Economic Development and Public Health

Alina Kovalenko *

May 2018

Abstract

Although rapid development of an area generally leads to better economic outcomes like higher standards of living and wages, it can have some negative social impacts as well. Using recent economic shocks associated with localized fracking booms, this paper documents one such externality - increased incidence of risk-taking behaviors. I exploit plausibly exogenous geographic distribution of shale deposits and temporal expansion of the drilling activity in the Marcellus region. Using detailed county-level data from 2002-2015, I find that counties with fracking potential are associated with higher rates of gonorrhea infections (an increase of 12%), as well as arrests for disorderly conduct and drunkenness (3% and 5% respectively). The pattern of results is consistent with substantial positive income effects and a small change in the composition of workers.

*University of Texas at Austin, alinakov@utexas.edu

1 Introduction

The inflows of physical, human, and financial capital that accompany the development of new technologies often have disproportionate impacts on specific geographic areas even when the technology itself is widely distributed. For example, even though people all across the country benefit from being able to purchase cheaper gas when a new source of oil is developed, it is the area with the physical drilling infrastructure that experiences the largest economic impact.

These developments can significantly alter the size and composition of economies in which they occur and are the focus of the "resource boom" literature. They can lead to higher-paying jobs, increased property values, and more government tax revenue (Bartik et al., 2017; Black et al., 2005). These effects can also be negative, however, resulting in pollution, crime and disruption of otherwise smoothly operating environments (James and Smith, 2017; Howarth et al., 2011). In particular, understanding the public health effects associated with such rapid developments is an important research question for policymakers and the affected communities.

Estimating these effects is challenging because the development of new technologies is usually highly correlated with other characteristics of the areas being developed that are responsible for the observed effects. The advancement of hydraulic fracturing ("fracking") and horizontal drilling technologies provides a useful natural experiment to answer this question. Until the development of these methods, there was no feasible way to extract the oil and gas reserves located in shale formations. Thus in contrast to easily accessible resources such as lumber or coal, the geographic distribution of shale deposits is unlikely to be related to the development of the communities on top of them. In other words, fracking could be viewed as an exogenous shock to the health outcomes of people living in areas affected by it.

Using detailed county-level data from 2002-2015, this paper examines how the recent shale development has affected local STD rates in the states of Ohio, Pennsylvania, and West Virginia. I exploit the plausibly exogenous geographic distribution of shale deposits and the temporal expansion of the fracking activity in a difference-in-difference-style model. The results demonstrate that shale-rich counties are associated with higher gonorrhea rates, while chlamydia rates remain similar to areas with no oil and gas deposits. Specifically, I find that counties on top of shales experience on average an increase of 4-5 new cases of gonorrhea compared to counties without any fracking potential, or approximately 12% relative to the average rate in my sample. Furthermore, using STD data broken down by gender and age in PA, I find that the effects are almost entirely concentrated in females between the ages of 15 and 24.

The findings are robust to a number of alternative specifications including changing the timing of the policy and using the intensity of actual drilling activity for defining treatment. This study offers policy implications for local communities and policy makers facing large inflows of human and physical capital. Given the positive estimates derived in this paper, the communities should anticipate a rise in STD rates following a resource boom. To the extent that increased infection rates could be averted through education and medical screening, opening additional health clinics and designing a public information campaign may be useful strategies.

This paper proceeds as follows. Section I provides the background on the fracking boom. Section II gives an overview of the existing literature. In sections III and IV, I describe the data and econometric specifications. Section V discusses main results and robustness checks, and section VI concludes.

2 Background

Over the past decade, the shale revolution - a result of technological breakthroughs in horizontal drilling and hydraulic fracturing - has dramatically changed the energy industry in the US resulting in an explosion of well drilling and oil and gas production. As shown in Figure 1, in the early 2000s most of the fuels were produced from vertically drilled (i.e. conventional) wells. By 2011, this trend reversed and most production came from horizontally drilled (i.e. fracked) wells.

The boom has also fueled a structural transformation of local economies throughout the entire country creating job opportunities and increasing incomes.

In this paper, I focus on the Marcellus and Utica shale formations which at least partially cover the states of Pennsylvania, Ohio, and West Virginia (Figure 2). Although this area had some history of conventional oil and gas production, it was not until the recent fracking boom that the activity became particularly intensive. Prior to mid-2000s, oil and gas production was relatively low (Figure 3). However, by 2015, there were over 7 trillion cubic feet of natural gas and nearly 41 million barrels of crude oil produced, an increase of around 900% and 350% respectively. Furthermore, since January 2012, natural gas production in the Marcellus and Utica regions has accounted for 85% of the increase in total natural gas production in the U.S.[1] It is also reflected in increased employment in oil and gas extraction industry, shown in the bottom left panel of Figure 3. The total number of jobs in the sector remained relatively constant at about 23,000 prior to the fracking boom. However, oil and gas employment went up from 23,979 in 2005 to 59,201 in 2015, coinciding with the timing of the shale energy boom.

While fracking is an effective way of accessing massive reserves, it is a more complex and labor-intensive process compared to conventional extraction methods.[2] It requires the ability both to construct the wells and support infrastructure, and afterwards to maintain the operation process around-the-clock. Brundage et al. (2011) estimate that the drilling of one new well involves 420 people across 150 occupations which is equivalent to 13.1 full time workers had they all performed all the tasks themselves. The specific nature of work increased the demand for low-skill young male workers, some of which come from non-local areas seeking temporary employment (Cascio and Narayan, 2015). Moreover, most upstream field workers work twelve hour shifts two-thirds of the time, normally on a 14 days on and 7 days off rotation. [3]

[1] Drilling Productivity Report, EIA.

[2] It is also a more capital-intensive process - one conventional vertical well requires an investment of between \$1 and \$3 million but one unconventional well can cost up to \$5 to \$8 million (https://blogs.siemens.com/measuringsuccess/stories/688/)

[3] Horizontal Drilling Workforce Study, PSAC: http://www.psac.ca/wp-content/uploads/hor_drill_wf_study.pdf

The rapid expansion and unique working conditions in the fracking industry provide a favorable environment for unhealthy behaviors and in particular, a higher prevalence of STDs. While empirical studies are limited, there is an abundance of anecdotal evidence to suggest that fracking booms raise public safety and health concerns in nearby communities. For example, in The Atlantic article titled "How Fracking Is Bad for Our Bodies" Jason Silverstein (2013) writes:

> *When mostly male newcomers flood a small town, they overwhelm community services. After all, the new population isn't met with more doctors, police officers, or teachers. [...] There's often a spike in crime rates, dropouts, alcohol abuse, sexually transmitted infections, and mental health problems.*

Such news stories are not rare. In his New York Times article, John Eligon (2013) describes the experience of the doctor practicing in the oil boomtown who says that "since the oil industry started growing rapidly in the region, he has had to treat many more sexually transmitted diseases." A number of descriptive case studies report that STDs have increased in areas where transient workers, mostly young men, enter a new town en masse to work in the oil and gas industry. For example, in its case study, Food and Water Watch found that in Pennsylvania the average number of cases of STDs was 62 percent higher in heavily drilled counties than in counties with no drilling [4]. A spike in STDs has also been reported in highly drilled areas in other states, including Carrizzo Springs, Texas; Mesa County, Colorado; and McKenzie County, North Dakota (Vaughan 2012; Eligon 2013).

In this study, I focus on two sexually transmitted infections - gonohrea and chlamydia. Chlamydia trachomatis infection is the most commonly reported nationally notifiable disease in the United States, with more than 1,500,000 cases cases reported to state and local health departments in 2015. The second most commonly reported disease is N. gonorrhoeae, with approximately 820,000 new cases occurring in the United States each year. Chlamydia is known

[4] The Social Costs of Fracking: A Pennsylvania Case Study: https://www.foodandwaterwatch.org/insight/social-costs-fracking

as a “silent” infection as most infected people are asymptomatic, while gonorrhea symptoms are relatively observable. In addition, unlike other STDs, gonorrhea has a short incubation period making it a better approximation of contemporaneous sexual behavior. The highest reported rates of these infections are among sexually active teenagers and young adults. Both diseases can be cured with antibiotics, although antimicrobial resistance in gonorrhea is becoming an increasing concern.

There are several theoretical mechanisms that could explain the relationship between fracking booms and STD rates. The first one comes through the composition of the population. If unconventional oil and gas industry attracts out-of-state workers, and if these workers are more likely to have STDs than the existing population, then STD rates in fracking counties will increase. Additionally, there is a reason to think that STD rates might increase if agglomeration or peer effects are present. While this paper does not explicitly investigate these mechanisms, I provide some preliminary evidence in section V.

3 Literature Review and Contribution

The economics literature contains many studies that look at what happens to a local economy following a rapid expansion of an industry or a large labor demand shock. Local shocks linked to natural resources in particular have been especially appealing as they are sudden, substantial in magnitude, and geographically focused, three features that contribute to a more credible empirical identification.

For the purposes of this paper, empirically identifying the relationship between rapid economic development and public health requires variation in growth of areas that is exogenous to health outcomes and other local conditions. Exploring mainly economic outcomes, previous work has relied on structural modeling (Hausman and Kellogg, 2015), border discontinuity (Boslett et al, 2016) and endowment-based approaches (Black et al, 2015; Alcott and Keniston, 2015; Michaels, 2011). Endowment-based design is especially attractive since if only pure geological characteristics of areas are used, the spacial variation comes entirely from natural factors dating back hundreds of years that are arguably unrelated to other local conditions.

The empirical approach I take most closely follows that exploited by Black et al. (2005) looking at the boom in the Appalachian coal industry in 1970s and Michaels (2011) focusing on the effects of oil abundance of the Southern US counties over the period of 1940-90. As both of these events occurred more than four decades ago when economic and social conditions were different, it is not clear how they relate to the current state of events. Unlike these papers, I exploit one of the largest local economic shocks in recent decades.

Local shocks induced by fracking have gained considerable attention, with numerous studies linking shale booms to increased employment and earnings (Jacobsen and Parker, 2016; Feyrer et al., 2017), environmental concerns (Howarth et al., 2011) and depressed housing values (Maniloff and Mastromonaco, 2014).[5] However, comprehensive studies exploring the indirect negative consequences of regional development are limited to public safety and crime, and the existing evidence is scant and mixed. While Putz et al. (2011) fail to find evidence that energy booms lead to higher crime rates in North Dakota, Haggerty et al. (2014) show a positive relationship between crime rates and resource booms in the western US. The most recent paper by James and Smith (2017) focuses on the booming regions throughout the US and is similar in its empirical approach to this study as they also divide areas into groups based on the shale resources underneath them. The authors find higher rates of property and violent crime in shale-rich counties compared to counties without any oil and gas extraction potential.

Despite the growing recent literature looking at the effects of localized fracking booms, to the best of my knowledge no study has looked at their impacts on the incidence of sexually transmitted diseases. I build on the previous research by using the quasi-natural experiment

[5] There are also a number of recent working papers exploiting localized fracking booms to look at educational outcomes (Cascio and Narayan, 2015), net welfare consequences (Bartik et al., 2017), and marriage rates (Kearney and Wilson, 2017)

of fracking breakthrough combined with detailed county-level data to investigate the effects of localized rapid shale development on public health of nearby communities. In doing so, I am also able to highlight some important heterogeneity across such demographic characteristics as age and gender. Lastly, this study emphasizes the importance of documenting unintended consequences of economic decisions. Understanding these factors and quantifying their effects will help identify effective policy interventions aimed to improve public health and safety.

4 Data and Summary Statistics

To explore the relationship between fracking booms and the prevalence of STDs, I would ideally have annual data at the county level on STD rates both before and after the increase in drilling activity, and preferably split by age, gender, race, and education level. Likewise, to measure the variation and intensity of fracking activity, I need frequent and reliable data on drilling and production of oil and gas at the county level. The best currently available data do not meet all of these requirements, but come close, as I describe below.

The county-level STD data for 2002-2015 come from the Centers for Disease Control and Prevention (CDC) NCHHSTP AtlasPlus application. [6] The STD measures are compiled from reports sent to CDC's Division of STD Prevention by the STD control programs and health departments in all 50 states. I focus on gonorrhea and chlamydia rates per 100,000 population defined as the number of new cases for the calendar year divided by population for that calendar year and then multiplied by 100,000. [7] The population denominators used to compute these rates are based on the U.S. Census Bureau population estimates. Unfortunately, the CDC county-level STD data cannot be broken down by demographic group (e.g. age, gender, or race). However, I obtained gonorrhea and chlamydia rates by gender and age from the Pennsylvania Health

[6] I requested earlier STD data from CDC as well, however it did not cover most of the sample, and I don't include it in my analysis.

[7] I don't focus on syphilis and HIV rates because the data for these outcomes are not disclosed for many counties due to the low number of cases.

Department, and I incorporate these data into my analysis as well.

Maps and shapefiles of the shale basins come from the Energy Information Agency (EIA). I overlay them with the county maps to determine whether a county lies on top of the shale. The detailed oil and gas data are generously provided by DrillingInfo under a special agreement I obtained with the help of the Economics Department at UT Austin. DrillingInfo is a private information services company which collects monthly well-level data from various state agencies. To identify fracking activity, I split the data by the type of well - horizontal, directional, vertical, or unknown - and following Feyrer et al. (2017), I classify horizontal and directional wells as drilled by fracking, and vertical wells as conventional. To be conservative, I group unknown well types with vertical wells in the conventional category. Geologically, oil and gas deposits are often co-located; therefore, in my analysis I consider both. I collect the data on the number of new wells spudded [8] and then aggregate it to the county-year level. [9] The distribution of the total number of spudded wells over the sample period by county and year is shown in Figures 4 and 5.

Employment and income data come from the Regional Economic Accounts of the Bureau of Economic Analysis (BEA). This employer-reported series covers more than 95% of US jobs. BEA's estimates of total employment are derived from the Bureau of Labor Statistics data with adjustments to account for employment not covered, or not fully covered, by the state unemployment insurance and the unemployment compensation for Federal employees programs. Lastly, I supplement my main sample with the county-level demographic data from the The Census Bureau's Population Estimates Program. These include population counts by age group, gender, and race.

The main sample includes counties in Ohio, Pennsylvania, and West Virginia. This region is particularly interesting and important as it lies on top of the largest shale formation in the US,

[8] In oil and gas industry, to "spud" a well means to begin drilling operations.

[9] In the database, this variable comes from both completion and production reports, and is stored separately depending on the source. As recommended by the DrillingInfo support team, I extract information on spudded wells from production reports, since it is consistently reported and available for the entire sample period.

and accounts for 40% of domestic shale gas production. Moreover, the area is more representative of the US as a whole in terms of economic and demographic characteristics. In line with existing research on resource booms, I exclude urban counties with over 250,000 residents to make the control group more comparable to the treatment group. This reduces the main sample to 2,590 county-year observations. Table 1 displays descriptive statistics for the whole sample, and control and treatment groups separately in the period before fracking started as well as after its expansion. I discuss it in more detail in the next section.

5 Empirical Model

5.1 Empirical Specification

The identification strategy used in this study relies on the fact that shale oil and gas deposits are unevenly distributed, generating substantial geographic variation in the fracking potential of different counties. Another source of variation comes from the timing of the onset of fracking activity, influenced by the heterogeneity in formations' geology and exploration techniques. Together, the interaction between geographic and temporal factors forms the basis for a difference-in-difference design. The basic analysis begins with the following regression model to examine the relationship between STD rates and fracking booms:

$$STD_{ct} = \alpha + \beta(Treat_c \times Post_t) + \gamma_c + \gamma_t + X_{ct}\theta + \epsilon_{ct}, \tag{1}$$

where the main outcome of interest, STD_{ct}, is the rate of gonorrhea or chlamydia in county c and year t. [10] The variable $Treat_c$ is an indicator equal to 1 if center of county c lies on top of the shale, and 0 otherwise. As shown in Figure 5, fracking technology became prevalent in the late 2000s. Thus, for the purposes of my analysis I will date the introduction of fracking to

[10] I do not use a log transformation because of the occurrence of zero observations in my sample.

2007, and therefore I set $Post_t$ equal to 1 for all years from 2007 onward.[11] Year fixed effects, represented by γ_t, control for observed and unobserved factors that are common to all counties in a given year. Likewise, γ_c are county fixed effects, which account for observed and unobserved differences across counties that are constant over time. [12] The vector X_{ct} represents county-level demographic and economic control variables (e.g. age, race, gender, population, income and employment measures). The idiosyncratic error ϵ_{ct} might be correlated within a county over time, so I account for it by clustering the standard errors at the county level.

The coefficient of interest, β, measures the average annual difference in STD rates in treatment counties relative to control counties, after fracking was developed, relative to before its initiation. Assuming the anecdotal evidence presented in earlier sections is true, I would expect β to be positive: in years following the boom, relative to years prior, counties with high fracking intensity should have seen greater growth in STD rates than counties not affected by the boom. Otherwise, if fracking has no significant effect on STD rates, I expect β to be 0. The advantage of this approach is that it generates variation in fracking potential that we don't expect to be driven by pre-existing characteristics affecting STD rates. However, not all of the counties lying on top of shales experienced extraction booms. Therefore, I am only able to identify the intent-to-treat effect (ITT), which could be underestimated compared to the effect of treatment on the treated (TTE).

Because counties lying on shales were unable to access them due to technological constraints before the implementation of fracking, we would not expect their development to be affected by the oil and gas deposits they contain. I next provide evidence in support of the research strategy and this identification assumption.

[11] I check the validity of this choice in the robustness section.

[12] I do not include county-specific trends in the model because they might over-control for the response to the treatment, an econometric issue discussed by Wolfers (2006). When I do estimate the models with county-specific trends included, the estimates decrease in size.

5.2 Identification

In this section, I discuss issues that could lead OLS estimation of Equation 1 to produce biased estimates of β. The key identification assumption in the difference-in-difference model is that STD rates in treatment and control groups would follow the same time trend in the absence of shales and fracking. While this assumption is difficult to prove directly, I attempt to explore its validity by conducting a number of checks. I begin by looking at whether the counties in the treatment and control groups are statistically indistinguishable in terms of their observed characteristics and outcomes of interest. Column 4 of Table 1 shows that STD rates, demographic and employment characteristics were similar in the treatment and control groups before treatment started. Out of the 15 estimated differences in background characteristics, the only ones that are significant at 1% are the differences in income per capita and percent of adult population with less than high school education. Also, as expected there was a small number of wells drilled in the treated counties in the pre-period which reflects the presence of exploratory drilling operations.

While similar levels of observable characteristics suggest that they might also be similar between the two groups in changes, it is important to check directly how the two groups evolve over time. This includes checking, for example, the possibility that counties on top of oil and gas shales might have been on higher trajectories in terms of STDs prior to the introduction of fracking. If this is the case, then the estimation of the impact of shale development on health outcomes would result in overestimation. In figure 6, I graph the STD rates separately for counties that lie on top of oil and gas shales and counties that do not. The graphs shows that gonorrhea and chlamydia rates for both groups moved relatively in parallel in the pre-period. Starting from 2009, gonorrhea rates for counties lying on top of oil and gas shales increased more than those for the counties that do not, coinciding with the surge in fracking activity. On the other hand, the comparative trends in the incidence of chlamydia infection rates do not seem to show such divergence.

To investigate pre-existing trends more rigorously, I use pre-fracking data from 2002 to 2006 to estimate the following model:

$$STD_{ct} = \alpha + \beta Treat_c + \lambda(Treat_c \times Year_t) + \gamma_t + \epsilon_{ct}. \tag{2}$$

In this specification, I regress STD outcomes on an indicator for whether the county is on top of shale, year dummies and an interaction between this indicator and year dummies. The results from estimating this equation are reported in Table 2. The coefficients on the interaction terms of the treatment indicator with the year dummies indicate whether there was a significant difference in STD rates between control and treatment counties in each year (compared to the omitted year dummy for 2002). As columns 1 and 2 show, the estimates are not statistically different from zero. I also note that some are positive and others are negative, lacking any consistent pattern. This conclusion is supported by the fact that, based on the F-tests presented in the bottom of the table, I cannot reject the hypothesis that all the interaction terms are jointly equal to zero. The presented evidence suggests that both groups were on similar trends of STD outcomes in the five years prior to the fracking boom.

6 Results

6.1 Main Results

I begin by examining the effect of the rapid development on STD rates by estimating equation (1). In doing so, I use Ohio, Pennsylvania and West Virginia counties grouped by geological characteristics. Table 3 presents estimates for gonorrhea and chlamydia rates defined as the number of new cases per 100,000 residents (adjusted for population in a given county). Columns 1 and 3 present the basic specification with county and year fixed effects only, while columns 2 and 4 include other control variables.

Overall, the results show a positive effect of fracking activity on both gonorrhea and chlamydia rates, although it is only statistically significant for the former. The coefficient estimate on the interaction variable can be interpreted as the average increase in the number of new STD cases per 100,000 population for a given county that is lying on top of shale compared to counties that are not. For example, the first column in Panel A suggests that a county on top of oil and gas deposits experiences on average an increase of 5 new cases of gonorrhea compared to counties without such reserves. Given the average rate of 40 per 100,000 people, this suggests that the fracking boom increases gonorrhea rates by roughly 12% relative to the average rate for the treatment group observed in my sample in the pre-fracking period. This effect is significant at the 5% level. Adding controls to the main specification yields similar estimates, although the significance drops to 10%. [13]

There is little evidence that drilling activity led to increased chlamydia rates with all specifications yielding insignificant estimates. The observed heterogeneity of results could potentially be explained by several factors. Unlike other STDs, gonorrhea has a short incubation period making it a better approximation of contemporaneous sexual behavior. For instance, chlamydia is asymptomatic for most people, while gonorrhea symptoms materialize within days of infection.[14] Moreover, using NHSLS data, Cunningham and Shah (2014) show that prostitution is significantly correlated with gonorrhea and not chlamydia for both men and women. To the extent that the reduced form relationship presented in Table 3 works through changes in transactional sex markets, my findings are consistent with the above mentioned paper as well.

Next, I test the relationship between the fracking boom and STD rates using the more detailed data I obtained from the Health Department of Pennsylvania. This gives me an opportunity to explore the heterogeneity of results by demographic group. These estimates are presented in

[13] I note that I tried including county-specific linear time trends in my model and it yielded lower estimates that were mainly insignificant. As discussed by Angrist and Pischke (2008), adding county-specific trends typically only works well in sufficiently big sample periods and with sharp changes in the outcome at the date of treatment.

[14] "Workshop Summary: Scientific Evidence on Condom Effectiveness for Sexually Transmitted Disease Prevention." National Institutes of Health, 2001.

Panel B of Table 3. The results suggest that the impact is larger for females and the younger population. For example, the magnitude of the estimated effect for gonorrhea for females ranges from around 2.9 to 5 new cases per year in the counties that have fracking potential compared to counties that do not. At the same time, the effect is practically non-existent for males.

These results, while surprising, provide evidence in support of the "peer effects" theory. Given that nearly 80% of oil and gas employees are male [15], they are unlikely to be the result of workers with higher incidences of STDs migrating to fracking counties. Instead, it suggests that young women are engaging in riskier behaviors than they would have in the absence of fracking activity. One explanation would be an increase in prostitution, but it could also be due to more non-transactional sex. The fact that STD rates do not change significantly for men could be due to higher propensities to engage in risky behaviors; if men who solicit prostitutes (or who regularly engage in casual sex) are more likely to have STDs, then it is not surprising to see a larger increase in STD rates for women. Finding more detailed demographic data in other states, as well as aggregating prostitution arrest data, would help pin down the underlying mechanism. This would be a promising area for future research.

6.2 Robustness Checks

Tables 4 provides a series of robustness checks that probe several assumptions of the main model and the interpretation of results. Each set of estimates includes year and county fixed effects and some of them also contain a full range of control variables. Panel A considers an alternative year for defining the beginning of the policy. In particular, I re-estimate equation (1) using 2008 and 2006 as the first year of the widespread fracking expansion. The results are similar to the estimates from the main model, and imply that a county on top of oil and gas deposits experiences on average an increase of 4.4-5 new cases of gonorrhea compared to counties without such reserves.

The advantage of defining treatment based on the presence of shales is that it is arguably

[15] Labor Force Statistics from the Current Population Survey: https://www.bls.gov/cps/cpsaat18.htm

exogenous to trends in STDs as the resources the shales contain did not have any economic value before the invention of fracking technology. The disadvantage, however, is that not all the counties with fracking potential actually experienced a drilling boom. Thus, assigning treatment based on the intensity of drilling could yield a more targeted treatment group. On the other hand, one could raise endogeneity concerns if the tendency to allow fracking is correlated with trends in STD rates. To account for these tradeoffs, I use an alternative treatment definition based on drilling activity. In particular, I estimate the following regression:

$$STD_{ct} = \alpha + \beta NewWells_{ct} + \gamma_c + \gamma_t + \epsilon_{ct}, \tag{3}$$

where $NewWells_{ct}$ is the number of new unconventional wells drilled in county c and year t. Panel B of Table 4 reports the results. Both gonorrhea and chlamydia estimates are positive, however only one specification for gonorrhea produces significant effects at the 10% level. The estimate in columns 1 and 3 imply that 10 extra unconventional wells are associated with an increase of 0.510 new cases of gonorrhea and 0.416 new cases of chlamydia infection per 100,000 residents, which approximately correspond to increases relative to the sample averages of 1.2% and 0.3%, respectively. A back-of-the-envelope calculation suggests that the continuous treatment evaluated at the 50-well cutoff used in the literature (for example Komarek, 2016) would lead to an increase of 4.2 gonorrhea cases per 100,000, which is close to my main estimates.

Lastly, I perform a placebo test by looking at a different health outcome that is arguably unaffected by fracking. While rigorous analysis of the changing population composition is outside of the scope of this study, I refer to the abundant anecdotal evidence for guidance. There is no evidence that I know of that links fracking booms to increased obesity rates. I obtained county-level data on the prevalence of obesity in 2004-2013 from the CDC's Division of Diabetes Translation. I re-estimate equation 1 changing the STD outcomes to the obesity rate and present the results in Panel C. As speculated, they are close to zero and insignificant, providing some confidence that the model is not picking up the impact of something else other than shale development on STDs.

6.3 Potential Mechanisms

There are two major mechanisms that could explain the increased STD rates that result from rapid shale development. The first is a change in population composition. If fracking jobs are filled by workers coming from out of state, and if these workers are more likely to have STDs than the existing population, then STD rates in fracking counties will increase even if no individual changes their behavior. In this case the net effect on public health is basically zero since the individuals are simply moving from one area to another, and thus there is no room for policy to improve welfare on the national level.

The second class of mechanisms rely on nonlinear agglomeration effects above and beyond those explained by the composition of the population; in other words, STD rates increase because some individuals are more likely to engage in behaviors such as drinking or soliciting prostitutes as a result of the fracking boom. This mechanism falls more in line with the anecdotal evidence mentioned in the background section and is supported by a substantial body of literature that links peer effects to riskier behavior (see, for example, Lundborg, 2006; Clark and Loheac, 2007) In this case, there is room for policy to improve welfare through improved education and medical infrastructure.

Unfortunately, the limited level of demographic detail available at the county level make distinguishing between these two alternatives difficult. One way to asses these effects is by using information on not just the presence of fracking activity but also its scale. If peer effects are important, then the effects of fracking activity on risky behavior should be increasing at an increasing rate. When I use a specification that includes a quadratic term in the number of wells, I find that the coefficient on this term is close to zero and insignificant for both gonorrhea and chlamydia infections, not providing support for the peer effects story (Table 5).

However, the heterogeneous results for PA described in the earlier section indicate that the peer effects story should not be completely ruled out. I find that the increases are concentrated entirely in females between the ages of 15-24. Given that the vast majority of oil and gas employees are male, these are unlikely to be explained by worker migration. Furthermore, this is the age range that we would expect to be most closely associated with risky sexual behaviors (including prostitution). Prostitution arrest data would be useful but are only available at the police station level, making aggregation difficult; this could be a fruitful avenue for future research.

7 Conclusion

Technological advancements often have outsized effects on the economic development of local economies. Often, these effects are positive, resulting in more jobs and higher wages. They can also be negative, however; rapid inflows of physical and human capital can put a strain on local infrastructure and lead to negative social externalities. Evaluating both positive and negative effects of such development is an important research question that is complicated by the fact that the areas which experience such growth are not chosen at random.

The recent development of hydraulic fracturing and horizontal drilling technologies provides an excellent natural experiment to analyze the effects of such rapid economic development. Because the oil and gas reserves located in shales were inaccessible prior to the development of this technology, communities sitting on top of shale reserves experienced no economic benefit relative to non-shale areas.

This paper contributes to the literature related to resource booms and the development that accompanies them by leveraging this natural experiment to analyze the effects of the fracking boom on sexually transmitted diseases. Using rich data from 2002 through 2015, I perform a difference-in-differences analysis of the Marcellus region, the largest shale formation in the United States. In doing so I exploit the plausibly exogenous geographic distribution of shale deposits and the temporal variation in the expansion of the drilling activity. The findings reveal that rapid shale development is associated with 4-5 additional cases of gonorrhea per 100,000 population, or an increase of 12%. These findings are robust to a large number of specifications and controls.

While this paper does not take an explicit stand on the mechanism through which resource booms lead to higher STD rates, preliminary evidence suggests that it is not simply due to reallocation of people with STDs. Using STD data broken down by gender and age in PA, I find that the effects are almost entirely concentrated in females between the ages of 15 and 24. Given that 80% of workers in the oil and gas industry are male, this effect is unlikely to be explained by workers with STDs moving to fracking communities. Instead it is consistent with the idea that fracking booms lead to a riskier behaviors (such as drinking, casual sex, and prostitution). This idea is supported by both anecdotal evidence and the economics and psychology literatures which document how peer effects can lead an increase in risky behavior.

This work has important policy implications for communities who have access to shale resources. CDC estimates show that there are about 20 million new STD infections in the US each year, which cost the American healthcare system nearly $16 billion in direct healthcare costs alone. Costs may be even higher for fracking communities, which are generally smaller and rural and therefore may lack the healthcare infrastructure to adequately deal with these issues. To the extent that improved education, access to preventative care, and testing services can limit the effects of the spread of STDs, small fracking communities may benefit significantly from implementing such policies.

There are several promising directions for future work in this area. While my main analysis focuses on the Marcellus region due to its large size and more accurate relation to the US economy as a whole, there are also several other large shales to which the methodology could be applied in future versions of this paper (such as those in Texas or North Dakota). In addition, there are several sources of data which could help pin down the mechansim underlying these results. Prostitution arrests, which are only available at the police station level but could plausibly be aggregated to the county level, would be a useful way to check whether communities are engaging in riskier behaviors as a result of fracking booms.

8 References

Angrist, Joshua D., and Jorn-Steffen Pischke. 2008. "Mostly harmless econometrics: An empiricist's companion." Princeton university press.

Allcott, Hunt, and Daniel Keniston. 2014. "Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America." National Bureau of Economic Research Working Paper 20508.

Bartik, Alexander W., Janet Currie, Michael Greenstone, and Christopher R. Knittel. 2017. "The Local Economic and Welfare Consequences of Hydraulic Fracturing." National Bureau of Economic Research Working Paper 23060.

Black, Dan, Terra McKinnish, and Seth Sanders. 2005. "The Economic Impact of the Coal Boom and Bust." *Economic Journal* 115 (503): 449-76.

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.

Brundage, Tracy L., Timothy W. Kelsey, Janice Lobdell, Larry L. Michael, Jerey Jacquet, James R. Ladlee, Jerey F. Lorson, and Thomas B. Murphy. 2011. "Pennsylvania Statewide Marcellus Shale Workforce Needs Assessment." Marcellus Shale Education and Training Center.

Cascio, Elizabeth U., and Ayushi Narayan. 2015. "Who Needs a Fracking Education? The Educational Response to Low-Skill Biased Technological Change." National Bureau of Economic Research Working Paper 21359.

Clark, Andrew E., and Youenn Lohac. 2007 ""It wasn?t me, it was them!" Social influence in risky behavior by adolescents." *Journal of health economics* 26 (4): 763-784.

Cunningham, Scott, and Manisha Shah. 2014. "Decriminalizing Indoor Prostitution: Implications for Sexual Violence and Public Health." National Bureau of Economic Research

Working Paper 20281.

Edlund, Lena, and Evelyn Korn. 2002. "A theory of prostitution." *Journal of political Economy* 110 (1): 181-214.

Eligon, John. 2013. "An Oil Boom takes a toll on health care." New York Times. Available at: http://www.nytimes.com/2013/01/28/us/boom-in-north-dakota-weighs-heavily-on-health-care

Gopalakrishnan, Sathya, and Allen H. 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.

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.

Haggerty, Julia, Patricia H. Gudeb, Mark Delorey, and Ray Rasker. 2014. "Long-Term Effects of Income Specialization in Oil and Gas Extraction: The U.S. West, 1980?2011." *Energy Economics* 45: 186-195.

Hausman, Catherine, and Ryan Kellogg. 2015. "Welfare and distributional implications of shale gas." National Bureau of Economic Research Working Paper 21115.

Hill, Elaine L. 2013. "Shale Gas Development and Infant Health: Evidence from Pennsylvania." Working paper, Cornell University.

Howarth, Robert W., Anthony Ingraffea, and Terry Engelder. 2011. "Natural Gas: Should Fracking Stop?" *Nature* 477 (7364): 271-75.

Jacobsen, Grant D. and Dominic Parker. 2014. "The Economics Aftermath of Resource Booms: Evidence from Boomtowns in the American West." *Economic Journal* 126: 1092-1128.

Jacquet, Jeffrey. 2009. "Energy boomtowns and natural gas: Implications for Marcellus Shale local governments and rural communities." *NERCRD Rural Development* 30: 59-78.

James, Alexander, and Brock Smith. 2017. "There Will Be Blood: Crime Rates in the Shale-Rich US Counties." *Journal of Environmental Economics and Management* 84: 125-152.

Komarek, Timothy M. 2016. "Labor market dynamics and the unconventional natural gas boom: Evidence from the Marcellus region." *Resource and Energy Economics* 45: 1-17.

Lundborg, Petter.2006. "Having the wrong friends? Peer effects in adolescent substance use." *Journal of health economics* 25 (2): 214-233.

Maniloff, Peter and Ralph Mastromonaco. 2014. "The Local Economic Impacts of Hydraulic Fracturing and Determinants of Dutch Disease." Working paper, Colorado School of Mines.

Michaels, Guy. 2011. "The long term consequences of resource?based specialisation." *The Economic Journal* 121 (551): 31-57.

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

Putz, Audrey, Alex Finken, and Gary A. Goreham. 2011. "Sustainability in Natural Resource-Dependent Regions that Experienced Boom-Bust-Recovery Cycles: Lessons Learned from a Review of the Literature." North Dakota State University, Fargo.

Silverstein, Jason. 2013. "How Fracking Is Bad for Our Bodies." Atlantic. Available at: https://www.theatlantic.com/health/archive/2013/10/how-fracking-is-bad-for-our-bodies/280384/

Vaughan, Vicky. 2012. "Boom takes toll on health care, social services." My San Antonio. Available at: http://www.mysanantonio.com/news/local˙news/ article/Boom-takes-toll-on-health-care-social-services-3717774

Weber, Jeremy G. 2012. "The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming." *Energy Economics* 34 (5): 1580-1588.

Wolfers, Justin. 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *The American Economic Review* 96 (5): 1802-1820.

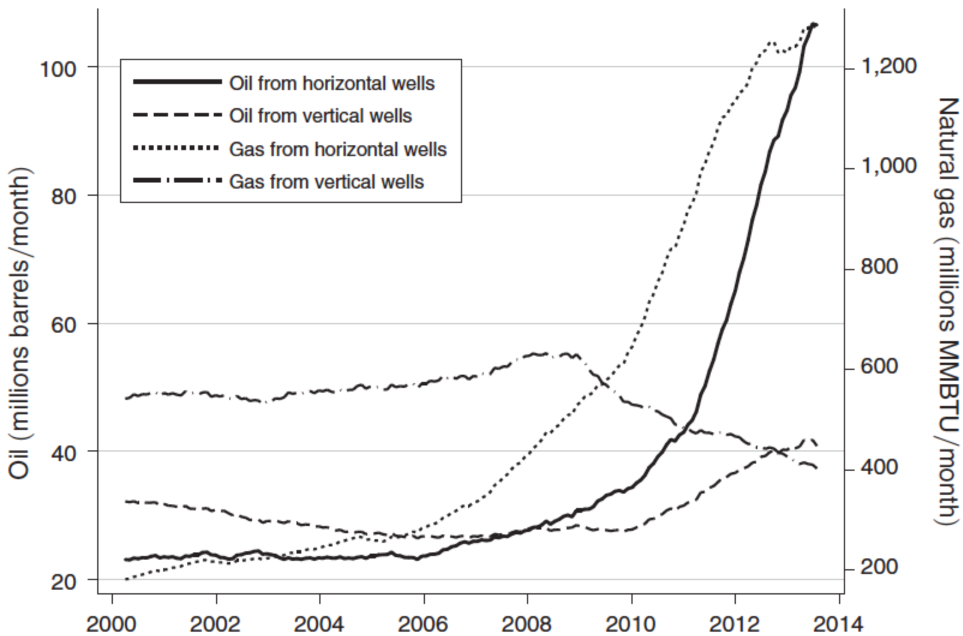

Figure 1: Annual Production by Drill Type
Source: Feyrer et al. (2017)

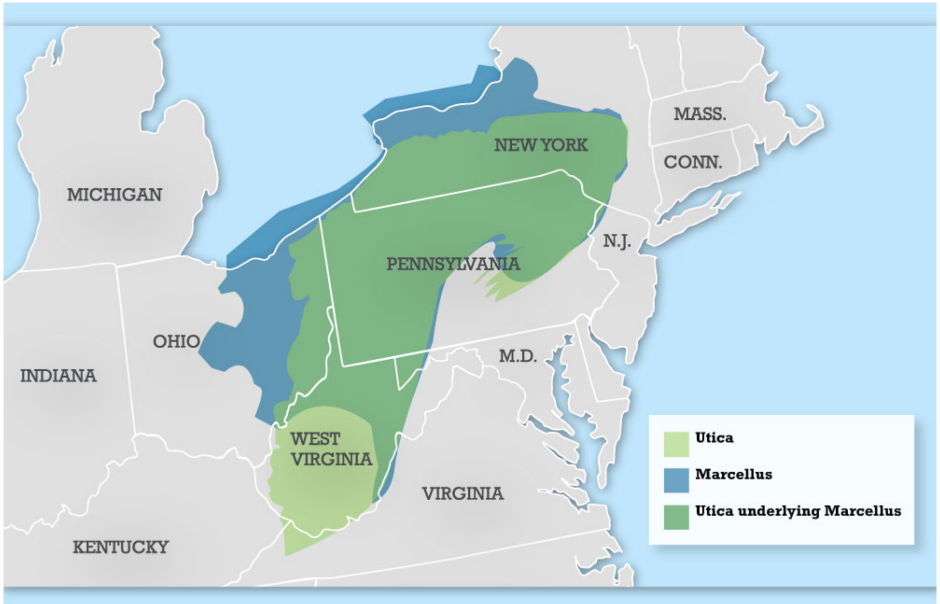

Figure 2: Shale formations in the Northeastern United States
Source: Energy Information Administration

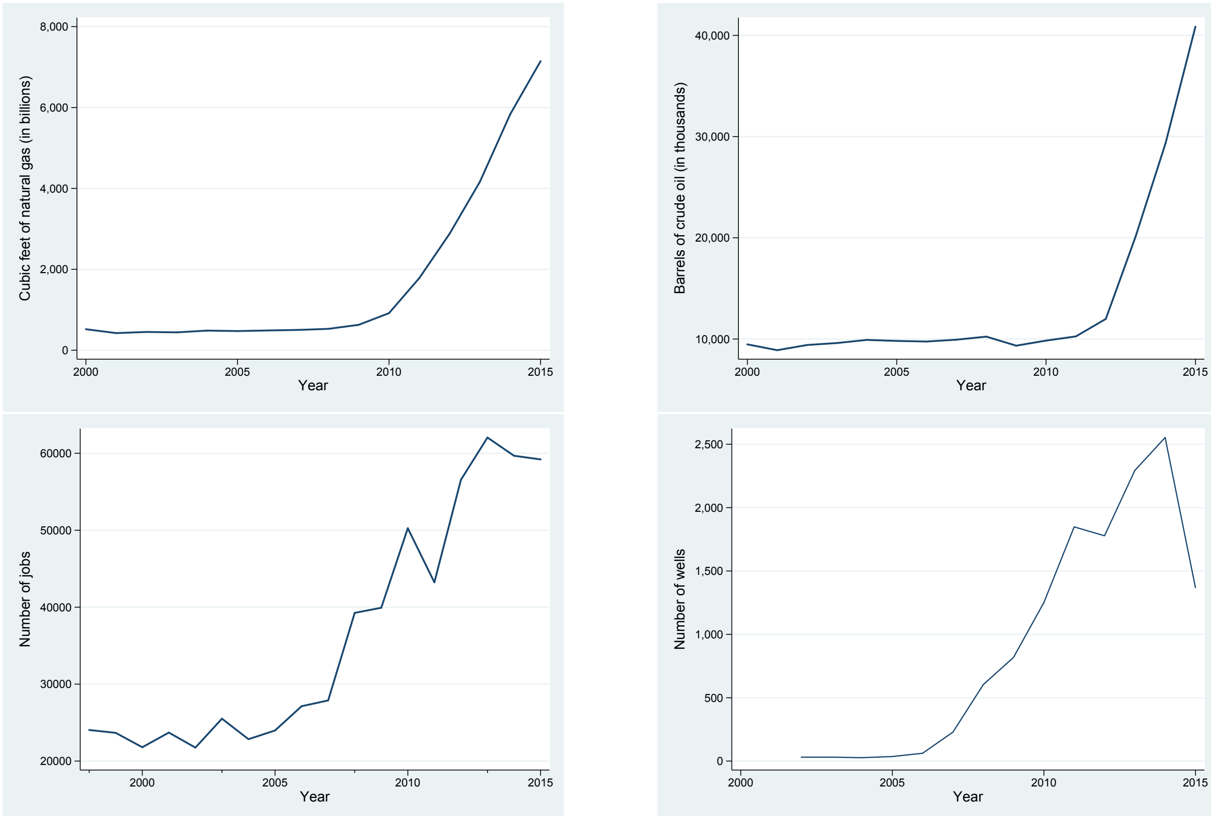

Figure 3: Natural gas (upper left) and oil (upper right) production, oil and gas employment (bottom left), and number of new unconventional wells (bottom right)
Source: Enegry Information Administration, Bureau of Economic Analysis and author's calculations of the state totals for Ohio, Pennsylvania, and West Virginia from DrillingInfo database

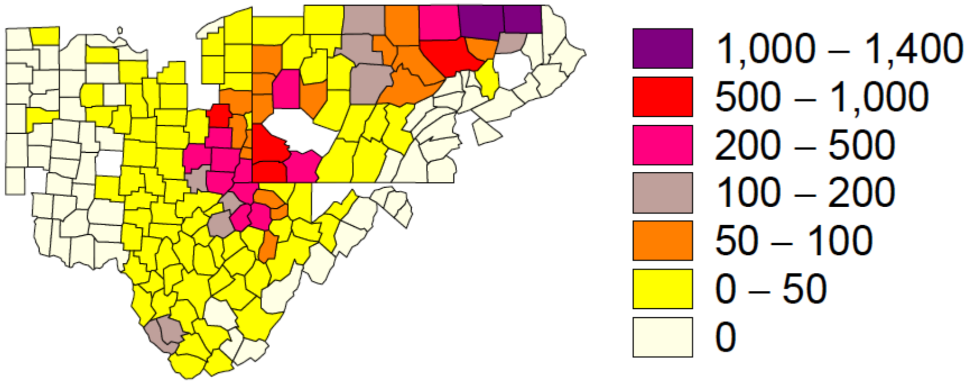

Figure 4: Total new unconventional wells spudded per county, 2002-2015
Source: Author's calculations from DrillingInfo well-level database

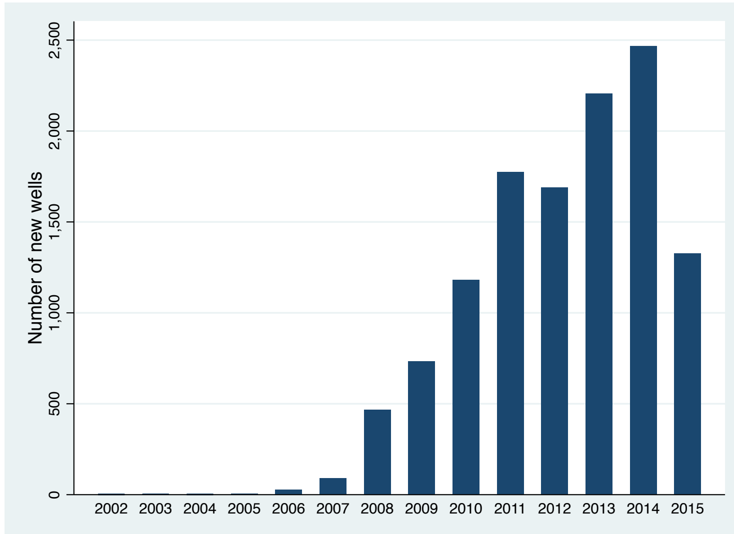

Figure 5: Total new unconventional wells spudded per year, 2002-2015
Source: Author's calculations from DrillingInfo well-level database

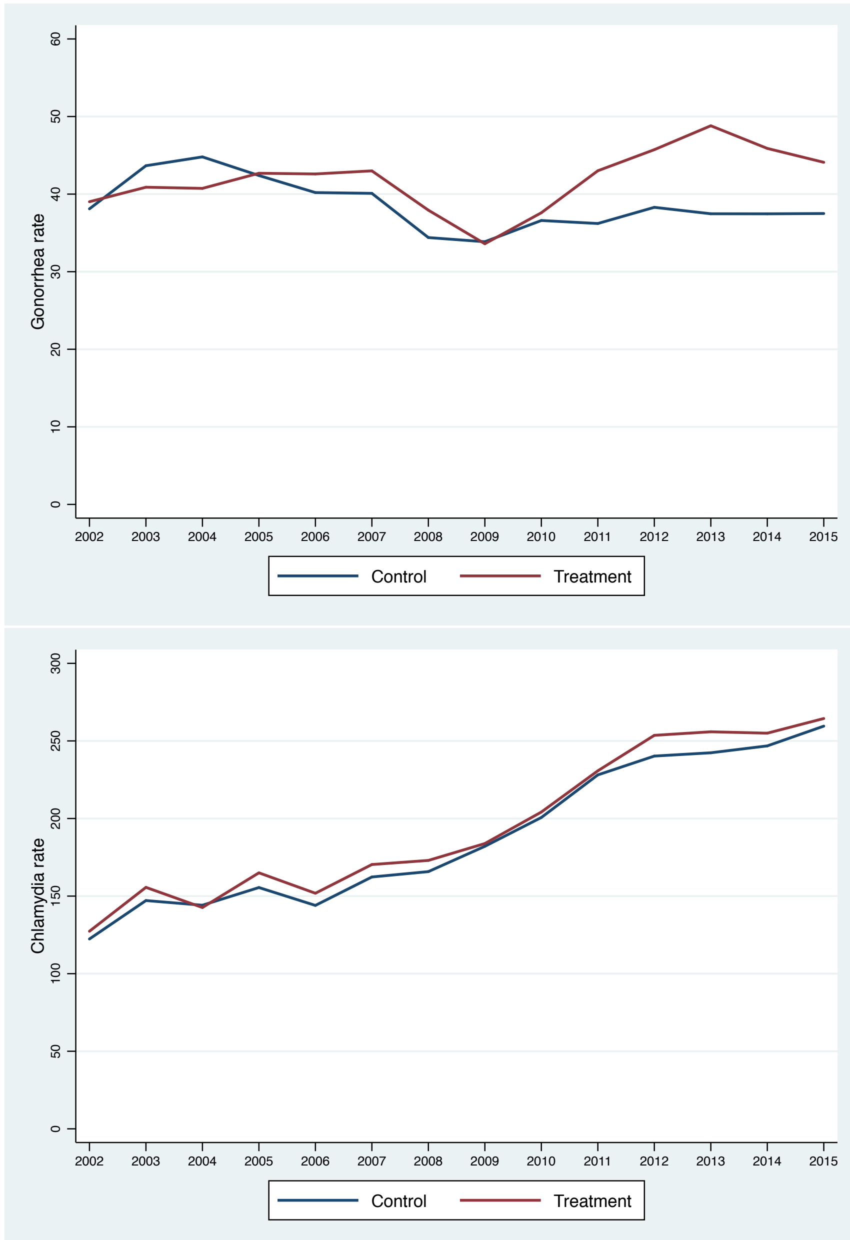

Figure 6: Trends in gonorrhea (top) and chlamydia (bottom) rates by group

Source: Author's calculations from CDC's county-level STD rates for Ohio, Pennsylvania, and West Virginia

Table 1: Summary Statistics

Variable	All counties	Before 2007 Control	Before 2007 Treatment	Before 2007 Difference	After 2007 Control	After 2007 Treatment
	(1)	(2)	(3)	(4)	(5)	(6)
Gonorrhea rate	40.50	42.11	40.69	1.42	37.55	41.95
	(40.05)	(3.91)	(3.36)	(3.13)	(33.60)	(38.68)
Chlamydia rate	193.286	143.85	147.04	-3.19	215.55	221.07
	(97.60)	(70.91)	(79.91)	(5.35)	(89.26)	(102.32)
Number of wells	4.74	0.00	0.07	-0.07**	0.02	11.09
	(23.25)	(0.07)	(0.54)	(0.03)	(0.18)	(34.72)
Total Employment	30,913.74	30,943.13	30,290.20	652.93	31,614.45	30,889.88
	(29,495.24)	(27,467.94)	(29,220.86)	(1,986.88)	(29,159.55)	(30,397.94)
Income per capita	30,856.56	27,856.28	25,623.59	2,232.69***	34,664.15	32,658.29
	(6,728.75)	(3,692.30)	(6,194.33)	(379.74)	(5,106.71)	(6,296.16)
Population	61,848.88	60,981.68	61,326.47	-344.79	63,260.03	61,659.18
	(52,175.54)	(48,567.73)	(52,357.97)	(3,545.46)	(52,099.06)	(53,158.84)
% Male	49.73	49.63	49.59	0.05	49.67	49.87
	(1.87)	(1.61)	(1.81)	(0.12)	(1.34)	(2.17)
% White	95.54	95.78	96.00	-0.23	95.05	95.47
	(3.25)	(3.24)	(2.86)	(0.21)	(3.61)	(3.22)
% Black	2.47	2.41	2.32	0.09	2.57	2.52
	(2.67)	(2.71)	(2.43)	(0.18)	(2.84)	(2.70)
% Age 15-34	24.40	24.98	25.03	-0.04	24.05	24.06
	(3.56)	(2.40)	(3.82)	(0.24)	(2.44)	(4.07)
% Age 35-60	35.13	35.82	35.77	0.04	34.80	34.74
	(2.03)	(1.49)	(2.04)	(0.13)	(1.70)	(2.16)
% Less than high school	17.65	19.85	22.64	-2.79***	12.27	14.28
	(7.19)	(4.94)	(7.26)	(1.02)	(3.61)	(5.61)
% High school	43.76	44.68	43.52	1.15	43.44	43.68
	(5.55)	(5.15)	(5.41)	(0.83)	(6.11)	(5.61)
% Some college	23.10	21.54	20.45	1.11*	26.27	25.01
	(4.49)	(4.10)	(3.69)	(0.62)	(4.33)	(3.53)
% College	15.50	13.93	13.40	0.53	18.00	17.11
	(6.53)	(5.11)	(5.81)	(0.87)	(6.27)	(7.16)
F-statistic				0.60		
p-value				0.8122		
Observations	2,590	315	610	-	567	1,098

Notes: Column 1 reports means and standard deviations (in parentheses) for the entire sample. Columns 2-3 and 5-6 present means and standard deviations (in parentheses) of characteristics of counties in treatment and control groups. Column 4 presents the differences between the control and treatment group in the pre-fracking period. Robust standard errors of these differences are reported in parentheses. Number of county-year observations in each group is reported in the bottom row. Educational attainment variables are only reported for 2000 and 2011-2015 (from Census and pooled ACS respectively).
* p<0.10, ** p<0.05, *** p<0.01

Table 2: Treatment-Control Differences in Pre-Fracking Time Trends in STD Outcomes

	Gonorrhea rate	Chlamydia rate
	(1)	(2)
Treat	1.0301	2.9584
	(10.2511)	(12.8336)
Treat × 2003	-3.9082	5.1514
	(3.2034)	(9.2943)
Treat × 2004	-5.4091	-5.7332
	(4.3273)	(9.0423)
Treat × 2005	-0.7853	6.6171
	(3.0204)	(9.0082)
Treat × 2006	1.1271	4.2059
	(3.143)	(10.0276)
F-statistic	0.21	0.25
p-value	0.9318	0.9091
Observations	925	925

Notes: The coefficients on the interaction of the treatment variable and year dummies for pre-fracking years are reported from equation (2). Robust standard errors clustered at the county level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 3: Main Results

	Gonorrhea		Chlamydia	
	(1)	(2)	(3)	(4)
Panel A: Main sample				
Total	5.0476**	4.7639*	0.7129	1.4787
	(2.4965)	(2.5372)	(6.9792)	(7.1639)
Panel B: Heterogeneous results for PA				
Total	5.7073**	4.2135*	10.2090	9.5547
	(2.5359)	(2.4857)	(8.9243)	(9.9243)
Female	4.8063*	2.9005	9.8704	10.0183
	(2.6699)	(2.1767)	(10.7660)	(11.8996)
Male	-0.7322	0.4647	3.0289	2.6227
	(2.2102)	(2.2754)	(9.5974)	(10.5483)
Age 15-24	10.1586	10.9565	35.4264	32.6111
	(11.1383)	(12.7606)	(45.3871)	(44.8552)
Age 25-34	0.8902	2.1334	5.3201	6.0484
	(7.9042)	(7.3216)	(8.0659)	(8.0551)
Age 35+	0.4006	-0.7986	-0.3634	0.8093
	(0.5146)	(0.5713)	(1.3845)	(1.2546)
Controls	No	Yes	No	Yes
County and Year FE	Yes	Yes	Yes	Yes

Notes: The coefficients on the interaction of the treatment and post variables are reported from equation (1). Control variables include population shares split by age, race, gender, total employment, income per capita, and population. Panel A includes 2,590 county-year observations from OH, WV, and PA. Sample used in panel B has 728 county-years. Robust standard errors clustered at the county level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 4: Robustness Checks

Dependent variable	Gonorrhea		Chlamydia		Obesity	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: changing the timing of treatment						
$\text{Post}_{2007} \times$ Treat	5.0476**	4.7639*	0.7129	1.4787		
	(2.4965)	(2.5372)	(6.9792)	(7.1639)		
$\text{Post}_{2008} \times$ Treat	5.0939**	4.8148*	0.9971	1.08407		
	(2.6076)	(2.6818)	(7.4125)	(7.5593)		
$\text{Post}_{2006} \times$ Treat	4.7367*	4.4533*	2.5609	3.9489		
	(2.6479)	(2.6990)	(6.42038)	(6.7185)		
Panel B: continuous treatment						
New wells	0.0510*	0.0478	0.0416	0.0392		
	(0.0294)	(0.0318)	(0.0707)	(0.0753)		
Panel C: placebo test						
$\text{Post}_{2007} \times$ Treat					-0.1327	-0.2154
					(0.2780)	(0.2622)
Controls	No	Yes	No	Yes	No	Yes
County and Year FE	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:*Control variables include population shares split by age, race, gender, total employment, income per capita, and population. Robust standard errors clustered at the county level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 5: Exploring the Mechanisms

	Gonorrhea rate	Chlamydia rate
	(1)	(2)
Panel A: Main sample		
Wells	0.1116***	0.1043
	(0.0404)	(0.1455)
Wells Squared	-0.0003	-0.0003
	(0.0002)	(0.0005)
Panel B: PA only		
Wells	0.0644	0.1465
	(0.0500)	(0.1791)
Wells Squared	-0.0001	0.0004
	(0.0002)	(0.0005)

Notes: Presented results come from the specification of STD rates on the continuous measure of new wells and its quadratic term. Robust standard errors clustered at the county level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01

A Fracking Boom in the Unites States

Shale oil and gas refer to fossil fuels trapped in shale formations. Shale formations are fine-grained sedimentary rocks that can be easily split into thin sheets. Some shales contain kerogen, a type of organic matter that yields oil and gas. Even though the presence of some shale formations was known in the past, they had no economic value attached to them, and the counties located over shales developed largely independently of the geology underneath. The "fracking boom" mainly came from improvement of extraction technologies which made it feasible and profitable to unlock large quantities of shale oil and gas resources. These two technologies are hydraulic fracturing and horizontal drilling. Hydraulic fracturing is a well stimulation technique which injects high pressure materials (water, sand and chemicals) into the targeted resource and creates fractures through which oil and gas can flow to the well. Horizontal drilling allows for the construction of wells that are initially vertical but turn to horizontal at depth. A single horizontally drilled and fractured well can be used to extract large quantities of resources as it can run for long distances underground parallel to the layers of shale.

While shale resources and production are found in several US regions, the seven most prolific areas shown in Figure 7 accounted for 92% of domestic oil production growth and all domestic natural gas production growth during 2011-14.[16] Some of the shales are located in areas which also have a history of conventional oil and gas production (for example, the Barnett shale in Texas), but others are in places without pre-existing large-scale oil and gas industry (for example, the Marcellus in Pennsylvania and the Bakken in North Dakota).

[16] April 2017 Drilling Productivity Report, EIA: https://www.eia.gov/petroleum/drilling

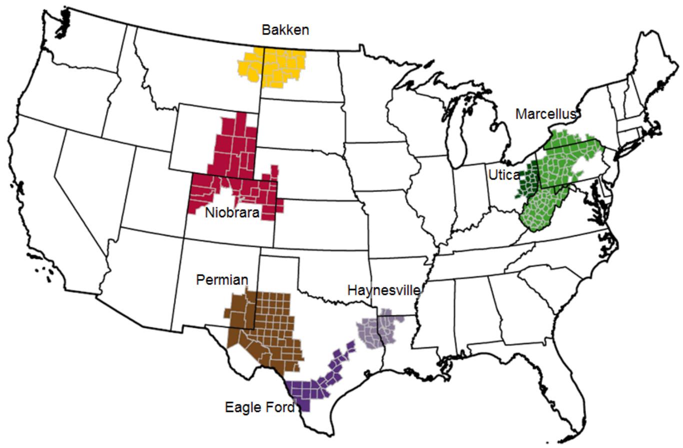

Figure 7: Distribution of Shale Resources
Source: Energy Information Administration (EIA)