities is unlikely to ever be available. Specifically, this expression says that the WTP for the change in amenities, expressed as a percentage of income, is equal to the difference between the percentage change in population, adjusted for the magnitude of moving costs, and the percentage change in real wages. In the subsequent empirical analysis, we will estimate the effect of fracking on housing prices and rents ($\Delta \ln r_t$) that are assumed to be an index for locally produced goods,[11] household wage and salary income ($\Delta \ln w_t$), and population ($\Delta \ln N_t$), respectively.[12] We go to the previous literature to obtain estimates of the 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 Suárez Serrato and Zidar (2016). The bottom line is that by combining the paper's empirical estimates with estimates of moving costs and the share of income devoted to housing, it is possible to derive an implementable expression for the WTP for the change in amenities in location a. The intuition behind this approach comes from the fact that, in

[9] Assuming that location a is small relative to location b allows us to assume that fracking has no general equilibrium impacts on prices in location b, and that we can rewrite the effect of fracking on differences between locations a and b, $\tilde{x}$, instead as the change in location a, Δx.

[10] A_{at} is the full vector of a location's amenities, and α measures the WTP for log changes in those amenities.

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

[12] In online Appendix Section I, we derive an explicit expression for the change in house prices in terms of parameters in equation (5).

spatial equilibrium, the marginal resident must be indifferent to relocating, which means that local housing prices will respond to changes in the utility provided by a location. The strength of this response will depend on both the elasticity of local housing supply and moving costs.

Using this expression, it is possible to develop an expression for the change in welfare for all the people who either reside or own property in location a before the change in amenities and local productivity occurs.[13] This is the population that has the greatest influence on whether fracking should be allowed in a community. In particular, let $\overline{W}_a$ 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 $\overline{W}_a \times \left(\Delta \ln w_{at} + \alpha \Delta \ln A_{at} - \beta \Delta \ln r_{at}\right)$, the welfare change for a landowner (who may or may not reside in location a) who owns one housing unit is $\bar{R}_a \times \Delta \ln r_{at} + \bar{Y}_a^{owner} \times \Delta \ln y_{at}^{owner}$, and WTP for all individuals who either reside or own property in location a before fracking is the sum of these two terms:[14]

$$(6) \quad \textit{WTP for allowing fracking} = \Delta V_{at}$$
$$\approx N_{at} \times \left(\overline{W}_a \times \Delta \ln w_{at} + \bar{Y}_a \times \Delta \ln y_{at} + \overline{W}_a \times \alpha \Delta \ln A_{at}\right).$$

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 the change in 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.[15] An appeal of this expression for WTP is that it is more realistic than the workhorse expression from the canonical Roback (1982) model (which is simply equal to the change in property values) because it does not make the unrealistic assumption that housing is supplied perfectly inelastically, and it reflects the fact that households face moving costs.

Nevertheless, this model is still stylized, and there are three caveats worth highlighting. First, the model assumes that workers are homogeneous, and relaxing this assumption would lead to additional welfare consequences. Renters with skills that are not well suited for fracking-related employment (e.g., the elderly) are

[13] 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. Ignoring the change in welfare for migrants can be justified by appealing to envelope-theorem-type arguments, as in Busso, Gregory, and Kline (2013). It also assumes that the average change in household income is attained by original residents and is not due to high earnings by immigrants. This assumption is supported by evidence from Bartik and Rinz (2018), which exploits longitudinal census data and a similar variation to this paper, and finds that the effects of fracking on labor market earnings for original residents of fracking-exposed locations are similar to the effects on overall average labor market earnings. 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 fracking in a community from the perspective of a local policymaker.

[14] Here, $\bar{Y}^{owner}$ is the average interest and dividend income for homeowners.

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

especially vulnerable; members of this group could experience declines in utility due to continued residence in a jurisdiction that allows fracking, and they could face moving costs that, in principle, could lock them into their current locations. 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 accrue to someone, our estimates of the effects of fracking on the change in housing prices or rent and dividend payments will not capture these benefits. Second, the model assumes that households have knowledge of and rational expectations about fracking'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 fracking will be more complicated. Of course, as new information about fracking's impacts (e.g., health effects, as in Currie, Greenstone, and Meckel 2017) emerges, households will update their WTP 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 fracking, including reductions in petroleum, natural gas, or electricity prices; effects on global warming or adoption of renewable technologies; and changes in geopolitics resulting from America's growing role as a fossil fuel producer. The model also assumes that fracking-affected areas are small enough relative to the US labor market that we can abstract from general equilibrium effects on overall wages in the United States.

II. A Primer on Hydraulic Fracturing and a New Research Design

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 our research design. The online Appendix provides more details.

A. *A Primer on Hydraulic Fracturing*

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." The oil and gas migrate to "pools" in permeable reservoir rocks such as limestone 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. When the drill reaches the layer of the pool (typically 1,000–5,000 feet below the surface for an onshore well), the drill bit is removed, and casing pipe is placed into the hole. Finally, 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. Alternatively, they may be pumped. For unconventional wells, drilling often continues to lower depths—sometimes exceeding 10,000 feet. Once the drill bit nears the shale formation, the bit begins to turn sideways, and drilling often continues in a horizontal fashion for more than 10,000 feet. This portion of the well is then cased and perforated. However, 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 wellhead. To break the pockets, a mixture of water, sand, and chemicals is pumped into the well under high pressure. 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. Due to both the costs of fracturing itself and associated horizontal drilling, fracked wells are usually several times more expensive than conventionally drilled wells. For example, the move away from conventional wells to fracked horizontal wells in the Woodford-Shale has increased average well costs from around $2 million to $5–6 million (Fitzgerald 2013). More broadly, the advent of fracking was associated with a roughly threefold increase in the average costs per well drilled in the United States (Fitzgerald 2013).[16] However, fracking has been dubbed farming for the relative certainty of producing hydrocarbons. Although national data on the number of wells that are fracked are unavailable, we can gain a sense for the emergence of fracking from the share of new wells that are drilled horizontally over shale formations; this share increased from 0.7 percent in 2000 to 25 percent in 2011 (data purchased from Drillinginfo, Inc. 2012). In part because of this rapid increase in fracking, the fraction of successful exploratory wells in the United States has risen from 41 percent in 2000 to 62 percent in 2010 (analysis of Energy Information Agency (EIA) and IHS data in Covert, Greenstone, and Knittel 2016).[17] Figure 1 shows that hydrocarbon production from horizonal wells over shale formations—a proxy for fracked wells—has increased substantially since 2000, while production from traditional wells has been stagnant.

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 a rough bowl shape, with the center then filled in by layers of sediment. If one of the layers is a shale layer, the basin can be referred to as a "shale" basin. A shale play is part of a shale basin where oil- and gas-producing firms have targeted a specific formation that exhibits similar geological and drilling characteristics. 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[18] 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.

[16] Firms have learned how to reduce horizontal drilling and hydraulic fracturing costs as they have gained more experience with these technologies (Covert 2015). Consequently, the current costs of fracking are likely lower than these figures and they may decline further going forward.

[17] Advances in three-dimensional imaging have also reduced dry holes for conventional wells.

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

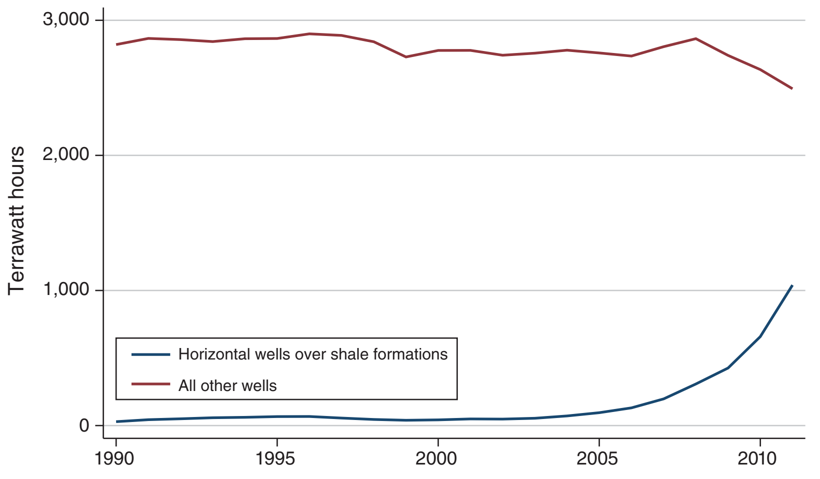

Figure 1. Hydrocarbon Production from Horizontal Wells over Shale Play

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 result of the technological innovations in using fracking 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.

Source: The data come from Drillinginfo, Inc. (2012).

Local Impacts of Hydraulic Fracturing Activity.—This paper builds on previous work that has striven to measure the economic benefits of fracking to local communities in terms of hydrocarbon production, employment, income, net migration, etc. (see, e.g., Feyrer, Mansur, and Sacerdote 2017; Fetzer 2014; Gopalakrishnan and Klaiber 2013; Muehlenbachs, Spiller, and Timmins 2015; Jacobsen 2019; Newell and Raimi 2018a, b; Weber 2012; Weinstein 2014). However, these benefits may come with substantial costs in terms of water and air pollution, traffic, crime, and damage to otherwise largely unperturbed physical environments (see, e.g., EPA 2015, Phillips 2014, Ground Water Protection Council and ALL Consulting 2009, National Energy Technology Laboratory (NETL) 2013, Rubinstein and Mahani 2015). We attempt to measure as many of these local impacts as possible, but ultimately, they cannot all be measured, and even if they could, their net impact on social welfare is unknowable. Our conceptual framework offers a way out of this conundrum by allowing us to develop estimates of the WTP for the total change in amenities and the net welfare impacts of allowing fracking in the community. This approach allows us to bridge the gap between the literature studying the benefits of fracking and the literature studying the costs of fracking, providing the first picture of the net impacts on local communities.

B. *A New Research Design*

This paper's empirical analysis aims to determine the consequences of fracking for a local community. The empirical challenge is to identify a valid counterfactual.

The difficulty is that places with fracking may differ from those without for reasons that also affect key outcomes. For example, 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 fracking, other things being equal.

The growing fracking literature offers a variety of identification strategies. Perhaps the most widely used is to compare areas over shale formations to areas without shale formations beneath them (see, e.g., Cascio and Narayan 2019; Fetzer 2014; Maniloff and Mastromonaco 2014; Weber 2012; Weinstein 2014). As we demonstrate below, however, these places differ on many dimensions in both levels and trends. Additionally, within shale plays there is substantial variation in how amenable different counties are to fracking, so these play-based strategies may be less powered relative to strategies that also exploit within-play variation. Others have taken advantage of a border discontinuity design, based on comparing areas in Pennsylvania, where fracking has been embraced, with bordering areas in New York, where it has been banned (Boslett, Guilfoos, and Lang 2016). This design is appealing for reasons of internal validity. However, as we show below, there is considerable heterogeneity in the effects of fracking across shale plays, so it is useful to develop an identification strategy that can be applied to multiple plays.

Our 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.

Cross-Sectional Variation in Prospectivity within Shale Plays.—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 fracking 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, Wrightstone, and Bowman 2012; Budzik 2013; Covert 2015; McCarthy et al. 2011). Rystad Energy is an oil and gas consulting firm that has created a "prospectivity" index of the potential productivity of different portions of shale plays based on a nonlinear function of the different geological inputs. We purchased Rystad's NASMaps product, which includes geographic information system (GIS) shapefiles of Rystad's prospectivity estimates for each North American shale play (Rystad Energy 2014). Figure 2 maps the Rystad prospectivity estimates for the major US shale plays. The geological variables included and the functional forms used to transform them into prospectivity scores differ for each shale, so scores cannot be compared across shale plays.

We aggregated the Rystad prospectivity measure to the county level by computing the mean and maximum Rystad score within each county. We then divided the counties in each shale play into Rystad score quartiles. Our preferred measure of potential fracking 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

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	Fort 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 fracking potential of the shale play became public.

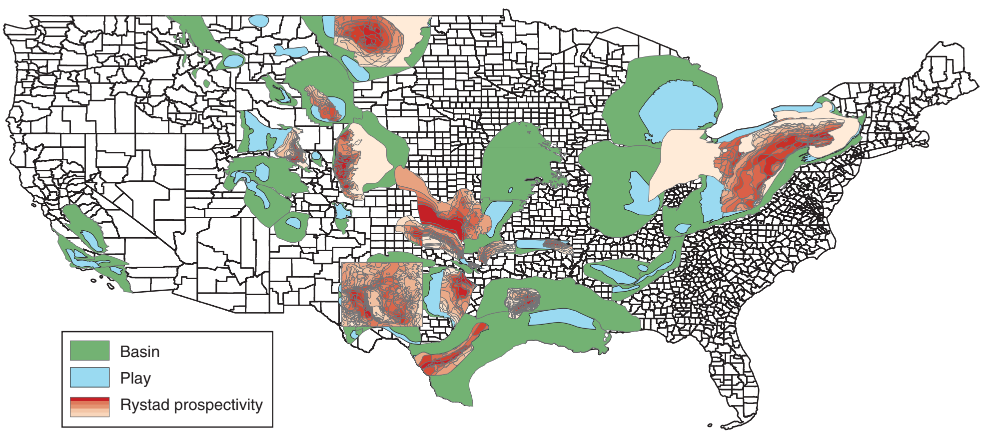

Figure 2. Shale Basins, Plays, and Prospectivity Scores

Notes: This figure overlays shale basins, shale plays, and Rystad prospectivity scores on 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.

Source: Shapefiles for US shale basins and plays come from EIA (2011), while prospectivity scores were purchased from Rystad Energy (2014).

production than the average quality. We also explore the sensitivity of the results to alternative measures of fracking exposure. Figure 3 shows a map of the county assignments. The online Appendix illustrates in greater detail how the Rystad prospectivity measure was used to divide counties into the top quartile and the bottom three quartiles.

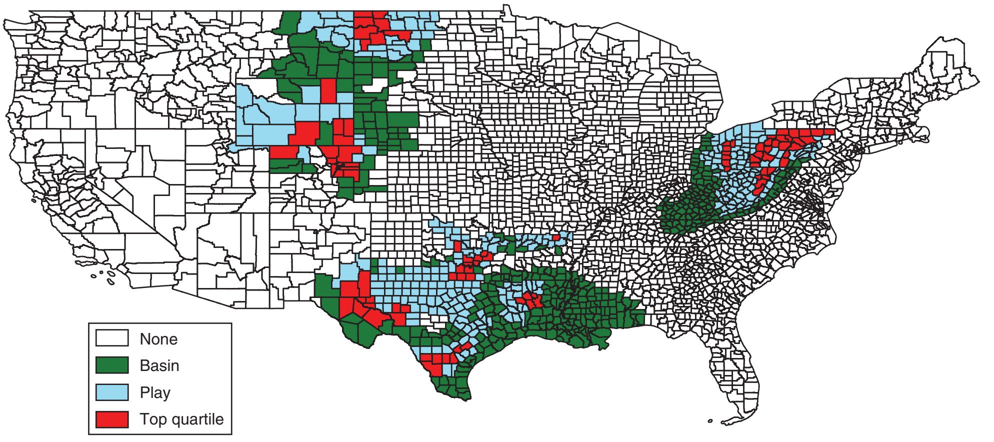

FIGURE 3. COUNTY PROSPECTIVITY SCORE CLASSIFICATIONS

Notes: This figure shows prospectivity score classifications for counties in the contiguous United States. 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.

Source: Author's analysis based on shapefiles for US shale basins and plays come from EIA (2011) and prospectivity scores from Rystad Energy (2014).

Temporal and Cross-Sectional Variation in the Discovery of Successful Fracking 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 fracking across shale plays. This time variation comes 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 fracking techniques in each play, as well as from local and national economic factors influencing oil and gas development. We determined the first date that the fracking potential of each of the 14 shale plays in the United States became public knowledge. When possible, these dates correspond to investor calls and production announcements when firms first began drilling operations involving fracking in an area or released information on their wells' productivity. The online 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 fracking 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. Fracking 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.

C. *Potential Spillovers and Alternative Identification Strategies*

Fracking opportunities in top-quartile counties might have spillovers on other counties, especially other counties in the same shale play, for at least two reasons.

First, counties in physical proximity to top-quartile counties may experience an increase in fossil fuel recovery within their own boundaries. Second, the increased economic activity in top-quartile counties can directly benefit nearby counties. In the presence of positive spillovers, the paper's identification strategy would underestimate the impacts on top-quartile counties because it relies on comparisons of top-quartile counties to neighboring counties in the same shale play. Furthermore, the identification strategy is not designed to produce estimates of the full impacts of fracking because it doesn't measure the impacts on non-top-quartile counties.

An alternative identification strategy would involve using propensity-score-matching to match all counties within shale plays to counties outside shale plays (Imbens and Rubin 2015). However, our exploration of this strategy showed that it is extremely difficult to balance covariates between the shale play counties and their matched comparisons. In light of the potential for confounding, we restrict reporting on the results from this approach to an abbreviated discussion in the online Appendix. For these results, see online Appendix Tables 14 to 17, which parallel Tables 3 to 8 in the main text.

III. Data Sources and Summary Statistics

Clearly, it would be impossible to estimate the effects of fracking on every potential outcome; however, we collected data on a large set of outcomes and will use these results to estimate the net welfare effects of fracking. This section briefly describes the data sources, with more details provided in the online Data Appendix, and provides some evidence on the validity of the research design.

A. *Data Sources*

Fracking Data.—Shapefiles of the locations of shale plays and basins, as well as historic oil and gas prices, come from the EIA.[19] Oil and gas production data for 1992 through 2011 were purchased from Drillinginfo, Inc. (2012) (formerly HPDI). The research design depends on the prospectivity estimates from Rystad Energy's NASMaps product purchased from Rystad Energy (2014).

Economic Outcomes.—Data on county-level economic outcomes come from several sources. The Bureau of Economic Analysis's Regional Economic and Information Systems (REIS) data on total employment and total annual earnings by type (BEA) 2014) are complemented by the Quarterly Census of Employment and Wages (QCEW) data on wages by industry (Bureau of Labor Statistics 2014).

Housing price data for 2009 through 2013 come from the American Community Survey (ACS), while housing price data for previous decades (2000s and 1990s), as well as data on the total number of housing units, come from the decennial

[19] For oil prices we use the Cushing, Oklahoma spot price for West Texas Intermediate (EIA 2011), and for natural gas we use the city-gate price. Shapefiles for the boundaries of shale plays and basins come from the EIA's Maps: Exploration, Resources, Reserves, and Production site (EIA 2011).

census.[20] 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.[21] The 2009–2013 ACS data need to be pooled to precisely estimate average county outcomes, so for a given county, these data are treated as a single year's observation.[22] 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 Bureau of Labor Statistics (2015). Migration data come from the Internal Revenue Service's county-county migration dataset, released as part of the Statistics on Income (IRS 2015).

Crime Data.—Crime data come from the Federal Bureau of Investigation (FBI) Uniform Crime Reporting (UCR) program (FBI 2015). Individual police agencies (e.g., Cambridge Police Department, MIT Police) report "index crimes" to the FBI, including murder, rape, aggravated assault, robbery, burglary, larceny, and motor vehicle theft. Reporting is not mandatory[23] and, consequently, not all agencies report all index crimes in all years. In order to define a consistent series, we use agencies that report[24] index crimes in most years[25] from 1992 through 2013. To ensure that the consistently reporting agencies are representative of the county as a whole, we include only counties in which sample agencies account for at least 20 percent of the total crimes in the county between 2011 and 2013.[26] 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.

Public Finance.—Data on local government spending and revenues come from the Census of Governments conducted every five years (years ending in 2 and 7) by the 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 above. We supplement these data using school district-level enrollment data from the Common

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

[21] All census and ACS data were retrieved from the National Historical Geographical Information System (Manson et al. 2018).

[22] The Census Bureau suppresses data for many counties in the one-year and three-year ACS releases. Data from very few counties are suppressed in the five-year ACS estimates.

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

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

[25] 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.

[26] 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 for 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.

Core (National Center for Education Reseaarch (NCES) 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,[27] 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.

B. *Summary Statistics*

Column 1 of Table 2 reports the county-level means of key variables. Panel A reports the values in 2000, before the widespread adoption of fracking, while panel B reports on the change between 1990 and 2000. 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 discuss the non-crime variables first and the crime variables next.

Column 2 compares counties over shale basins with other US counties and shows that there are important differences between these two sets of counties. Counties within a shale basin had worse economic outcomes in 2000; for example, per capita income in 2000 is almost 30 percent (0.285 natural log points) lower in shale basin counties. Indeed, nine of the ten reported variables are statistically (and economically) different between the two sets of counties. Panel B reveals that shale basin counties were growing more slowly than the rest of the country from 1990 to 2000; eight of the ten variables are statistically significantly different across the two sets of counties and the magnitudes are economically large in most cases. Overall, the results in column 2 suggest the need for an alternative to a difference-in-difference specification that is based on comparing shale basin counties with the rest of the United States.

Column 3 explores the validity of comparing counties in shale plays with counties in the same shale basin but not necessarily in the same shale play. The entries report the results from regressions of the variable in the row against an indicator for whether the county is in a shale play, an indicator for whether the county is in a shale play interacted for an indicator for whether the shale play is not 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. The differences in income levels and income changes are even larger than in column 2, and across the ten variables there are again statistically

[27] 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 online Appendix Table 12, we also report the robustness of our results to estimating long-difference models of the change between 1997 and 2012.

Table 2—Comparison of Pre-trends and Levels across Treatment and Control Counties

	Mean value in US (1)	Basin versus rest of US (2)	Play versus basin (3)	Rystad top quartile versus play (4)	Rystad top quartile versus p-score-matched sample (5)	Quartiles 1–3 versus p-score-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	−0.116	−0.157
		(0.037)	(0.031)	(0.050)	(0.064)	(0.046)
log (real median home rental prices)	6.621	−0.179	−0.024	0.055	−0.094	−0.090
		(0.032)	(0.030)	(0.045)	(0.060)	(0.036)
log (total housing units)	9.427	−0.159	0.413	0.082	−0.211	−0.353
		(0.055)	(0.087)	(0.143)	(0.169)	(0.109)
log (total employment)	9.532	−0.245	0.401	0.057	−0.315	−0.420
		(0.060)	(0.104)	(0.161)	(0.178)	(0.117)
log (total income per capita)	13.605	−0.285	0.411	0.036	−0.333	−0.428
		(0.062)	(0.103)	(0.171)	(0.194)	(0.121)
Share of population with bachelor's degree or more	0.241	−0.041	0.003	0.042	−0.003	−0.028
		(0.010)	(0.016)	(0.025)	(0.026)	(0.013)
Share of population ages 18–64	0.619	−0.003	−0.011	−0.003	−0.002	0.003
		(0.003)	(0.004)	(0.007)	(0.009)	(0.005)
log (real total government revenue: 2002–1992)	11.513	−0.274	0.373	0.050	−0.333	−0.421
		(0.059)	(0.101)	(0.159)	(0.177)	(0.114)
log (real total government expenditures: 2002–1992)	11.516	−0.283	0.373	0.063	−0.329	−0.431
		(0.060)	(0.102)	(0.162)	(0.181)	(0.115)
Total value of hydrocarbon production: 2000–1992	56.238	81.575	78.569	108.280	98.973	−2.655
		(19.983)	(17.700)	(58.527)	(68.077)	(43.437)
F-statistic		25.0	7.8	1.6	3.2	2.9
p-value		0.00	0.00	0.10	0.00	0.00
Counties exposed	—	715	316	64	64	252
Observations	2,842	2,842	791	401	1,410	1,625
Panel A2. Crime variables						
Property crimes per 100,000 residents	3,937	167	−1,480	−572	−1,774	−700
		(222)	(365)	(334)	(257)	(273)
Violent crimes per 100,000 residents	537	−56	−156	−64	−213	−94
		(49)	(59)	(81)	(76)	(55)
F-statistic		1.2	5.5	2.6	21.2	2.2
p-value		0.32	0.00	0.05	0.00	0.09
Counties exposed		523	266	56	56	210
Observations	2,020	2,020	573	338	888	1,068

(*continued*)

and economically large differences between these sets of counties. The entries suggest that this comparison is also unlikely to be a good basis for a credible quasi-experiment.

In contrast, the entries in column 4 support the validity of this paper's identification strategy. 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 indicator for whether the shale play that lies under the county is not in the balanced sample of shale plays, and play fixed effects on the subset of counties in plays. The coefficient

TABLE 2—COMPARISON OF PRE-TRENDS AND LEVELS ACROSS TREATMENT AND CONTROL COUNTIES (*continued*)

	Mean value in US (1)	Basin versus rest of US (2)	Play versus basin (3)	Rystad top quartile versus play (4)	Rystad top quartile versus *p*-score-matched sample (5)	Quartiles 1–3 versus *p*-score-matched sample (6)
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	0.051	0.006
		(0.026)	(0.014)	(0.028)	(0.020)	(0.021)
log (real median home rental prices)	0.012	0.055	−0.027	0.003	−0.003	0.003
		(0.016)	(0.006)	(0.008)	(0.018)	(0.014)
log (total housing units)	0.124	−0.036	−0.054	0.009	−0.038	−0.049
		(0.005)	(0.008)	(0.012)	(0.014)	(0.008)
log (total employment)	0.178	−0.042	−0.028	0.028	−0.023	−0.050
		(0.007)	(0.012)	(0.016)	(0.019)	(0.013)
log (total income per capita)	0.259	−0.045	−0.071	0.036	−0.023	−0.059
		(0.007)	(0.013)	(0.018)	(0.021)	(0.014)
Share of population with bachelor's degree or more	0.040	−0.012	0.002	0.013	0.009	−0.006
		(0.003)	(0.003)	(0.005)	(0.005)	(0.004)
Share of population ages 18–64	0.001	0.005	0.000	−0.006	−0.003	0.002
		(0.002)	(0.003)	(0.004)	(0.004)	(0.002)
log (real total government revenue: 2002–1992)	0.286	−0.064	−0.112	0.042	−0.029	−0.071
		(0.011)	(0.019)	(0.027)	(0.028)	(0.021)
log (real total government expenditures: 2002–1992)	0.290	−0.029	−0.124	0.034	−0.029	−0.062
		(0.011)	(0.020)	(0.029)	(0.031)	(0.023)
Total value of hydrocarbon production: 2000–1992	7.934	6.848	4.032	28.929	4.234	−26.252
		(4.150)	(7.247)	(18.096)	(23.299)	(19.522)
F-statistic		14.2	9.0	1.4	2.4	4.1
p-value		0.0	0.0	0.2	0.0	0.0
Counties exposed		715	316	64	64	252
Observations	2,842	2,842	791	401	1,410	1,625
Panel B2. Crime variables (change 1992–2000)						
Property crimes per 100,000 residents	−1,365	323	−365	125	401	28
		(177)	(200)	(220)	(202)	(197)
Violent crimes per 100,000 residents	−246	56	60	41	95	7
		(48)	(40)	(51)	(35)	(43)
F-statistic		1.1	3.9	0.2	2.5	0.0
p-value		0.35	0.01	0.88	0.06	1.00
Counties exposed		523	266	56	56	210
Observations	2,020	2,020	573	338	888	1,068

Notes: This table shows coefficients from regressions of baseline outcomes (panel A) and pre-trends (panel B) on different measures of exposure to fracking activity. Column 1 shows the mean value for the entire United States. 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 *p*-score-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 counties in the bottom three quartiles and the corresponding *p*-score-matched counties for each shale play. All specifications include both the fracking exposure measure and the fracking 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 as 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 US$. Robust standard errors are reported in parentheses in columns 2–4. Columns 5 and 6 cluster standard errors at the county level.

and standard error associated with the top-quartile indicator are reported in the table and are based on the balanced sample of counties. A comparison of pretreatment levels and trends finds little evidence of differences between counties within a shale play that have a top-quartile Rystad prospectivity measure and other counties in the same play. The null of equality between the reported variables cannot be rejected in either levels or trends.[28] Importantly, not only are fewer variables statistically significant in column 4, but the magnitudes of the coefficients decline sharply relative to columns 2 or 3.

The last two columns compare top-quartile counties and non-top-quartile counties to their propensity-score-matched counterparts. Column 5 shows that the propensity-score technique performs better than simple comparisons of counties over shale plays or basins to other counties. However, this strategy does not perform as well as our within-play comparison in column 4; in particular, housing market trends in top-quartile counties appear to differ substantially from those in the propensity-score-matched sample. Column 6 shows that it is even more difficult to find matches for counties in quartiles 1 through 3; all variables but hydrocarbon production are statistically different across the two groups.

Turning to the crime variable levels and pre-trends in panels A2 and B2, column 2 shows that there are small differences in levels of crime rates, but larger differences in trends, in counties within shale basins compared to the rest of the United States. In particular, counties within shale basins had rising property and violent crime between 1992 and 2000 relative to counties outside shale basins. Column 3 shows that comparing counties within shale plays to other counties within the same shale basin increases the magnitude of the difference in crime levels in panel A2 markedly, but actually leaves the magnitude of the differences in crime trends unchanged. Column 4 shows that when comparing Rystad top-quartile counties to other counties within the same shale play, we cannot reject the joint null of similar property and violent crime trends between top-quartile and other shale play counties in either levels or trends. The estimated difference in levels for property crime is statistically significant. However, the standard errors for both trend variables are extremely large, meaning that we cannot rule out quite large pre-trends in crime in top-quartile counties. Consequently, our crime results must be interpreted cautiously.

Finally, we turn to the propensity-score-matching comparisons in columns 5 and 6. Each of the shale play county groups exhibits statistically significantly lower crime rates compared to their propensity-score-matched counterparts. These findings suggest that the propensity-score-matching procedure is not successful in generating an adequate match for counties exposed to fracking, or for the control group of counties that are less likely to be exposed.

Although the column 4 results generally confirm the similarity of the top-quartile Rystad measure counties and other counties in the same shale play, all reported

[28] One of the few variables that remains different in levels across all columns is total hydrocarbon production. We believe this is because some locations with high potential for fracking 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.

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.

IV. Empirical Strategy

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

A. *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):

$$(7) \quad y_{cpt} \;=\; \mu_{pt} + \gamma_c + \delta\big(\mathbf{1}\big[\textit{Post-fracking}\big]_{pt} \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c\big) + \epsilon_{cpt}.$$

The specification includes year-by-play, μ_{pt}, and county fixed effects, γ_c. The two key covariates are (i) $\mathbf{1}\big[\textit{Post-fracking}\big]_{pt}$, which is an indicator that equals 1 in the year that fracking is initiated in shale play p and remains 1 for all subsequent years (this variable equals one for all counties that intersect a shale play after its first-frac date), and (ii) $\mathbf{1}\big[\textit{Rystad top quartile}\big]_c$, which 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 1 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 fracking. It measures the change in the difference in y_{cpt} between counties with high and low Rystad prospectivity values within shale plays, after fracking 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,[29] and that it assumes that fracking affects only 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 fracking on economic and other outcomes may evolve over time in ways that a pure mean shift model fails to capture.

[29] Although we are not able to reject the joint null hypothesis that 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.

Hence, we also estimate a richer specification that directly confronts these two potential shortcomings of equation (7). Let τ_{pt} be a play-specific time trend measuring the number of years since fracking began in play p. We then estimate

$$
\begin{aligned}
(8) \qquad y_{cpt} = {} & \mu_{pt} + \gamma_c \\
& + \beta_1 \big(\tau_{pt} \cdot \mathbf{1}[\mathit{Rystad\ top\ quartile}]_c \big) \\
& + \delta_0 \big(\mathbf{1}[\mathit{Post\text{-}fracking}]_{pt} \cdot \mathbf{1}[\mathit{Rystad\ top\ quartile}]_c \big) \\
& + \delta_1 \big(\tau_{pt} \cdot \mathbf{1}[\mathit{Post\text{-}fracking}]_{pt} \cdot \mathbf{1}[\mathit{Rystad\ top\ quartile}]_c \big) + \epsilon_{cpt}.
\end{aligned}
$$

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 fracking τ years after the start of fracking is $\delta_0 + \delta_1 \times \tau$. 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.[30]

To account for possible heteroskedasticity, we weight the equations for county-level outcomes with the square root of the sample size used to compute the value (e.g., the total number of owner-occupied housing units for the county-level mean housing price).[31] The reported standard errors are clustered at the county level to allow for arbitrary serial correlation in residuals from the same county. Online Appendix Tables 4 and 8 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 VD.

There are differences in the number of pre- and post-fracking years across shale plays, including some that have none or very few post-fracking 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 nine shale plays with first-frac dates in 2008 or earlier. The subsequent analysis reports both treatment effects estimated using all available data and treatment effects estimated using the balanced sample. In the former sample, the years outside the balanced sample contribute to the identification of the county and play-year fixed effects.[32] Among these nine shale plays, there are a total of 65 top-quartile counties and 310 counties

[30] The variable $\tau_{pt} \cdot \mathbf{1}[\mathit{Rystad\ top\ quartile}]_c$ is collinear with the county-specific time trends, so that variable is dropped in these specifications.

[31] 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.

[32] The unbalanced sample comprises 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

outside the top quartile.[33] We report estimates of fracking's impact on outcomes evaluated 3, 4, or 5 years (depending on the data source) after fracking's initiation from this balanced sample.

B. *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 ACS to estimate long-difference models using the pooled 2009–2013 ACS as the post-period and the 2000 decennial Census as the pre-period.[34] The long-difference specification may be especially appealing in the case of housing prices. As discussed in Section IIB, asset prices quickly reflect information about the future, so with annual housing data, assigning a first-fracking date after information about fracking 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 fracking information is available anywhere in the country and the last year is after the estimated first-fracking date for the shale play where fracking arrived last, is appealing. The estimating equation is derived by first differencing equation (8), which gives

$$\begin{aligned}(10)\qquad & y_{cp,2013/2009} - y_{cp,2000} \\ &= \gamma_p + \delta\big(\mathbf{1}\big[\textit{Post-fracking}\big]_{pt} \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c\big) + \epsilon_{cpt}.\end{aligned}$$

The parameter δ is a difference-in-differences mean shift estimate of the effect of fracking, and maps directly to δ in equation (8).

---

the full sample, so, for example, the specification corresponding to equation (8) takes the following form to ensure that the treatment effects are identified from the balanced sample only:

$$\begin{aligned}(9)\qquad y_{cpt} = {} & \mu_{pt} + \gamma_c + \beta_1\tau \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c \\ & + \beta_2\big(\mathbf{1}\big[\textit{Unbalanced sample}\big]_{ct} \cdot \tau \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c\big) \\ & + \delta_0\big(\mathbf{1}\big[\textit{Post-fracking}\big]_{pt} \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c\big) \\ & + \delta_1\big(\tau \cdot \mathbf{1}\big[\textit{Post-fracking}\big]_{pt} \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c\big) \\ & + \delta_2\big(\mathbf{1}\big[\textit{Unbalanced sample}\big]_{ct} \cdot \mathbf{1}\big[\textit{Post-fracking}\big]_{pt} \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c\big) \\ & + \delta_3\big(\mathbf{1}\big[\textit{Unbalanced sample}\big]_{ct} \cdot \tau \cdot \mathbf{1}\big[\textit{Post-fracking}\big]_{pt} \cdot \mathbf{1}\big[\textit{Rystad top quartile}\big]_c\big) \\ & + \epsilon_{cpt}.\end{aligned}$$

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

[33] 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.

[34] 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.

Note that the long-difference approach is unable to adjust the estimates for differences in short-term preexisting trends in outcomes between the top-quartile and other counties within a play.[35]

C. *A Note on Prospectivity as a Potential Instrument*

The approaches discussed above are reduced form, and an alternative empirical approach would be to use the Rystad prospectivity score as an instrumental variable. However, we think that our reduced-form approach provides more economically interpretable and policy-relevant estimates for two reasons. First, policymakers likely don't have detailed information regarding the exact quantity of future oil and gas production when they decide whether to allow fracking. Instead, they are likely to be aware of whether their county has substantial fracking potential. Consequently, our reduced-form estimates of the average amenity and welfare impacts in high-fracking areas answer the policy-relevant question. Second, the theoretically correct endogenous variable for the key housing price regressions is the present value of *expected* hydrocarbon production from fracking because these markets are forward looking, but this variable is not directly observed. Consequently, estimates of WTP for amenity changes or for allowing fracking that rely on current production as the endogenous variable would be difficult to interpret. Further, it is difficult to develop meaningful estimates about future production since fracking remains a new technology and recovery rates are changing rapidly (see, e.g., Covert 2015). For these two reasons, we instead choose to focus on the reduced-form relationships between outcomes and our "instrument:" the interaction of an indicator for whether the county is in the top quartile of prospectivity within the shale play with a variable capturing when fracking techniques were adapted to that individual play. Nevertheless, despite these concerns about the interpretation of the instrumental variable estimates, for reference we have included two-stage least squares estimates of the effects of fracking where the value of current-year hydrocarbon production is the endogenous variable in online Appendix Section VB (see online Appendix Table 21 for the estimates themselves).

V. Results

A. *Oil and Natural Gas Production Effects*

We begin our empirical analysis with an event study-style version of equation (7), where the indicator variable, $\mathbf{1}[\textit{Post-fracking}]_{pt}$, 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-frack year in the relevant shale play. We plot the coefficients associated with the interaction of this vector and $\mathbf{1}[\textit{Rystad top quartile}]_{c}$; these coefficients measure the difference in outcomes between top-quartile and other counties

[35] The initiation of fracking 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. Online Appendix Section VIC describes the housing characteristics we use in more detail.

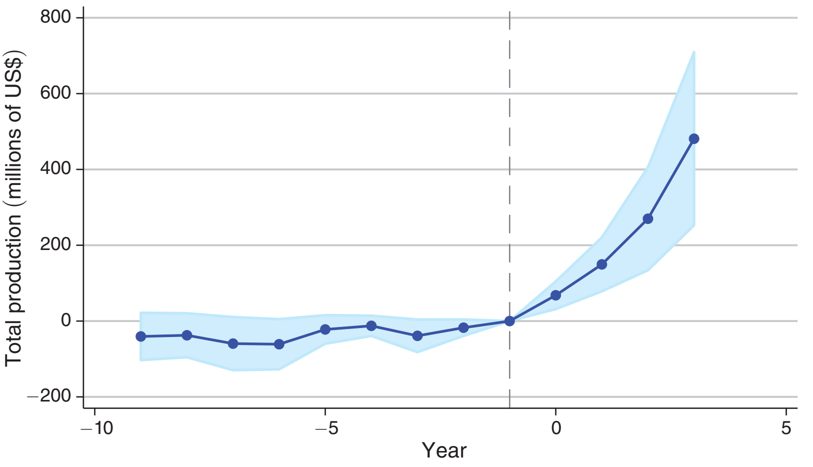

Figure 4. Event Study Analysis of County-Level Value of Hydrocarbons

Notes: This figure plots results from an event study analysis of the difference in the county-level value of hydrocarbon production between high-fracking-potential counties and other counties in shale plays before and after fracking began. The reported coefficients come from fitting a modified version of equation (7) where we interact $\mathbf{1}[\textit{Rystad top quartile}]_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frack 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 fracking in or before 2008 and event years common to all these shale plays (i.e., event years observed for all shale plays that began fracking in or before 2008). Data on hydrocarbon production from 1992 to 2011 come from Drillinginfo Inc. (2012). The shaded blue region shows 95 percent confidence intervals calculated using standard errors clustered at the county level.

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.

Figure 4 shows the evolution of 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 fracking techniques in the top-quartile counties, relative to the other counties. Additionally, the figure makes clear that, following the initiation of fracking, 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 summarizes the findings from Figure 4 more parsimoniously. 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-fracking years. Column 2 allows for differential pre-fracking event time trends in top-quartile counties and then includes a term to test whether these potentially differential top-quartile trends change after fracking is initiated. Column 3 uses the balanced sample of counties described above and replaces the top-quartile, pre-fracking

Table 3—Impact of Fracking on the Value of Hydrocarbon Production

	(1)	(2)	(3)
Total value of oil and gas production			
$\mathbf{1}(\textit{Fracking exposure}) \times \mathbf{1}(\textit{Post})$	242	36	36
	(68)	(47)	(23)
$t \times \mathbf{1}(\textit{Fracking exposure})$		3	
		(6)	
$t \times \mathbf{1}(\textit{Fracking exposure}) \times \mathbf{1}(\textit{Post})$		124	125
		(37)	(38)
Fracking exposure effect at $\tau = 3$	242	409	410
	(68)	(123)	(115)
Fracking exposure group	Top quartile	Top quartile	Top quartile
Control group	Quartiles 1–3	Quartiles 1–3	Quartiles 1–3
Fracking exposure level shift	Yes	Yes	Yes
Fracking exposure trend	No	Yes	Yes
Fracking exposure trend break	No	Yes	Yes
County fixed effects	Yes	Yes	Yes
County-specific trends	No	No	Yes
Year-play fixed effects	Yes	Yes	Yes
Restricted to balanced sample	No	No	Yes

Notes: This table reports regressions of oil/gas production variables on fracking exposure. Fracking 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 Drillinginfo, Inc. (2012) 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-fracking level shift, and a post-fracking 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-frack 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. If the year is after the first-frac date for the shale, defined as the first year that there is any fracking within the county's shale play, then $\mathbf{1}(\textit{Post}) = 1$. The coefficients and standard errors for Fracking Exposure Effect at $\tau = 3$ correspond to the $\mathbf{1}(\textit{Fracking exposure}) \times \mathbf{1}(\textit{Post})$ coefficient plus 3 times the $t \times \mathbf{1}(\textit{Fracking exposure}) \times \mathbf{1}(\textit{Post})$ coefficient. Standard errors clustered at the county level are reported in parentheses. *Sample:* Columns 1 and 2 include 8,100 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.

event time trend variable with county-specific calendar time trends. The bottom of the table reports the estimated treatment effect from each of these models three years after fracking begins.

It is apparent that the initiation of fracking 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 fracking 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 fracking in top-quartile counties. To put this estimated effect into some context, the median population in top-quartile counties prior to fracking activity is about 22,000, indicating an increase in hydrocarbon production of roughly \$19,000 per capita.

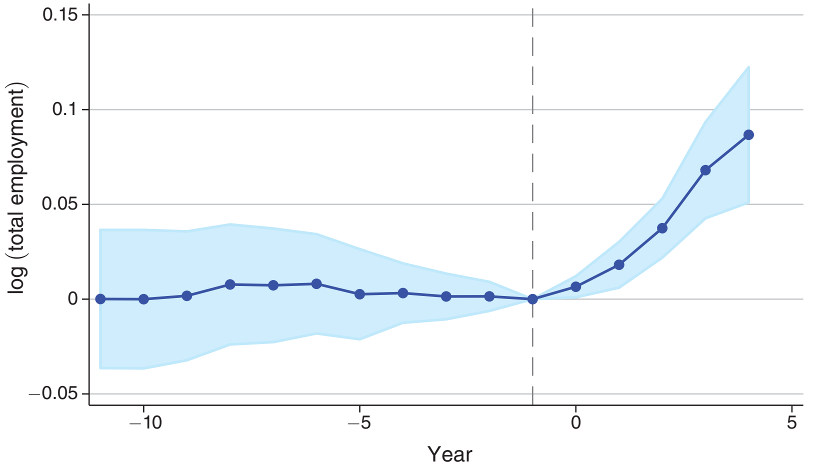

Figure 5. Event Study Analysis of Total Employment

Notes: This figure plots results from an event study analysis of the difference in log (total employment) between high-fracking-potential counties and other counties in shale plays before and after fracking began. The reported coefficients come from fitting a modified version of equation (7) where we interact $\mathbf{1}[Rystad\ top\ quartile]_c$ with a vector of event year indicators, τ_{pt}. Event years are defined as the calendar year minus the first-frack 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 fracking in or before 2008 and event years common to all these shale plays (i.e., event years observed for all shale plays that began fracking in or before 2008). Data on county-level total employment from 1990 to 2012 come from the Local Area Personal Income data from the REIS data produced by the BEA. Specifically, we use the variable CA25-10. The shaded blue region shows 95 percent confidence intervals calculated using standard errors clustered at the county level.

B. *Labor Market and Amenity Effects*

Figure 5 and online Appendix Figure 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 fracking's initiation. Since there are positive pre-trends for both outcomes, these graphs suggest that specifications that allow for differential pre-trends and a trend break after the initiation of fracking will produce the most reliable estimates.

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 reports only the estimated treatment effect four years after the initiation of fracking, 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 IRS county-to-county migration flows data.[36]

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

Table 4—Impact of Fracking on Employment and Aggregate Income: Time-Series Specifications

	(1)	(2)	(3)
Panel A. log(total employment)			
Fracking exposure effect at $\tau = 4$	0.037	0.055	0.051
	(0.016)	(0.029)	(0.019)
Panel B. Income			
log(total income)			
Fracking exposure effect at $\tau = 4$	0.049	0.061	0.033
	(0.015)	(0.031)	(0.024)
B1. log(total wage/salary income): 54 percent of total personal income			
Fracking exposure effect at $\tau = 4$	0.054	0.105	0.066
	(0.021)	(0.035)	(0.031)
B2. log(total rents/dividends): 18 percent of total personal income			
Fracking exposure effect at $\tau = 4$	0.077	0.109	0.085
	(0.020)	(0.041)	(0.030)
B3. log(total transfers): 10 percent of total personal income			
Fracking exposure effect at $\tau = 4$	0.011	0.001	−0.005
	(0.012)	(0.020)	(0.008)
B4. log(total proprietor's income): 18 percent of total personal income			
Fracking exposure effect at $\tau = 4$	0.073	−0.060	0.005
	(0.044)	(0.067)	(0.068)
Panel C. Migration			
C1. log(in-migration)			
Fracking exposure effect at $\tau = 4$	0.044	0.073	0.005
	(0.017)	(0.038)	(0.042)
C2. log(out-migration)			
Fracking exposure effect at $\tau = 4$	−0.001	0.007	−0.047
	(0.013)	(0.031)	(0.035)
Fracking exposure group	Top quartile	Top quartile	Top quartile
Control group	Quartiles 1–3	Quartiles 1–3	Quartiles 1–3
Fracking exposure level shift	Yes	Yes	Yes
Fracking exposure trend	No	Yes	Yes
Fracking exposure trend break	No	Yes	Yes
County fixed effects	Yes	Yes	Yes
County-specific trends	No	No	Yes
Year-play fixed effects	Yes	Yes	Yes
Restricted to balanced sample	No	No	Yes

Notes: This table reports regressions of aggregate economic outcomes on fracking 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 panels A and B come from the REIS data produced by the BEA. Migration measures in panel C come from the IRS's county migration data. Column 1 allows for a level shift in fracking-exposed counties. Columns 2 and 3 allow for pre-trends, a post-fracking level shift, and a post-fracking trend break in counties exposed to fracking. In columns 1 and 2, all fracking 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 plus 4 times the top-quartile trend break coefficient. Standard errors clustered at the county level are reported in parentheses. *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 9,246 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 8,694 observations from 378 total counties, of which 59 Rystad top-quartile and 237 outside-top-quartile counties are in the balanced sample. Panel B4, column 3 includes 4,736 observations from 296 total counties, of which 59 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 Fracking 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)
Fracking exposure group	Top quartile
Control group	Quartiles 1–3
Play fixed effects	Yes

Notes: This table reports long-difference regressions of the change in county aggregate economic outcomes between 2000 and 2009–2013 on a measure of fracking exposure. Fracking 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 is quartiles one through three. The fracking exposure measure is included by itself, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frack date after 2008. The reported estimates are for the balanced sample. Robust standard errors are reported in parentheses. *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.

Panels A and B indicate that Rystad top-quartile counties experience sharp improvements in economic activity after the initiation of fracking, relative to other counties in the same play. In the preferred specifications presented in columns 2 and 3, the estimates indicate increases in employment of about 5.1–5.5 percent. The income results reveal gains of 3.3–6.1 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 suggest 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 ACS 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. Panels A and B suggest a 4.8 percent increase in employment, a 2.6 percentage point gain in the employment-to-population ratio, a 0.6 percentage point decline in the

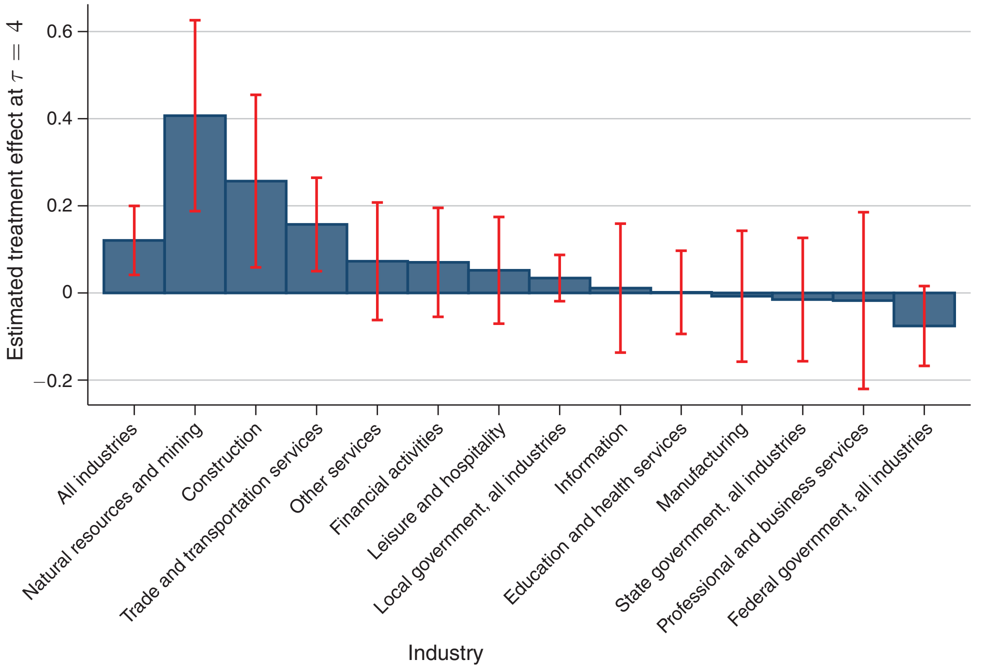

Figure 6. Employment Effects by Industry

Notes: This figure plots estimates of the effect of fracking on log (employment) by industry five years after the start of fracking. Each bar reports results of fitting equation (8) for the given industry, which corresponds to column 2 in the tables. Equation (8) allows for differential pre-trends in event time, as well as a trend break in outcomes and a mean shift for Rystad top-quartile counties. The model also includes play-year and county fixed effects. All Rystad top-quartile variables are interacted with an indicator for being in the unbalanced sample. The reported estimates correspond to the balanced sample. Data on employment by industry from 1990 to 2013 come from the QCEW produced by the Bureau of Labor Statistics (2014). Counties are included in the sample if the given employment variable is non-missing in all years from 1990 to 2013. Red bars report 95 percent confidence intervals calculated using standard errors clustered at the county level.

unemployment rate, and a 5.8 percent rise in mean household income.[37] Finally, panel C indicates that there was a 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 fracking 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. This is larger than the 4–5 percent increase in employment in Tables 4 and 5, but the QCEW assigns employment to a county according to the place of work, rather than the place of residence as is the case for the data files used in Tables 4 and 5.[38] Natural resources and mining is the industry

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

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

with the largest increase in employment, an increase of more than 40 percent. There are also statistically significant increases in employment in construction and transportation. No industry has a statistically significant decline in employment.[39]

Hydraulic fracturing is also likely to lead to changes in the composition of the workforce and population because many of the jobs associated with fracking are held by men in their 20s and 30s. Online Appendix Table 3 explores how the 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), as well as 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 fracking is associated with increases in crime rates that may be associated with an influx in young men (for example, see Edlund et al. 2013).[40] Online Appendix Figure 5 shows the event study plot for log violent crime. The estimates are imprecise and difficult to make strong conclusions from. Panels A, B, and C of Table 6 report the results of the same three specifications used in Tables 3 and 4 for total crime per hundred thousand residents, violent crime per hundred thousand residents, and property crime per hundred thousand residents, respectively. In the simplest model, without controls for differential pre-trends in top-quartile counties, the estimated effect on all kinds of crime per capita is positive, and is significant for violent crime. However, once controls for pre-trends are added, the estimates become less precise and the sign actually turns negative. This lack of precision and sensitivity to specification makes it difficult to come to firm conclusions regarding the effects of fracking on crime.[41]

We also attempted to ask whether air quality in top-quartile counties was affected by fracking-related activity. Unfortunately, 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,[42] only 13 of 65 top-quartile counties and 66 of 370 shale play counties have non-missing data in all years between 2000 and 2011.

[39] Despite the large estimated increase in wage and salary income per household in Table 5, which might make manufacturing firms less competitive in fracking counties, the estimated change in manufacturing employment is very small. There are a few possible explanations. One is that given capital adjustments costs and other frictions, any effect on manufacturing may appear only a number of years after fracking starts. Alternatively, lower natural gas prices may help keep local manufacturers competitive despite the rise in wages. Fetzer (2014) proposes this channel and finds evidence consistent with lower natural gas prices being an important mechanism keeping manufacturing in fracking counties.

[40] See http://geology.com/articles/oil-fields-from-space/.

[41] Note that although the effects of fracking on crime rates are unclear, the estimated effects of fracking on the total level of crime are more consistent. In results available upon request, fracking is estimated to increase the total level of all types of crime in all specifications.

[42] The measure is average total suspended particulate matter (TSP), imputed using PM10 or PM2.5 when TSP is not available.

Table 6—Impact of Fracking on Crime

	(1)	(2)	(3)
Panel A. Total crime per 100,000 residents			
Fracking exposure effect at $\tau = 5$	172	−96	−171
	(201)	(138)	(192)
Panel B. Violent crime per 100,000 residents			
Fracking exposure effect at $\tau = 5$	56	29	−4
	(28)	(66)	(75)
Panel C. Property crime per 100,000 residents			
Fracking exposure effect at $\tau = 5$	116	−125	−166
	(177)	(123)	(209)
Fracking exposure group	Top quartile	Top quartile	Top quartile
Control group	Quartiles 1–3	Quartiles 1–3	Quartiles 1–3
Fracking exposure level shift	Yes	Yes	Yes
Fracking exposure trend	No	Yes	Yes
Fracking exposure trend break	No	Yes	Yes
County fixed effects	Yes	Yes	Yes
County-specific trends	No	No	Yes
Year-play fixed effects	Yes	Yes	Yes
Restricted to balanced sample	No	No	Yes

Notes: This table reports regressions of crime rates on fracking exposure. Fracking 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 fracking exposure variables are included by themselves, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frack date after 2008. The reported estimates are for the balanced sample. Crime data come from the FBI UCR system. Crime reports from law enforcement agencies are aggregated to the county level. Data from a law enforcement agency are included only 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-fracking level shift, and a post-fracking 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-frack 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 plus five times the top-quartile trend break coefficient. Standard errors clustered at the county level are reported in parentheses. *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 7,480 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.

C. *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 (10). The estimates suggest that fracking 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 by roughly 24 percent, and welfare and hospital expenditures by about 24 percent (although this

Table 7—Impact of Fracking 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 (elementary/secondary education spending per pupil)	0.008
	(0.034)
Fracking exposure group	Top quartile
Control group	Quartiles 1–3
Play fixed effects	Yes

Notes: This table shows regressions on the change in government spending and revenues between 2002 and 2012 on fracking 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 is quartiles one through three. The fracking 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 Censuses of Governments. Panels A1, A2, 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. *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.

increase is not statistically significant by conventional criteria). Interestingly, we find only a small, and noisily estimated, 2.5 percent increase in education expenditures. Panel D, which reports the change in log elementary and secondary education per pupil, shows 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, other revenues of 26 percent, and sales tax revenues of 59 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. This is consistent with recent case study evidence from Newell and Raimi (2018a, b), although they find important heterogeneity across municipalities, which we also explore further below.[43]

Overall, the Table 7 results indicate that fracking 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.

D. *Robustness*

We gauge the robustness of the results to alternative definitions of fracking exposure and alternative approaches to controlling for local economic shocks. Panels A and B of online Appendix Table 4 and panel B of online 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 fracking. 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

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

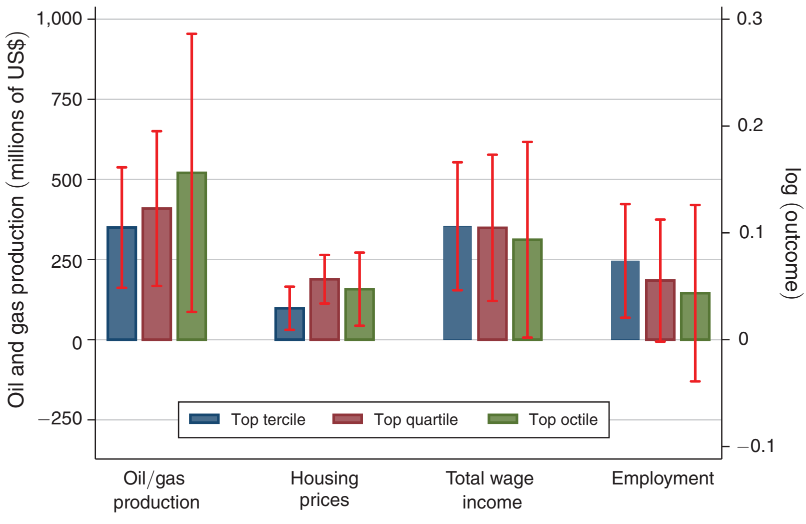

Figure 7. Estimates by Fracking Exposure Measure

Notes: This figure plots estimates of the effect of fracking using different definitions of fracking exposure: being in the top tercile, quartile, or octile of maximum prospectivity within the shale play. Estimates are presented for four outcomes: millions of dollars of oil and gas production, median housing prices for owner-occupied housing units, total wage and salary income, and total employment. The bars for hydrocarbon production, total wage and salary income, and employment report results of fitting equation (8) for the given outcome variable, which corresponds to column 2 in the tables. Equation (8) allows for differential pre-trends in event time, as well as a trend break in outcomes and a mean shift for Rystad top-quartile counties. The model also includes play-year and county fixed effects. The bar for median housing prices reports results of fitting equation (10). 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 hydrocarbon production from 1992 to 2011 come from Drillinginfo, Inc. (2012). Data on county-level total employment and wage and salary income from 1990 to 2012 come from the Local Area Personal Income data from the REIS data produced by the BEA. Data on median housing prices come from the decennial Census and the ACS (Manson 2018). Red bars report 95 percent confidence intervals calculated using standard errors clustered at the county level.

(Conley 1999).[44] 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. Figure 7 then graphically reports results from online Appendix Tables 4 and 8, showing how our estimates vary with different measures of fracking exposure for our four key outcomes variables: hydrocarbon production, housing values, total wage and salary income, and employment.

Panel A of online Appendix Table 4 suggests that the conclusions about the effect of fracking on hydrocarbon production are qualitatively unchanged by these alternative approaches. It is reassuring that the estimated effect is increasing in the

[44] 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 kilometers 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.

stringency of the indicator definition for fracking amenability in the cases of both the maximum- and mean-based definitions. Furthermore, the estimates are larger for the maximum-based definition. The Conley standard errors tend to be larger than conventional ones, but their use does 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 of online Appendix Table 4 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 by allowing the Eagle Ford to influence the estimated treatment effect. When the maximum Rystad prospectivity score is used, fracking is estimated to increase total income by 6–9 percent, and the effect is statistically significant in seven of the nine 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.[45] Panel B of online 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 fracking. The broader lesson here seems to be that even within shale plays, the economic benefits of fracking are concentrated in the subset of counties that are most suitable for drilling, although the imprecision of the estimates makes definitive conclusions unwarranted.[46]

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 fracking include these spillovers, which may result, for example, from individuals commuting from non-top-quantile to top-quantile counties. 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. We investigate these possibilities in Section VIB.

E. *Housing Price and Quantity Estimates*

A central component of the welfare calculation is the effect on the housing market. Table 8 provides details on the results of these regressions. Panel A reports on the impact of fracking's initiation in Rystad top-quartile counties from the estimation of the long-difference-in-differences specification outlined in equation (10). The estimates indicate that median and mean housing values for owner-occupied homes increased by 5.7 percent due to fracking. Further, the

[45] Panel A of online 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.

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

Table 8—Impact of Fracking on Housing Outcomes

	(1)
Panel A. Housing values	
A1. Median housing value	0.057
	(0.018)
A2. Mean housing value	0.057
	(0.018)
A3. Mobile housing units: median housing value	0.079
	(0.037)
Panel B. Rental prices	
B1. Median rental price	0.020
	(0.010)
B2. Mean rental price	0.029
	(0.011)
Panel C. Housing quantities	
C1. Total housing units	0.011
	(0.012)
C2. Total mobile homes	0.022
	(0.028)
C3. Share of housing units vacant	−0.010
	(0.005)
C4. Acres of agricultural land	−0.099
	(0.144)
Fracking exposure group	Top quartile
Control group	Quartiles 1–3
Play fixed effects	Yes

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 fracking exposure. Fracking 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 is quartiles one through three. The fracking exposure measure is included by itself, as well as interacted with an indicator for being in the unbalanced sample, defined as having a first-frack date after 2008. The reported estimates are for the balanced sample. Housing data for 2013–2009 come from the ACS. Housing data for 2000 come from the decennial census. Agricultural land data for 2002 and 2012 come from the 2002 and 2012 Censuses of Agriculture, respectively. All housing values are converted to 2010 dollars. Observations are weighted by the number of owner-occupied (renter-occupied) units in the county. Non-mobile-specific regressions are adjusted for changing owner-occupied (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. *Sample:* Included are 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.

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.[47]

[47] Online 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.

Panel C of online Appendix Table 4 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.[48] In total, 17 of the 18 estimates fall in a range of roughly 2 to 6 percent. 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 remain significant at a 95 percent level. Overall, we conclude that the initiation of fracking led to meaningful increases in housing prices in counties especially amenable to fracking, relative to other counties in the same shale play.

These estimated effects on county-level housing prices are large, relative to the magnitude of the effects of other substantive local changes on housing prices that have been documented in the previous literature. It is instructive that there is an extensive literature documenting the capitalization of various amenities into local housing prices and that 5.7 percent is a large county-level effect.[49] For example, Chay and Greenstone (2005) finds 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, Cellini, Ferreira, and Rothstein (2010) finds that school facility investments lead to 4.2–8.6 percent increases in housing prices but over the smaller geographic unit of school districts. While Currie et al. (2015) finds 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, but not statistically significant.[50] It is noteworthy, however, that the vacancy rate for housing units declined by 1.0 percentage point.[51]

[48] Although adding state fixed effects tends to reduce the estimated effect of fracking on housing prices, online Appendix Table 7 shows that adding state fixed effects does not dramatically influence many of the point estimates in the individual plays. The most notable change is that the estimate of the impact of fracking on housing prices for the Marcellus is reduced from roughly 9 percent to about 6 percent.

[49] The 5.7 percent average effect obscures important within-county variation in housing price changes, and, indeed, this is an important finding in Muehlenbachs, Spiller, and Timmins (2015).

[50] 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 online Appendix Table 13, we report specifications where we replace 2002 with 1997 as the base year. The point estimate for the effect of fracking on agricultural land quantities becomes 0.067 and is, again, imprecisely estimated. The sensitivity of the agricultural land results suggests that they must be interpreted with caution.

[51] A shortcoming of the housing supply data is that the end-of-period data are an average calculated from 2009–2013, and this includes several years when fracking was only in its early stages in some shale plays. Hence, we also examined the effect of the initiation of fracking on the number of housing unit construction permits issued. Online Appendix Figure 6 is an event study graph that suggests that there was an increase in permits with the introduction of fracking, but that this increase does not become apparent until three years after fracking was initiated (panel C of online Appendix Table 2 shows the same findings in a regression framework).

F. *Heterogeneity across Shale Plays*

The empirical design also allows us to estimate play-specific effects from fracking. We report on the nine shale plays with first-frack dates in 2008 or before that are included in the pooled results.[52] The event study plots for hydrocarbon production (online Appendix Figure 7) suggest that in eight of the nine shale plays in the pooled results, hydrocarbon production in top-quartile counties was largely flat prior to fracking and then took off after the commencement of fracking.[53] The lone exception is the Woodford-Anadarko play, which for largely idiosyncratic reasons experienced an increase in production in advance of fracking and a decline afterward.[54]

Table 9A reports the econometric results across the nine shale plays. Here, we focus on three sets of outcomes: hydrocarbon production, labor market-related outcomes such as average household income, and housing prices.[55] To help us understand these results, column 1 of Table 9B reproduces the overall estimates for the outcomes in Table 9A and column 2 of Table 9B reports the F-statistic and associated p-value from a test that the nine shale estimates in Table 9A are equal.

As suggested by the event study graphs, we estimate large increases in hydrocarbon production in seven of the nine plays; the estimates are statistically significant in five of them. Similarly, we estimate sizable increases in income per household in seven of nine plays, four of which are statistically significant. In contrast, the gains in housing prices appear to be concentrated in two of the nine plays. Specifically, the housing 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 the three key outcome variables (total value of hydrocarbon production, household income, and home values). With only nine observations, it is difficult to make precise statements about the sources of the observed heterogeneity.

[52] In the online Appendix, we also include the Eagle Ford shale play. Although fracking began there in 2009, which is beyond the cutoff for our pooled results, the Eagle Ford, located in the southern part of Texas, has attracted a lot of attention, so we report results on the Eagle Ford for completeness.

[53] The ten shale plays reported in online Appendix Figure 7 are the nine plays with first-frac dates in 2008 or before, plus the Eagle Ford, which has a first-frac date in 2009.

[54] 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. Consequently, even if top-quartile counties are expected to have much more fracking than others, with only one draw there is a nontrivial 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 fracking, 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.

[55] Given the substantial heterogeneity suggested by these results, it is also interesting to explore whether this heterogeneity extends to other outcomes. Online 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 fracking on the Bakken have been much larger than the effects of fracking on other plays. Table 5 also reports results for the Eagle Ford, which has a first-frac date in 2009, and so is not included in the results in the main text.

TABLE 9A—PLAY-SPECIFIC ESTIMATES

	Bakken (1)	Barnett (2)	Fayetteville (3)	Haynesville (4)	Marcellus (5)	Woodford-Anadarko (6)	Woodford-Ardmore (7)	Woodford-Arkoma (8)	Permian plays (9)
Panel A. Average characteristics of top-quartile counties									
Population (2000)	6,307	109,202	24,046	24,576	112,911	45,516	19,537	9,955	15,221
Oil share of hydrocarbon production value (2011)	0.94	0.42	0.00	0.01	0.07	0.34	0.48	0.01	0.64
Panel B. Hydrocarbon production									
B1. Total value of hydrocarbon production	972	322	69	1,730	185	−452	123	199	169
	(414)	(183)	(78)	(903)	(70)	(65)	(70)	(158)	(134)
Panel C: Labor markets									
C1. log (mean household total income)	0.293	0.045	0.099	0.080	0.049	0.069	−0.013	0.000	0.170
	(0.083)	(0.025)	(0.110)	(0.053)	(0.012)	(0.084)	(0.079)	(0.134)	(0.049)
C2. log (mean household wage and salary income)	0.286	0.031	−0.014	0.078	0.078	0.079	−0.028	0.075	0.177
	(0.100)	(0.030)	(0.133)	(0.064)	(0.014)	(0.102)	(0.095)	(0.161)	(0.059)
C3. log (mean household rent, dividend, and interest income)	0.833	0.061	0.671	0.078	0.086	−0.171	0.116	0.495	−0.006
	(0.313)	(0.095)	(0.417)	(0.201)	(0.045)	(0.319)	(0.297)	(0.505)	(0.183)
C4. log (population)	0.130	0.071	−0.014	−0.045	0.018	0.060	0.042	−0.038	−0.007
	(0.045)	(0.053)	(0.115)	(0.055)	(0.024)	(0.117)	(0.075)	(0.089)	(0.039)
Panel D: Housing prices									
D1. log (median home values)	0.228	−0.046	0.018	−0.071	0.089	−0.074	−0.032	0.051	0.029
	(0.086)	(0.030)	(0.111)	(0.057)	(0.014)	(0.091)	(0.082)	(0.138)	(0.051)
Panel E: Annual change in WTP for amenities and welfare per household, using change in mean home values (dollars)									
E1. Change in amenities	−1,695	−3,399	28	−2,971	−1,083	−2,925	2,765	614	−4,638
	(2,298)	(3,140)	(624)	(1,620)	(953)	(887)	(2,742)	(986)	(1,686)
E2. Change in welfare	11,694	−838	3,430	−157	2,804	−580	2,426	3,685	1,431
	(2,496)	(1,931)	(944)	(2,038)	(910)	(841)	(2,569)	(1,360)	(2,207)
Top-quartile counties	8	5	1	5	28	1	4	2	11
Outside-top-quartile counties[a]	27	41	13	21	95	10	5	7	34

Notes: This table reports play-specific summary statistics and estimates of the effect of fracking for plays with first-frac dates in or before 2008. Panel A reports the average population and oil share of hydrocarbon production of top-quartile counties within each shale play."

Panels B, C, and D show estimates from regressions of outcome variables on Rystad top-quartile variables interacted with dummies for being in particular shale plays. Panel B reports time-series specifications corresponding to equation (8). Panel B allows for pre-trends, a level shift, and a trend break in the top-quartile indicators, and also includes 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 three years after the start of fracking. 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-occupied (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-frack date after 2008. The reported coefficients are for the balanced sample. Panel B data come from Drillinginfo, Inc. (2012) well data aggregated to the county level. Panel C and D data come from the 2000 decennial census and the 2009–2013 ACS. 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.

Panel E reports estimates of the effect of fracking 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 idiosyncratic preferences for location (s) of $\beta = 0.65$ and $s = 0.30$, respectively. Panel E shows estimates where the change in housing costs is measured using the estimated percentage change in mean 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 fracking 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.

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

Table 9B—Joint F-Test of Play-Specific Estimates

	All (1)	Joint F-test (2)
Panel A. Average characteristics of top-quartile counties		
Population (2000)	64,860	
Oil share of hydrocarbon production value (2011)	0.33	
Panel B. Hydrocarbon production		
B1. Total value of hydrocarbon production	409	F-stat: 11.4
	(123)	p-value: 0.00
Panel C. Labor markets		
C1. log (mean household total income)	0.058	5.4
	(0.012)	p-value: 0.00
C2. log (mean household wage and salary income)	0.075	5.5
	(0.012)	p-value: 0.00
C3. log (mean household rent, dividend, and interest income)	0.093	1.7
	(0.038)	p-value: 0.09
C4. log (population)	0.027	1.4
	(0.016)	p-value: 0.20
Panel D. Housing prices		
D1. log (median home values)	0.057	6.0
	(0.012)	p-value: 0.00
Panel E. Annual change in WTP for amenities and welfare per household, using change in mean home values (dollars)		
E1. Change in amenities	−1,405	
	(734)	
E2. Change in welfare	2,537	
	(758)	
Top-quartile counties	65	
Outside-top-quartile counties[a]	253	

Notes: This table reports summary statistics and estimates of the effect of fracking for the overall sample. Column 1 shows the summary statistic or estimate for all counties with first-frac dates in or before 2008. Column 2 presents results from the joint F-test that the estimates are equal for all counties with first-frac dates in or before 2008 (when relevant). Panel A shows summary statistics on county population and the oil share of hydrocarbon production.

Panels B, C, and D show estimates from regressions of key outcome variables on Rystad top-quartile variables in the overall sample and then reports joint F-tests for whether the estimated effects for each outcome are equal across shale plays. 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 includes 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 three years after the start of fracking for panel B and four years after the start of fracking 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-occupied (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 Drillinginfo, Inc. (2012) well data aggregated to the county level. Panel C and D data come from the 2000 decennial Census and the 2009–2013 ACS. 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.

Panel E reports estimates of the effect of fracking on amenities and welfare in dollars. 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 idiosyncratic preferences for location (s) of $\beta\ =\ 0.65$ and $s\ =\ 0.30$, respectively. Panel E shows estimates where the change in housing costs is measured using the estimated percentage change in mean 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 fracking 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.

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

G. *Quantitative Comparison to the Literature*

Before turning to the welfare analysis, it is important to place these reduced-form results in the context of the larger literature. We note from the outset, though, that the meaning of these comparisons is limited by the fact that this paper relies on a new research design, and we believe it is the most comprehensive in terms of coverage of shale plays and outcome variables. Nevertheless, there are some striking similarities and dissimilarities to previous work. We specifically focus on the three papers that estimate the most similar parameters—Boslett, Guilfoos, and Lang (2016); Feyrer, Mansur, and Sacerdote (2017); and Jacobsen (2019)—and discuss how they compare to our estimates. Feyrer, Mansur, and Sacerdote (2017) is interested in estimating the geographic extent of the labor market impacts of fracking. The authors instrument for local oil and gas production using predicted values from a model of county-level oil and gas production. They find that a $1 million increase in oil and gas production per capita increases earnings in exposed counties by $79,751 per capita, or 2.1 percent per $10,000 in oil and gas production per capita. In our instrumental variable estimates reported in online Appendix Table 21, we find that a $10,000 increase in oil and gas production per capita increases earnings per capita by a similar 3.1 percent.

Boslett, Guilfoos, and Lang (2016) exploits variation in exposure to fracking from New York's 2008 moratorium on fracking in New York State. These authors use house transaction data from five counties along the New York and Pennsylvania border, near the area of northeastern Pennsylvania and southern New York that has high fracking potential. They find that housing prices rise 10.1 percent more in Pennsylvania counties where fracking is allowed relative to New York counties where it is not allowed. As before, these results are broadly consistent with our finding that fracking increased housing prices in exposed counties, but the magnitude differs somewhat (we find a 5.7 percent effect on housing values).

To our knowledge, Jacobsen (2019) is the only other paper in the literature to explore both the labor and the housing market impacts of fracking. Defining areas exposed to fracking based on the ex post change in oil and gas production, Jacobsen (2019) finds that in nonmetropolitan areas more exposed to fracking, wage and salary income per capita rise 13.8 percent, population rises 3.9 percent, home values rise 9.9 percent, and rents rise 3.4 percent.[56] These patterns are qualitatively similar, but all of the estimates are 50–100 percent greater in magnitude than the corresponding difference-in-differences estimates in this paper that rely on geological and time variation.

The estimates in all three of the papers discussed in this section are qualitatively similar to those in this paper: the papers consistently find that fracking increases housing values 5 to 12 percent, increases wage and salary income 7 to 15 percent, and increases population 2 to 4 percent. However, in every case, although the order of magnitudes is the same, the exact quantitative values differ nontrivially. There are three possible explanations for these differences. First, the Boslett, Guilfoos, and

[56] For comparison, we use Jacobsen's (2009) estimates for 2011, which most closely correspond to our estimates, which usually average outcomes between 2009 and 2013.

Lang (2016) and Jacobsen (2019) papers use different samples, with Boslett, Guilfoos, and Lang (2016) using a sample from only five counties in New York and Jacobsen (2019) using only smaller, nonmetropolitan areas. If there are heterogeneous treatment effects, then fracking may have different effects in nonmetropolitan areas of New York and Pennsylvania than in other regions. Indeed, our estimate of the effect of fracking on housing prices for the Marcellus shale, which covers New York and Pennsylvania, in Table 9A is 8.9 percent, which is much closer to the estimate in Boslett, Guilfoos, and Lang (2016). The larger estimated effects of fracking on local labor market outcomes found in Jacobsen (2019) could be driven by larger effects of fracking in smaller labor markets. Even in the case of Feyrer, Mansur, and Sacerdote (2017), which is also national in scope, their empirical strategy relies on different sources of variation than ours, which can also result in different estimated effects of fracking. Second, in the case of reduced-form studies like Jacobsen (2019) or Boslett, Guilfoos, and Lang (2016), the implicit first-stage effect on oil and gas production may be different. Finally, as described above, although all three papers use difference-in-difference style strategies as in this paper, the underlying source of variation in exposure to fracking is different in each of these papers, and readers will naturally make their own judgments about the credibility of the various research designs.

VI. Interpretation and Local Welfare Consequences of Fracking

What are the net local welfare consequences of fracking? To this point, the paper has reported on a wide range of outcomes; some indicate that, on average, Rystad top-quartile counties have benefited from the initiation of fracking, while others reveal less positive impacts. Guided by the conceptual framework outlined in Section I, this section develops measures of WTP for the change in local amenities and for the net local welfare consequences of the initiation of fracking, using the estimated changes in housing prices and rents, income, and population from the previous section.

A. *Local Welfare Estimates*

While there is little question that fracking increases local productivity, a central question in the debate about fracking 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 change caused by fracking, as well as the WTP for the change in amenities. We use the local labor market model developed in Section I. 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 (5), WTP for the change in amenities can be expressed as

$$(11) \qquad \alpha\widehat{\Delta \ln A_{at}} = s\widehat{\Delta \ln N_{at}} - \left(\widehat{\Delta \ln w_{at}} - \beta\widehat{\Delta \ln r_{at}}\right),$$

where $\Delta \ln N$ is the change in local population, 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), $\Delta \ln r$.[57] 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.

To gain further intuition about the expression of WTP for amenity changes, consider the case where WTP for amenities is zero: in this case, 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, that is, $(\widehat{\Delta \ln w_{at}} - \beta\widehat{\Delta \ln r_{at}})/s$, 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 moving costs/locational preference (i.e., 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. With these estimates in hand, it is straightforward to develop an estimate for the WTP for allowing fracking (i.e., the net welfare change for original residents) by using equation (6), which also incorporates income from lease payments received by households, and the previous section's empirical estimates. Table 10 reports this too. The entries in columns 1–4 report the mean annual WTP for the change in amenities and for allowing fracking per household for households originally living or owning land in top-quartile counties. The entries in columns 5–8 report the aggregate present value of the WTP measures for original households in all 65 top-quartile counties using a 5 percent discount rate and assuming that fracking is allowed permanently and that the estimated annual changes in amenities, income, housing costs, etc. are constant and last forever. Columns 1–2 and 5–6 use the increase in rental prices (2.9 percent) as the measure of the change in housing costs, and columns 3–4 and 7–8 use the increase in housing prices (5.7 percent).

All estimates in the table assume that $\beta = 0.65$, the share of household wage and salary income spent on locally produced goods, following Albouy (2008), and that $s = 0.30$, the standard deviation of idiosyncratic location preferences or moving costs, which is the population-share weighted average of the values for

[57] 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 fracking, then $\Delta \ln p_j = (1/(1-\beta))\Delta \ln r$ and the percentage changes 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; expectations about future growth associated with fracking; or segmented housing markets.

Table 10—Welfare Estimates

Annual impacts per household (in $ per household)				Aggregate impacts for top-quartile counties (in billions of $)			
Δ in housing costs = 2.9%		Δ in housing costs = 5.7%		Δ in housing costs = 2.9%		Δ in housing costs = 5.7%	
WTP for change in:							
Amenities (1)	Welfare (2)	Amenities (3)	Welfare (4)	Amenities (5)	Welfare (6)	Amenities (7)	Welfare (8)
Panel A. All households							
−2,225	1,716	−1,405	2,537	−74	57	−47	85
(842)	(610)	(734)	(758)	(28)	(20)	(24)	(25)
Panel B. Owner-occupants							
—	1,716	—	2,537	—	41	—	61
—	(610)	—	(758)	—	(15)	—	(18)
Panel C. Renters							
—	363	—	363	—	3	—	3
—	(429)	—	(429)	—	(4)	—	(4)
Panel D. Absentee landlords							
—	1,354	—	2,174	—	13	—	21
—	(432)	—	(590)	—	(4)	—	(6)

Notes: This table reports estimates of the effect of fracking on amenities and welfare in dollars. Different panels report values for different household types, with panel A reporting results for all household types and panels B, C, and D reporting results for how the change in welfare is split between households in owner-occupied housing units, renter-occupied housing units, and units with absentee landlords, respectively. Columns 1–4 report results in terms of annual income per household. Columns 5–8 aggregate these figures to the total impact of fracking 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. Columns 1–2 and 5–6 report results where the change in housing costs is measured using the estimated percent change in mean rents (0.029), while columns 3–4 and 7–8 show estimates where the change in housing costs is measured using the estimated percentage change in mean home prices. For each measure of the change in housing costs, we report both the estimated change in amenities (columns 1, 3, 5, and 7) and the estimated change in total welfare (columns 2, 4, 6, and 8). All columns use our preferred parameter values for the standard deviation of household idiosyncratic preferences and the share of income spent on housing of $s = 0.30$ and $\beta = 0.65$, respectively. The calculations are converted to dollars using the mean household wage and salary income in top-quartile counties of $45,668 and mean household interest and dividend income in top-quartile counties of $3,822. Standard errors incorporate both sampling variance in the estimated parameters and uncertainty in the values of s and β. For these calculations, the assumed standard deviations are 0.09 for idiosyncratic preferences (s) and 0.1 for the share of income spent on housing (β). For more information, see discussion in the text.

non-college-educated and college-educated workers of 0.27 to 0.47, respectively, estimated in Diamond (2016).[58] Throughout, we use the estimated 7.5 percent change in mean wage and salary income, a 13.1 percent change in interest and dividend income, and a 2.7 percent change in population (based on the Table 5 results).[59]

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

[59] Presumably, rent and dividend incomes from fracking are received primarily by households that own houses and can benefit from lease payments resulting from their mineral rights. Consequently, we need to scale up our estimate of the effect of fracking on rent and dividend income to account for the fact that not everyone owns their

There is substantial uncertainty in the true values of s and β. To account for this fact, we incorporate uncertainty in the values of s and β into the standard errors of our amenity change and welfare estimates. Specifically, we assume that s has a standard deviation of 0.09 (the average of the standard error of the point estimates for college-educated and non-college-educated workers in Diamond 2016), that β has a standard deviation of 0.1, and that the covariance between these terms and between these parameters and our estimated parameters is 0.

These assumptions can then be used, along with the sampling variation in $\Delta \ln N_{at}$, $\Delta \ln w_{at}$, Δr_{at}, and Δy_{at}, to compute the standard error of the estimates of WTP for the amenity change and of WTP for allowing fracking (i.e., welfare). As described in more detail below, the standard errors are relatively small, giving us greater confidence in the estimates. Further, online Appendix Table 20 provides a different set of amenity and welfare estimates when alternative reasonable values of s and β, different measures of the $\Delta \ln N_{at}$, $\Delta \ln w_{at}$, and Δr_{at}, and alternative empirical specifications are used; the qualitative patterns hold, suggesting that values of these key parameters do not drive the results. However, it must be kept in mind that the exact magnitudes of the amenity and welfare effects of fracking will depend on the true values of s and β, which are ultimately unknown.

The Table 10 estimates suggest that the initiation of fracking decreases local amenities. Using the preferred assumptions, the estimated annual WTP is −\$1,405 per household when the change in housing prices is used as a proxy for local prices and −\$2,225 when the change in rental rates is used. The estimated effects on amenities are precisely estimated, with the standard errors allowing us to rule out zero effect on amenities. If we assume that the decline in amenities is permanent, then the present value of the decline in local amenities is −\$47 billion using housing prices and −\$74 billion using rental rates.[60] Finally, we note that, in principle, these estimates capture 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 fracking 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 fracking equals about \$1,700 to \$2,500 per household annually (i.e., 3.3 to 4.9 percent of mean annual household income).[61] As is the case with the estimated WTP for amenities, the estimated welfare effects are quite precise, with the estimated standard errors allowing us to rule out positive welfare effects smaller than \$500 per household. If the changes in amenities and economic opportunities are permanent, columns 6 and 8 suggest that the aggregate increase in welfare is in the neighborhood of \$57 billion to \$85 billion for the top-quartile Rystad counties.

home. The estimated 13.1 percent change in rent and dividend income is the 9.3 percent change estimated in Table 5 multiplied by one, divided by the share of households that own their houses in Rystad top-quartile counties.

[60] This calculation uses the 2000 census population for each county.

[61] We use the mean annual household income in 2000 of counties in shale plays that are outside the top quantile of \$51,818 in 2010 (US\$).

It is instructive to compare the welfare gains implied in Table 10 with those that a standard Roback analysis would produce. Specifically, the estimated welfare gain is about \$160 million in the average county, or \$10.4 billion among all top-quartile counties, when the welfare change is equal to the change in home values as in the canonical Roback model. The reason for this much smaller estimated welfare effect 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. The results from this paper's model and the standard Roback model both indicate that the value of the greater economic opportunities outweighs the decline in local amenities, suggesting large and meaningful average net gains for the top-quartile counties.

Are these estimates plausible? Recall that our estimate of the impact of the introduction of fracking 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 with a 5 percent discount rate. There are 65 top-quartile counties, so the estimated national welfare gain of \$57 billion to \$85 billion is a little over 15 percent of the discounted value of the increase in hydrocarbon production of \$520 billion from these counties. Thus, at least with this basis for 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 fracking. 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, although decisive evidence would likely require more detailed micro data.[62]

It is nevertheless possible to provide some preliminary evidence on heterogeneity in the welfare measures by homeownership and across shale plays. Panels B–D of Table 10 explore how the estimates of WTP for allowing fracking differ for individuals who own their own home (and work in the labor market), renters, and absentee landlords.[63] It is striking that the WTP for allowing fracking is five to seven times higher among homeowners than it is for renters. Of course, these welfare

[62] Using longitudinal census data, Bartik and Rinz (2018) provides information on one dimension of heterogeneity, occupation, and finds that residents originally working in manual-intensive occupations gained much more from the introduction of fracking than others.

[63] The welfare change for each renting household is equal to the sum of the change in wages and amenities subtracting off the change in rents. Plugging in equation (11) for the change in amenities, the change in amenities for renters simplifies to $s\Delta \ln N$. As a result of this simplification, the change in welfare for renters does not depend on the change in rents and does not vary depending on which measure of the change in rents is used. The change in welfare for absentee landlords is equal to the sum of the increase in rents and the increase in rent and dividend payments. The welfare change for owner-occupied households is equal to the sum of the welfare change for renters and absentee landlords, as residents who are owner-occupants can be thought of as absentee landlords who rent the housing to themselves.

estimates should be viewed cautiously, because they require even stronger assumptions, including the assumption that renters and owners have identical preferences and abilities to benefit from fracking's labor market effects and that all landlords live outside of fracking-exposed areas and are unexposed to the labor market or amenity effects of fracking. Even with these assumptions in mind, these results suggest that the welfare effects of fracking are likely to vary substantially across individuals, even within top-quartile counties.

In Table 9A, panel E, we report the estimated change in WTP for amenities and for allowing fracking separately by shale play. The estimates are qualitatively consistent across shale plays, with six of nine shale plays experiencing declines in amenities or quality of life and six of nine experiencing welfare improvements. The largest estimates come from the Bakken's (primarily in North Dakota and Montana) annual WTP of \$11,700 and the Woodford-Arkoma's (Oklahoma) annual WTP of \$3,700, although there are also large net gains in the Fayetteville (in Arkansas and Oklahoma), the Woodford-Ardmore, and the Marcellus (largely in Pennsylvania, West Virginia, and Ohio). Interestingly, the shale plays with negative welfare estimates have values that are small in magnitude and statistically insignificant. Overall, the play-specific estimates are very demanding of the data and hence substantially less precise than the aggregate estimates.

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 variables would be important predictors of WTP to allow fracking. Among the three largest gainers, one (Bakken) is dominated by petroleum and the other two (Fayetteville and Marcellus) are dominated by natural gas production. 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 that influence the welfare estimates. Overall, it is apparent that the question of where fracking offers the largest net benefits cannot be answered decisively with just ten data points.[64]

Two final points are noteworthy. First, these revealed preference estimates of WTP to allow fracking (and for amenity changes) are ultimately determined by households' knowledge. If new information causes households to update their estimates of fracking'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 fracking reflect only local changes in welfare. The national and global welfare effects of fracking include potentially very important consequences for petroleum, natural gas, and electricity prices; local air pollution; global warming; and geopolitics, as well as general equilibrium effects on labor markets. All of these impacts are outside the scope of this paper; however, none of them become relevant if local communities do not allow fracking within their jurisdictions.

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

B. *Spillovers*

To the extent that fracking activity in a county has spillover effects on other counties in the same shale play, our identification strategy will underestimate the benefits and costs of fracking. To investigate this hypothesis, we also explored models using propensity-score-matching to select the "control" counties to compare to both counties in the top quartile of the Rystad prospectivity measure and counties in the lower three quartiles (Imbens and Rubin 2015). As discussed above, we were unable to select control counties with covariates that balanced using propensity scores. We present these propensity-score estimates in online Appendix Section D.3.1 and online Appendix Tables 14 through 17. The estimates do suggest gains in income and employment in the bottom three quartiles, though they are smaller in magnitude than for the top-quartile counties. These results suggest that spillovers may in fact mean that our method understates the magnitude of both the benefits and the costs of fracking, but the imbalance in covariates means that these results should be interpreted cautiously.

VII. Conclusions

This paper has developed a measure of the net welfare consequences of fracking on local communities that accounts for both its benefits and costs. To do so, we utilize a new identification strategy based on geological variation in shale deposits within shale plays and differences in the timing of the initiation of fracking across plays. Further, we set out a Roback-Rosen-style locational equilibrium model and use it to derive an expression for WTP for allowing fracking in a local community that is a function of the parameters that can be estimated with the identification strategy.

There are three primary findings. First, counties with high fracking potential experience a boom in oil and natural gas production. This boom is characterized by sharp increases in a broad set of economic indicators, including gains in total income (3.3–6.1 percent), employment (3.7–5.5 percent), housing prices (5.7 percent), and housing rental rates (2.9 percent). Second, there is evidence of deterioration in the noneconomic quality of life or total amenities, including higher violent crime rates. Using the model's results, we estimate that annual WTP for fracking-induced changes in local amenities is roughly equal to −\$1,400 per household annually, or −2.7 percent of mean annual household income. Third, the net welfare effects of allowing fracking appear to be substantial and positive for local communities. Again using an expression derived from the model, we estimate that across all US shale plays, WTP for allowing fracking is about \$2,500 per household annually, or about 4.9 percent of mean household income among original households in counties with high fracking potential. Importantly, there is also evidence of substantial heterogeneity in WTP across shale plays.

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 fracking 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 deposits both in North America and in 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 fracking have benefited on average, although there is evidence of important heterogeneity in the local net benefits. Understanding the sources of this heterogeneity is a first-order question for researchers and policymakers interested in assessing the impacts of allowing fracking in their community.

REFERENCES

Albouy, David. 2008. "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas." NBER Working Paper 14472.

Allcott, Hunt, and Daniel Keniston. 2018. "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America." *Review of Economic Studies* 85 (2): 596–731.

Bartik, Alexander W., Janet Currie, Michael Greenstone, and Christopher R. Knittel. 2019. "The Local Economic and Welfare Consequences of Hydraulic Fracturing: Dataset." *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20170487.

Bartik, Alexander W., and Kevin Rinz. 2018. "Moving Costs and Worker Adjustment to Changes in Labor Demand: Evidence from Longitudinal Census Data." https://alex-bartik.squarespace.com/s/JMP_21feb2018_mod-bwx7.pdf.

Boslett, Andrew, Todd Guilfoos, and Corey Lang. 2016. "Valuation of Expectations: A Hedonic Study of Shale Gas Development and New York's Moratorium." *Journal of Environmental Economics and Management* 77: 14–30.

Budzik, Philip, and Steven P. Mathis. 2013. "Hydraulic Fracturing: Breaking Rocks to Unlock Energy." *American Society of Mechanical Engineers* webinar, December 5, 2013.

Bureau of Economic Analysis (BEA). 1990–2014. *Regional Economic Information System: Local Area Personal Income and Employment*, CA04 and CA25 series, Washington, DC: US Bureau of Economic Analysis. https://www.bea.gov/regional/downloadzip.cfm (accessed on July 17, 2016).

Bureau of Labor Statistics. 2014. "Quarterly Census of Employment and Wages: County High-Level QCEW NAICS-Based Data Files from 1990–2013." https://www.bls.gov/cew/datatoc.htm.

Bureau of Labor Statistics. 2015. "Table 24: Historical Consumer Price Index for All Urban Consumers (CPI-U): U.S. City Average, All Items." In *CPI Detailed Report—January 2015*, edited by Malik Crawford, Jonathan Church, and Bradley Akin, 91–94. Washington, DC: US Department of Labor.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103 (2): 897–947.

Cascio, Elizabeth U., and Ayushi Narayan. 2019. "Who Needs a Fracking Education? The Educational Response to Low-Skill Biased Technological Change." NBER Working Paper 21359.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125 (1): 215–61.

Census Bureau. 2014a. *Building Permits Survey*. https://www.census.gov/construction/bps/ (accessed May 23, 2009).

Census Bureau. 2014b. *Census of Governments: Finances*. https://www.census.gov/programs-surveys/gov-finances/data/datasets.all.html (accessed November 4, 2014).

Chay, Kenneth Y., and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy* 113 (2): 376–424.

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

Covert, Thomas R. 2015. "Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale." http://home.uchicago.edu/~tcovert/webfiles/fracking.pdf.

Covert, Thomas, Michael Greenstone, and Christopher R. Knittel. 2016. "Will We Ever Stop Using Fossil Fuels?" *Journal of Economic Perspectives* 30 (1): 117–38.

Currie, Janet, Michael Greenstone, and Katherine Meckel. 2017. "Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania." *Science Advances* 3 (12): e1603021.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker. 2015. "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings." *American Economic Review* 105 (2): 678–709.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000." *American Economic Review* 106: 479–524.

Drillinginfo, Inc. 2012. *DI Desktop: US Oil and Gas Well Production Data.* Used with permission. https://info.drillinginfo.com/products/di-basic/.

Edlund, Lena, Hongbin Li, Junjian Yi, and Junsen Zhang. 2013. "Sex Ratios and Crime: Evidence from China." *Review of Economics and Statistics* 95 (5): 1520–34.

EIA (**Energy Information Administration**)**.** 2011. *Maps: Exploration, Resources, Reserves, and Production*. https://www.eia.gov/maps/maps.htm.

EPA (**Environmental Protection Agency**)**.** 2015. *Assessment of the Potential Impacts of Hydraulic Fracturing for Oil and Gas on Drinking Water Resources—Executive Summary*. Washington, DC: EPA, Office of Research and Development.

FBI (**Federal Bureau of Investigation**)**.** 2015. *Uniform Crime Records* (*UCR*) *and Law Enforcement Officers Killed and Assaulted* (*LEOKA*) *1992–2014.* https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.-2015/resource-pages/aboutcius_final.

Fetzer, Thiemo. 2014. "Fracking Growth." Centre for Economic Performance (CEP) Working Paper 1278.

Feyrer, James, Erin T. Mansur, and Bruce Sacerdote. 2017. "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution." *American Economic Review* 107 (4): 1313–34.

Fitzgerald, Timothy. 2013. "Frackonomics: Some Economics of Hydraulic Fracturing." *Case Western Law Review* 16 (4): 1337–62.

Gopalakrishnan, Sathya, and H. Allen Klaiber. 2013. "Is the Shale Energy Boom a Bust for Nearby Residents? Evidence from Housing Values in Pennsylvania." *American Journal of Agricultural Economics* 96 (1): 43–66.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy* 118 (3): 536–98.

Ground Water Protection Council and ALL Consulting. 2009. *Modern Shale Gas Development in the United States: A Primer*. Washington, DC: US Department of Energy, Office of Fossil Energy, National Energy Technology Laboratory.

Hornbeck, Richard, and Enrico Moretti. 2019. "Who Benefits From Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality." faculty.chicagobooth.edu/richard.hornbeck/research/papers/TFP_HM.pdf.

Hsiang, Solomon M. 2010. "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America." *Proceedings of the National Academy of Sciences* 107 (35): 15367–72.

Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* 1st ed. New York: Cambridge University Press.

IRS (**Internal Revenue Service**)**.** 2015. "U.S. Population Migration Data." Statistics on Income (SOI) Tax Stats 1990–2012. https://www.irs.gov/statistics/soi-tax-stats-migration-data (accessed August 11, 2015).

Jacobsen, Grant D. 2019. "Who Wins in an Energy Boom? Evidence from Wage Rates and Housing." *Economic Inquiry* 57 (1): 9–32.

Kline, Peter, and Enrico Moretti. 2014. "People, Places and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs." *Annual Review of Economics* 6: 629–62.

Maniloff, Peter, and Ralph Mastromonaco. 2014. "The Local Economic Impacts of Hydraulic Fracturing and Determinants of Dutch Disease." Colorado School of Mines, Division of Economics and Business Working Paper 2014-08.

Manson, Steven, Jonathan Shroeder, David Van Riper, and Steven Ruggles. 2018. "IPUMS National Historical Geographic Information System: Version 13.0 [Database]." IPUMS NHGIS, Minneapolis, MN. http://doi.org/10.18128/D050.V13.0 (accessed June 20, 2018).

McCarthy, Kevin, Katherine Rojas, Martin Niemann, Daniel Palmowski, Kenneth Peters, and Artur Stankiewicz. 2011. "Basic Petroleum Geochemistry for Source Rock Evaluation." *Oilfield Review* 23 (2): 32–43.

Moretti, Enrico. 2011. "Local Labor Markets." In *Handbook of Labor Economics*, Vol. 4B, edited by Orley Ashenfelter and David Card, 1237–313. Amsterdam: North Holland.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The Housing Market Impacts of Shale Gas Development." *American Economic Review* 105 (12): 3633–59.

NCES (**National Center for Education Statistics**). 2015. "Common Core Data." Local Education Agency (School District) Universe Survey (1987, 1992, 1997, 2002, 2007, 2012). https://nces.ed.gov/ccd/pubagency.asp (accessed September 25, 2015).

NETL (**National Energy Technology Laboratory**). 2013. *Modern Shale Gas Development in the United States: An Update*. Washington, DC: US Department of Energy.

Newell, Richard G., and Daniel Raimi. 2018a. "US State and Local Oil and Gas Revenue Sources and Uses." *Energy Policy* 112: 12–18.

Newell, Richard G., and Daniel Raimi. 2018b. "The Fiscal Impacts of Increased U.S. Oil and Gas Impacts on Local Governments." *Energy Policy* 117: 14–24.

Phillips, Susan. 2014. "Pennsylvania's Frack Ponds Now Number More than 500." *StateImpact Pennsylvania* (*NPR*), November 11. https://stateimpact.npr.org/pennsylvania/2014/11/11/pennsylvanias-frack-ponds-now-number-more-than-500/.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–78.

Rubinstein, Justin L., and Alireza Babaie Mahani. 2015. "Myths and Facts on Wastewater Injection, Hydraulic Fracturing, Enhanced Oil Recovery, and Induced Seismicity." *Seismological Research Letters* 86 (4): 1060–67.

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

Suárez Serrato, Juan Carlos, and Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106 (9): 2582–624.

Weber, Jeremy G. 2012. "The Effects of a Natural Gas Boom on Employment and Income in Colorado, Texas, and Wyoming." *Energy Economics* 34 (5): 1580–88.

Weinstein, Al. 2014. "Local Labor Market Restructuring in Shale Booms." *Journal of Regional Analysis and Policy* 44 (1): 71–92.

Wynveen, Brooklynn J. 2011. "A Thematic Analysis of Local Respondents' Perceptions of Barnett Shale Energy Development." *Journal of Rural Social Sciences* 26 (1): 8–31.

Zagorski, William A., Gregory R. Wrightstone, and Douglas C. Bowman. 2012. "The Appalachian Basin Marcellus Gas Play: Its History of Development, Geologic Controls on Production, and Future Potential as a World-Class Reservoir." In *Shale Reservoirs—Giant Resources for the 21st Century*, edited by J.A. Breyer, 172–200. Tulsa, OK: American Association of Petroleum Geologists.

This article has been cited by:

1. Alexander Bartik, Scott Nelson. 2016. Credit Reports as RRsumms: The Incidence of Pre-Employment Credit Screening. *SSRN Electronic Journal* . [Crossref]