Unearthed: An Exploration of Shale Development on House and Income Inequality

Shahrzad Ghourchian*

June 9, 2021

Abstract

Recent advances in shale development have produced both positive and negative outcomes for local communities, with higher employment and income known to be the most significant effects. Despite the stated importance of the distribution of economic gains among local populations in previous literature, adequate research on the shale boom's impact on inequality and affordability does not exist. I employ the difference-in-difference (DiD) method to study the unintended social consequences of the hydraulic fracturing boom in Oklahoma, the second-largest producer of oil and gas in the country, over the period of 2004–2017. Given the magnitude of the economic adjustment with the advent of hydraulic fracking in Oklahoma, it is a prime setting to study the impact of shale development on inequality and affordability. I find that shale counties experienced higher housing prices and lower affordability power compared to non-shale counties. However, the estimation fails to find any statistically significant effect on inequality. Oklahoma provides a unique setting because of the size of the fracking industry and the speed with which fracking grew. Nevertheless, the setting is not too unique: The results will be relevant for any locality facing a natural resource boom.

Keywords: Shale oil and gas, housing prices, hydraulic fracturing, inequality, affordability

JEL Codes: Q0, R0, I3

*Department of Economics, Siena College, Loudonville, NY 12211, USA; Tel: +1-305-987-1975; E-mail: sghourchian@siena.edu. The author thanks participants at the Eastern Economic Association 45th Annual Conference, the Georgetown Center for Economic Research 2019 Biennial Conference, the Southern Economic Association 90th Annual Meeting, and the 46th Eastern Economic Association conference for their helpful comments and suggestions. The usual disclaimer applies.

1 Introduction

Recent technological advancements in the extraction of shale oil and gas have made previously inaccessible resources economically profitable. Aside from macroeconomic benefits, previous studies show that energy developments create both benefits and costs for local populations. While there is a general consensus in the literature that energy-driven labor demand shocks have significant effects on wages and employment (e.g. Pless, 2010 and Marchand, 2015) and despite the stated importance of the distribution of economic gains among local populations in previous literature (e.g., Weber, 2012; Christopherson and Rightor, 2012; Marchand, 2015), adequate research on the shale booms impact on inequality and affordability does not exist. I study the unintended social consequences of the hydraulic fracturing boom in Oklahoma, the second-largest producer of oil and gas in the country[1]. Shale development in this state provides a unique opportunity to examine the effect of natural resource booms on associated regional economic and social impacts.

In 2003, Devon Energy, one of the biggest Oklahoma-based oil companies, combined its expertise in horizontal drilling with Mitchell Energy's knowledge of hydraulic fracturing (or "fracking"). Previously inaccessible natural gas and oil suddenly became available and economically profitable. Higher prices and the combination of horizontal drilling with fracking changed the economics of oil and gas production (Brown, 2014). As Pless (2010) argues, job creation is likely the most significant effect of shale development, since workers are needed for legal and regulatory work, construction, and infrastructure development, resulting in higher local employment and income not only for those who work in the industry but also for others through the spillover effect. The result, depending on population and housing density in the area, is either an increase in demand for housing and a subsequent increase in the value of local properties (Bennett and Loomis, 2014) or a housing shortage, at least in the short run, that can lead to higher rents or affordability crises for low-income households, driving low-income renters to leave the area (Christopherson and Rightor, 2012; Schafft et al., 2014). Moreover, within the local labor market, the gains generated from the positive labor demand shock induced by the energy boom may be distributed either uniformly or unevenly across the earnings distribution, with overall inequality either rising or falling (Marchand, 2015).

In Oklahoma, the industry employed 53,500 people in 2016, who earned a combined total of $5.6 billion, and the spillover has impacted nearly every sector of the state economy. Each new direct oil and gas job supports slightly more than two additional jobs statewide, and almost one in five wage and salary workers and self-employed workers are employed directly or indirectly by the oil and gas sector. Including spillover effects, approximately $32.6 billion (27%) of total state household earnings come from the energy sector. Oil- and gas-driven income gains have pushed state per capita income to 95% of the United States average in

[1] According to the Research Foundation Report of the State Chamber in 2016.

recent years, up from 85% just a decade ago. Moreover, in 2015, Oklahomans earned an estimated $1.7 billion in oil and gas royalties[2]. Given the magnitude of the economic adjustment with the advent of hydraulic fracking in Oklahoma, it is a prime setting to study the impact of shale development on inequality and affordability.

This article studies the distributional and affordability effects of the shale boom on local communities using the difference-in-difference (DiD) method. I first uses county-level data between 2004 and 2017 to show the local economic effect of shale gas development on housing prices. Next, using the annual data from the Internal Revenue Service (IRS), I calculate the Gini coefficient and housing affordability index for each county to estimate the shale development's impact on inequality and affordability. The results show that households living in shale counties experienced higher housing prices and lower affordability power compared to non-shale counties. However, the estimation fails to find a statistically significant effect on inequality.

This research is economically important for many reasons since it presents one of the first empirical studies to investigate the effect of the energy industry on local communities, particularly on income inequality and affordability rather than the poverty rate, which has been commonly studied in previous literature. Note that it is important to distinguish between poverty and inequality. As is explained in the *World Bank* website:

> *Theoretically, it is possible to have no poverty but high inequality (high concentration of income among a small group of very rich households). Inequality is a broader concept than poverty. It is defined in terms of an entire population, not just the portion below a certain poverty threshold.*

The distribution of economic gains from oil and gas extraction is a major concern, especially with regard to those who are already economically disadvantaged. As shale development continues to expand, understanding its potential impact on surrounding populations will be a key component of effective policy formation (Gopalakrishnan and Klaiber, 2014); estimating the economic gains for locals can help policymakers decide how much incentive (or disincentive) to provide for extraction (Weber, 2012) while maintaining a competitive advantage over other shale players. Moreover, recent fluctuations in oil and gas prices due to the COVID-19 pandemic raised a lot of concerns regarding the industry's impact on locals, as drastic movements in the prices of energy-related goods can generate labor demand shocks in local labor markets (Marchand, 2012). Last but not least, there is empirical evidence in favor of the idea that economic inequality generates violent criminal activity; an individual may be more likely to commit a violent crime when they feel economically or socially alienated from a majority group (James and Smith, 2017, and Merton, 1938).

This paper proceeds as follows: Section 2 provides a background, a description of the economic impacts of the shale development boom, and a literature review on the subject. Section 3 describes the data and

[2]These data were gathered from the State Chamber of Oklahoma's Research Foundation Report, September 2016.

methodology used in this study, and Section 4 provides the results of my empirical work. Finally, Section 5 contains concluding remarks and the policy implications of this paper.

2 Shale Boom Background and Literature Review

2.1 Background

Exploration for and increases in natural gas and oil production can affect local communities in many ways, both directly, through amenities or disamenities associated with extractions, and indirectly, through positive and negative spillovers. Policymakers are especially concerned about these effects, since a vast number of Americans live close to oil and gas wells, promoting a growing body of research on this topic. Changes in housing prices as a result of unconventional oil and gas development commonly appear in literature as indicators of communities' perceptions of the benefits and damages of such development, as they aggregate and monetize the preferences of home buyers and sellers (Krupnick and Echarte, 2017). Moreover, increases in income for locals, which can come from lease and royalty payments from gas and oil companies to land and mineral rights owners, might increase housing demand and, therefore, housing prices. Migrating oil and gas industry workers can also create positive housing demand shocks (Farren, 2014). Oil and gas production contributes a significant amount of tax to local governments; these revenues can then be capitalized in housing prices through improvements in public goods and services, such as schools, roads, and parks (Boslett et al., 2016).

It is also possible for nearby housing to capitalize due to disamenities associated with extraction. Environmental degradation is one of the main concerns about hydraulically fractured wells for various reasons. Fracking consists of shooting a mixture of water, chemicals, and sand into wells to create fissures in rock formations that free the trapped gas. Therefore, shale extraction is highly water-intensive, and due to its chemical content, the wastewater needs to be disposed of properly. Consequently, shale development brings vast water management concerns, especially how to avoid potentially contaminating groundwater nearby groundwater-dependent homes (Muehlenbachs et al., 2015) and earthquakes induced by wastewater injection (Cheung et al. 2018). Moreover, a housing shortage, at least in the short run, can cause higher housing rents, which may lead low-income households to be unable to afford their housing, causing affordability crises. Other shale development disamenities that may push local housing prices down include noise pollution, deterioration of local roads and bridges, a decrease in air quality, an increase in traffic and nighttime light, an increase in the crime rate, etc. These types of possible negative impacts of shale development on local welfare affects local governments, which may find them difficult to address.

The main impact of shale development upon the local economy, however, is known to be the increase in economic activity and, hence, an increase in labor demand. Natural gas extraction is highly labor-intensive, and as Pless (2010) argues, job creation is likely the most significant effect since workers are needed for legal and regulatory work, construction, and infrastructure development, causing higher employment and income not only for those who work in the industry but also for others through the spillover effect[3]. Inequality could, therefore, be the result of the income gap between landowners who receive royalties and those who do not (Brasier et al., 2011) or the income gap between workers in specialized, highly skilled occupations and those involved in lower-paid, unskilled occupations (Lawrie et al., 2011)[4]. To my knowledge, this is the first study to examine the impact of shale development on housing affordability and inequality, despite many studies providing evidence of the shale boom's impact on income and housing prices.

2.2 Literature Review

Given the importance of oil and gas production to local communities, a growing body of literature using surveys, interviews, and various data sets and empirical methods can be found on the overall impact of energy industry development on local households[5], on employment and income[6], and on housing prices[7]. However, although increases in economic activity and hence employment and income are positioned to be the most important channel through which oil and gas extraction impacts local communities (e.g., Pless, 2010; Weber, 2012), only a few studies address the distribution of money within these communities, i.e. inequality, or who benefits from development activities in the United States. The results of these studies are mixed.

For instance, Black et al. (2005) analyze the effect of the coal boom and bust on poverty and find strong evidence that poverty decreased during the coal boom and increased during the coal bust using data for Kentucky, Ohio, Pennsylvania, and West Virginia. They argue that for areas that are economically depressed, as the coal towns of this four-state region were in 1970, attracting industrial employment can help increase local wages and reduce poverty without crowding out existing industrial employment. Michaels (2010) finds no evidence of any adverse effect of specialization in oil on income inequality, despite it having increased per capita income. Schafft et al. (2014), however, argue that respondents in areas with high levels of drilling are significantly more likely to perceive the effects of local economic gains but also report increased

[3] Which also causes an increase in demand for housing and thereby increases the value of local properties (Bennett and Loomis, 2014).

[4] As Marchand (2015) explains: "within a local labour market, the gains generated from the positive labour demand shock induced by an energy boom may be distributed either uniformly or unevenly across the earnings distribution. In the case where all individuals proportionately benefit from the gains of an energy boom, overall inequality is expected to remain unchanged. If, however, these benefits accrue only to certain individuals at either the top or the bottom of the distribution, overall inequality would either rise or fall respectively".

[5] e.g. Considine et al., 2009, Anderson and Theodori, 2009, Brasier et al., 2011, Litovitz et al., 2016, Schafft et al., 2014, Jacobsen, 2019, Ratledge and Zachary, 2017.

[6] e.g. Michaels, 2011, Lawrie et al., 2011, Christopherson and Rightor, 2012, Weber, 2012, Brown, 2014, Tunstall, 2015.

[7] e.g. Boxall et al., 2005, Sanders, 2012, Muehlenbachs et al., 2013, Gopalakrishnan and Klaiber, 2014, Bennett and Loomis, 2014, Muehlenbachs et al., 2015, Balthrop and Hawley, 2015, Boslett et al., 2016, Weber and Hitaj, 2015, Delgado et al., 2016, Weber et al., 2016, Cheung et al., 2018.

inequality, heightened vulnerability of disadvantaged community members, and pronounced strains on local infrastructure. Jacobsen (2019) presents evidence of the distributional effects of energy extraction, James and Smith (2017) document a rise in income inequality that coincides with the timing of the energy boom, and Hwang and Paarlberg (2019) argue that the energy boom increases levels of income inequality.

For Canada, Marchand (2015) finds that the energy boom aggregated local inequality but reduced local poverty, while Fortin and Lemieux (2015) find that less-educated workers experienced a larger growth in wages than university graduates, which reduced returns to education and overall inequality. They argue that the extractive resources sector boom helps account for about two thirds of the divergence in the growth in mean wages between these provinces and the rest of the country and that all wage percentiles increase more for women than men in both periods, resulting in a small decline in the gender gap.

For countries other than the United States and Canada, López-Feldman et al. (2007) examine distributional and poverty effects of natural resource extraction at the national, regional, and community levels in Mexico. Their results show that inequality in the distribution of natural resource income is relatively high, and welfare transfers are usually unequally distributed (most households do not receive them), but they are directed disproportionately to poor households. Buccellato and Mickiewicz (2009) find that regional oil and gas abundance is associated with high within-region inequality in Russia, and Loayza et al. (2013) find evidence that producing districts in Peru have better average living standards than otherwise similar districts (larger household consumption, lower poverty rate, and higher literacy). Moreover, they find that income inequality increases in both producing and non-producing districts. For Australia, Lawrie et al. (2011) examined a range of socioeconomic indicators of resource booms, such as income, cost of living, housing affordability, welfare receipts, and unemployment, and find that growing levels of inequality suggest that those on lower incomes may experience a significant disadvantage. Bhattacharyya and Williamson (2013) investigate the effects of resource booms on income distribution in Australia over the century from 1921 to 2004 and find that the very top end of the income distribution benefits from commodity booms disproportionately more than the rest of society. Moreover, they argue that non-renewable resources, such as minerals and petroleum, increase inequality. Caselli and Michaels (2013) argue that the development of the oil sector increased per capita income in oil abundant counties in Brazil without increasing local income inequality.

3 Data and Methodology

3.1 Data

For housing values, I use the county-level annual housing price index, obtained from the Federal Housing Finance Agency (FHFA), as my measure of changes in housing values. The FHFA House Price Index (HPI) is a broad measure of the movement of single-family house prices. Covering 49 counties, the HPI is a weighted, repeat-sales index, meaning that it measures average price changes in repeat sales or refinancing on the same properties. Since the HPI is not adjusted for inflation, the housing price index was then converted to real values using the Consumer Price Index (CPI) excluding shelter. To address concerns regarding the redistribution of economic gains for local communities, I use Internal Revenue Service (IRS) data to measure the impact on income inequality using the Gini index. The publicly available IRS data is in group form. Therefore, I rely upon the studies of Cowell and Mehta (1982), Cowell (2000), and Frank (2009) to construct a compromise Gini coefficient[8]. I then use the median household income from the IRS data and the real housing price index for the frequently used housing affordability index, the income-to-housing price ratio. For counties' characteristics, such as income per capita, population density, and housing density, I use the Census 2010 data.

3.2 Empirical Estimation

There is a growing body of research using natural experiments to study the impact of resource booms. The difference-in-difference (DiD) method, for instance, has become an increasingly popular way to estimate causal relationships between shale booms and local economies. DiD estimation consists of identifying a specific intervention or treatment and then comparing the difference in outcomes after and before the intervention for groups affected by it to the same difference for unaffected groups (Bertrand et al., 2002). In this study, I define the treatment group to be counties that experienced shale extraction (which I call shale counties) and the treatment to be the shale development boom that happened in 2008[9]. I use a linear DiD regression to compare housing prices, inequality, and affordability in shale counties in Oklahoma before and after the boom to non-shale counties over the period of 2004 to 2017. To avoid endogeneity[10], shale counties were selected based on their geographic location, as it is common in the literature (e.g. Michaels, 2010). A problem that arises with using other counties in the same state as the control group is that there may be geographic spillovers associated with energy extraction such that what are considered control groups may, in fact, be receiving a treatment effect, which would understate the treatment effect (Munasib et al., 2015). To avoid this problem, I follow the literature (e.g., Black et al., 2005; Weber, 2011; Munasib et al., 2015)

[8] Appendix A shows how compromise Gini coefficient was calculated.

[9] Majority of papers that studied the impact pf shale boom used the year 2008. Weber (2012), however, uses 2007 instead of 2008 because of the macroeconomic shock to the U.S. that occurred in 2008. Since in the state of Oklahoma housing prices did no experience as much fluctuations as other states in the United States, I believe it is safe to follow the literature and use 2008 as the year shale boom happened. Graphing the percentage change in the number of horizontal wells also shows that 2008 was the year that shale boom happened in Oklahoma. However, I use falsification test as robustness check with 2007 and 2009 as the year for when shale boom, the treatment, happened.

[10] Weber (2012) explains: "county residents may fight gas drilling to avoid a possible decrease in property values or quality of life, with wealthier counties perhaps the best equipped to win the fight. Similarly, gas companies may target gas formations in counties based on characteristics not observed by the econometrician and that affect the outcome of interest, thereby introducing potential bias."

and construct and report results for a second sample (of 32 counties) that excludes non-shale counties that share a border with a shale county. Weber (2012) argues that shocks to high-population counties will cause large changes in absolute employment and income relative to smaller counties. To address this problem, I apply same estimation to a third data set with all the counties except the only two metropolitan counties, Oklahoma county and Tulsa county, to avoid their excessive influence on the regression results. The results using this sample which includes 47 counties are robust[11].[12]

The DiD estimator is based on a strong identifying assumption: the availability of a treatment and a control group that would have had a similar trend without the treatment. Abadie (2005) argues that this assumption may be implausible if pre-treatment characteristics that are thought to be associated with the dynamics of the outcome variable are unbalanced between the treated and the untreated. To address this issue, I present results using the solution, proposed by Abadie's (2005) and frequently used in the literature, to include a set of covariates; these covariates are used to describe how the average effect of the treatment varies with changes in observed characteristics. I also follow the literature (e.g., Linden and Rockoff, 2008; Caselli and Michaels, 2013) to examine preexisting differences in counties' characteristics using a cross-sectional estimator. For this matter, first I limit the data to before 2008 and use:

$$ln(Y_{ct}) = \alpha_t + \pi_1 Shale_{ct} + \epsilon_{ct} \qquad (1)$$

This equation shows whether the log of counties' characteristics, Y_{ct}, such as housing prices, inequality, affordability, income per capita, housing density, and population density, is a function of the shale development that happened in that county, $Shale_c$. α_t is a year fixed effect. To ensure robust results, I use both absolute level and relative values of the dependent variables. The identifying assumption is that the relative values would have followed parallel trends across counties in the absence of shale boom. I define relative value by diving the absolute value of the variable of interest by the average of the variable in each given year among all counties. The results using relative values as dependent variable are robust.

Given the similarity of pre-treatment trends in shale and non-shale counties, I then use a linear DiD model to estimate shale boom effect on local communities:

$$ln(Outcome_{ct}) = \beta_0 + \beta_1 Shale_c + \beta_2 Post2008_t + \beta_3 Shale_c * Post2008_t + \mu_c + \upsilon_t + \epsilon_{ct} \qquad (2)$$

$Post2008_t$ is an indicator for post-treatment, the interaction between $Shale_c$ and $Post2008_t$ estimates shale development impact on shale counties compared to non-shale, μ_c and υ_t are county and year fixed effects,

[11] These results are available upon request.

[12] Appendix B contains the names of these 49 counties included in the first sample and 32 counties included in second sample.

and ϵ_{it} is an error term. The dependent variable, $ln(Outcome_{ct})$, represents outcomes of interest: housing price index, Gini coefficient, and housing affordability index. The coefficients in equation (2) allow us to measure the shale boom impact in Oklahoma; β_0 shows the average outcome in non-shale counties before shale boom, β_1 measures the differences between shale and non-shale counties before 2008, and β_2 captures changes in outcome in non-shale counties after the boom. β_3 is the coefficient of interest which measures the average shale development effect on shale counties by differencing the changes in outcomes in shale counties after 2008 with non-shale counties[13].

Previous studies have used various control variables when employing DiD method, most commonly, geographical variables (such as latitude and longitude and distance to the closest city or state capital). While geographic location may matter for long-term growth trajectories, growth over a particular decade likely reflects the structure of the economy in the baseline year and how subsequent shocks interact with this structure (Weber, 2012). I, therefore, included income per capita, housing density and population density using the Census 2010 data as a set of control variables that allows counties with different characteristics to have different outcome. To make the specification even more robust, I follow the literature to include control variables for neighboring counties as well:

$$ln(Outcome_{ct}) = \beta_0 + \beta_1 Shale_c + \beta_2 Post2008_t + \beta_3 Shale_c * Post2008_t + \delta_1 X_c + \delta_2 C_c + \upsilon_t + \epsilon_{ct} \tag{3}$$

where X_c is a set of control variables capturing counties' observable characteristics and C_c is neighboring counties' observable characteristics.

4 Results

As explained before, all the results are reported for two samples: the first sample includes all 49 counties while the second sample excludes non-shale counties that share a border with a shale county. First, table 1 provides the summary statistics of the data used in this paper between 2004 and 2017. The upper part of the table shows the mean and standard deviation (SD) for all 49 counties and the bottom part shows the mean and SD for the second sample that includes 32 counties. For the first sample (49 counties), the housing price index is 65.12 on average with a standard deviation (SD) of 6.47, the Gini coefficient is 0.19 on average with an SD of 0.07, and the housing affordability index is 836.72 with an SD of 121.27, and the second sample (32 counties), the housing price index, the Gini coefficient, the housing affordability index

[13]Recent studies that use DiD method often use data from many years rather than only pre and post treatment, like Card and Krueger (1994), pioneers of the natural experiment did, causing the conventional standard errors to be biased downward (e.g. Moulton, 1990, Bertrand, 2004). One of the suggested solutions in the literature that was used here for addressing such a problem is to cluster standard errors at the county level.

and median household income are similar to the first sample. The population and housing density average and SD, however, are lower in the second sample.

Using equation (1) and limiting the sample to before 2008, table 2 displays that there were no preexisting differences in any of the counties' characteristics (neither in housing prices, inequality, or affordability, nor in income per capita, housing, or population density). The bottom part of table 2 shows the results using the same equation for the second sample (32 counties); I find no preexisting differences even when reducing the control sample to only counties with no border with shale counties. Finding no evidence of preexisting differences, I move to equation (2) to estimate the shale boom's impact on the housing price index, Gini coefficient, and housing affordability index. Note that the Gini coefficient measures the extent to which the distribution of income among households deviates from perfectly equal distribution, 0 or absolute equality, to 1 or absolute inequality, and the housing affordability index, which is constructed using the median income over the housing price index in each county, is a ratio that shows lower affordability power as the index goes down.

Table 3 presents the DiD estimation results using equation (2) for the housing price index (columns 1 to 3), Gini coefficient (columns 4 to 6), and housing affordability index (columns 7 to 9) with robust standard errors in parentheses. Columns (1), (4), and (7) show the basic results with only county and year fixed effects but no other control variable, while columns (2), (5), and (8) show the results using counties' observable covariates, and column (3), (6), and (9) include neighboring counties' characteristics as well. As shown in the first three columns, the DiD estimation show that shale development had a positive impact on housing prices; in the average shale county, the value of the housing price index increased by 5.5% due to the shale boom. The coefficient remains the same even after including covariates, column (2), and neighboring counties' characteristics, column (3). These results are statistically significant at a level of 1%. For income inequality, this model fails to find any significant differences between shale counties and non-shale counties associated with the shale boom; the coefficients on $Shale * Post2008$ are not statistically different from zero at any level of confidence in columns (4), (5), or (6). The three last columns show the boom's impact on the housing affordability index. The DiD results suggest a negative impact on the affordability index; basic results with county and year fixed effect, column (7), and results including covariates, column (8), show that the shale boom led to a 5.5% decrease in the affordability index. Adding control variables for neighboring counties increases the coefficient to 0.06 (7%), with higher statistical significance, the level of 1%. All three estimations prove a negative effect on the affordability index; being a shale county is associated with lower affordability power for households.

Table 4 shows that these results are robust using the second sample. Excluding control counties that share borders with shale counties gives the same results: Shale development led to an increase in the housing price index and a decrease in housing affordability index, which translates to higher housing prices and lower affordability. Note that in columns (1) to (3), results regarding housing price index, coefficients of interest are slightly higher (by less than 1%). Both tables (using the same equation but different samples) fail to detect any statistically significant effect on inequality in boom counties. This contradicts some previous studies, such as Weber (2011), but confirms some others, like Michaels (2011).

5 Robustness Check

5.1 Falsification Test

Although cross-sectional analysis results show no evidence of any preexisting differences in shale and non-shale counties, following Linden and Rockoff (2008), I use the falsification test as a robustness check to investigate parallel trend assumption using alternative years, 2007 and 2009, for the year that the shale boom happened. Table 5 shows the results using equation (1) with 2007 as the treatment year. These results are robust: Shale boom increased the housing price index and decreased the housing affordability index which means higher housing prices and lower affordability power in shale counties compared to non-shale. Table 6 shows the same robust results using 2009. Both falsification tests present results confirming that the shale development boom had a positive effect on the housing price index and a negative effect on the housing affordability index while the difference in income inequality between shale counties and non-shale counties is not statistically different from zero.

5.2 Quadratic Time Trend

To relax DiD's constant growth assumption, I follow Cosgrove et al. (2015) to include a quadratic time trend in DiD estimation[14]:

$$ln(Outcome_{ct}) = \beta_0 + \beta_1 Shale_c + \beta_2 Post2008_t + \beta_3 Shale_c * Post2008_t + \beta_4 t + \beta_5 t^2 + \mu_c + \upsilon_t + \epsilon_{ct} \quad (4)$$

Table 7 shows the results using this equation, with and without counties' and their neighbors' observable characteristics. Whether I control for covariates or not, the results remain the same: Coefficients on table 7 show that housing prices in shale counties were more expensive than other counties. Moreover, although increase in employment and income are known to be the most significant impact of shale boom, inequality differences between shale and non-shale counties were not significant. According to the last three columns, households living in shale counties experienced affordability problems due to the shale boom.

[14]Cosgrove et al. (2015) argue that including this quadratic time trend is particularly important since our analysis includes periods of both increasing and decreasing prices for natural gas and the 2007–2009 recession.

6 Conclusion and Policy Implication

Recent developments in the oil and gas industry, combined with an increase in prices, have led to an expansion in oil and gas production in many regions across the United States, generating public concerns about the associated costs and benefits to local communities. Despite a large body of literature, there have not been enough studies on the industry's impact on inequality and affordability. In this study, I examine the impact of the shale boom on local communities in Oklahoma between 2004 and 2017 using county-level data and the DiD method.

First, I use a sample limited to data before 2008 to provide evidence of no significant preexisting differences in shale and non-shale counties. Then, I use a linear DiD estimation to show the boom's effect on the housing price index, income inequality, and the housing affordability index. The results suggest that the shale boom was associated with appreciation in housing values in the state of Oklahoma for shale counties compared to non-shale counties; an average shale county experienced an increase of 5.5% in its housing price index and housing affordability index. Although previous literature provides evidence for higher employment and income in shale counties due to the boom, I find that the effect of the boom on income inequality was not significantly different from zero. The results are consistent across different samples, with or without covariates.

The results of this research have long-term implications for policymakers since oil and gas activities are rapidly expanding and their prices have recently been fluctuating drastically. There is, however, still a lot of room for future research on this subject. Today, one of the empirical challenges is the lack of data on royalty payments and migration to or out of these counties; many of today's oil and gas industry workers in Oklahoma moved when the shale boom happened. Therefore, comprehensive data on payments and migration patterns would present an interesting area of study. Moreover, it is noteworthy that there are many different channels through which the energy industry may impact local communities. While these results provide convincing evidence of the shale boom's impact on households, they do not specify any channel or the magnitude of their positive or negative contributions.

Oklahoma remains highly dependent on energy production, as many other energy-dependent states do. With this in mind, to overcome the resource curse, local and state governments require thoughtful long-term planning. Policymakers should choose between short-term taxation and long-term development. In the long run, a larger industry will be a far greater generator of government tax revenues than an industry stunted by high taxes or costly regulations (Considine et al., 2009). How to best use these tax revenues is a key policy concern.

References

[1] Alberto Abadie. "Semiparametric difference-in-differences estimators". In: *The Review of Economic Studies* 72.1 (2005), pp. 1–19.

[2] Brooklynn J Anderson and Gene L Theodori. "LOCAL LEADERS'PERCEPTIONS OF ENERGY DEVELOPMENT IN THE BARNETT SHALE." In: *Southern rural sociology* 24.1 (2009).

[3] Andrew T Balthrop and Zackary Hawley. "I can hear my neighbors' fracking: The effect of natural gas production on housing values in Tarrant County, TX". In: *Energy Economics* 61 (2017), pp. 351–362.

[4] Ashley Bennett and John Loomis. "Are housing prices pulled down or pushed up by fracked oil and gas wells? A Hedonic price analysis of housing values in Weld County, Colorado". In: *Society & natural resources* 28.11 (2015), pp. 1168–1186.

[5] Sambit Bhattacharyya and Jeffrey G Williamson. "Distributional impact of commodity price shocks: Australia over a century". In: (2013).

[6] Dan Black, Terra McKinnish, and Seth Sanders. "The economic impact of the coal boom and bust". In: *The Economic Journal* 115.503 (2005), pp. 449–476.

[7] Andrew Boslett, Todd Guilfoos, and Corey Lang. "Valuation of expectations: A hedonic study of shale gas development and New York's moratorium". In: *Journal of Environmental Economics and Management* 77 (2016), pp. 14–30.

[8] Peter C Boxall, Wing H Chan, and Melville L McMillan. "The impact of oil and natural gas facilities on rural residential property values: a spatial hedonic analysis". In: *Resource and energy economics* 27.3 (2005), pp. 248–269.

[9] Kathryn J Brasier et al. "RESIDENTS'PERCEPTIONS OF COMMUNITY AND ENVIRONMENTAL IMPACTS FROM DEVELOPMENT OF NATURAL GAS IN THE MARCELLUS SHALE: A COMPARISON OF PENNSYLVANIA AND NEW YORK CASES". In: *Journal of Rural Social Sciences* 26.1 (2011), p. 32.

[10] Jason P Brown. "Production of natural gas from shale in local economies: a resource blessing or curse?" In: *Economic Review (Kansas City)* (2014), pp. 119–119.

[11] Tullio Buccellato and Tomasz Mickiewicz. "Oil and gas: a blessing for the few. Hydrocarbons and inequality within regions in Russia". In: *Europe-Asia Studies* 61.3 (2009), pp. 385–407.

[12] Francesco Caselli and Guy Michaels. "Do oil windfalls improve living standards? Evidence from Brazil". In: *American Economic Journal: Applied Economics* 5.1 (2013), pp. 208–38.

[13] Ron Cheung, Daniel Wetherell, and Stephan Whitaker. "Induced earthquakes and housing markets: Evidence from Oklahoma". In: *Regional Science and Urban Economics* 69 (2018), pp. 153–166.

[14] Susan Christopherson and Ned Rightor. "How shale gas extraction affects drilling localities: Lessons for regional and city policy makers." In: *Journal of Town & City Management* 2.4 (2012).

[15] Timothy Considine et al. "An emerging giant: prospects and economic impacts of developing the Marcellus Shale natural gas play". In: (2009).

[16] Brendan M Cosgrove et al. "The economic impact of shale gas development: A natural experiment along the New York/Pennsylvania border". In: *Agricultural and Resource Economics Review* 44.2 (2015), pp. 20–39.

[17] Frank A Cowell and Fatemeh Mehta. "The estimation and interpolation of inequality measures". In: *The Review of Economic Studies* 49.2 (1982), pp. 273–290.

[18] Frank Alan Cowell. "Measurement of inequality". In: *Handbook of income distribution* 1 (2000), pp. 87–166.

[19] Michael S Delgado, Todd Guilfoos, and Andrew Boslett. "The cost of unconventional gas extraction: A hedonic analysis". In: *Resource and Energy Economics* 46 (2016), pp. 1–22.

[20] Michael D Farren. *Boomtowns and the nimbleness of the housing market: The impact of shale oil and gas drilling on local housing markets.* Tech. rep. 2014.

[21] Nicole M Fortin and Thomas Lemieux. "Changes in wage inequality in Canada: An interprovincial perspective". In: *Canadian Journal of Economics/Revue canadienne d'économique* 48.2 (2015), pp. 682–713.

[22] Mark W Frank. "Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures". In: *Economic Inquiry* 47.1 (2009), pp. 55–68.

[23] Sathya Gopalakrishnan and H Allen Klaiber. "Is the Shale Energy Boom a Bust for Nearby Residents? Evidence from Housing Values in Pennsylvania". In: *American Journal of Agricultural Economics* 96.1 (2014), pp. 43–66.

[24] Hyunseok Hwang and Laurie E Paarlberg. "The Energy Boom: Boon or Bane for Local Philanthropy?" In: *Social Science Quarterly* 100.5 (2019), pp. 1899–1915.

[25] Grant D Jacobsen. "Who wins in an energy boom? Evidence from wage rates and housing". In: *Economic Inquiry* 57.1 (2019), pp. 9–32.

[26] Alexander James and Brock Smith. "There will be blood: Crime rates in shale-rich US counties". In: *Journal of Environmental Economics and Management* 84 (2017), pp. 125–152.

[27] Alan J Krupnick and Isabel Echarte. “Housing Market Impacts of Unconventional Oil and Gas Development”. In: *Resources for the Future* (2017).

[28] Misty Lawrie, Matthew Tonts, and Paul Plummer. “Boomtowns, resource dependence and socio-economic well-being”. In: *Australian Geographer* 42.2 (2011), pp. 139–164.

[29] Leigh Linden and Jonah E Rockoff. “Estimates of the impact of crime risk on property values from Megan’s laws”. In: *American Economic Review* 98.3 (2008), pp. 1103–27.

[30] Aviva Litovitz et al. “Estimation of Regional Air-Quality Damages from Marcellus Shale Natural Gas Extraction in Pennsylvania”. In: *Air Quality*. Apple Academic Press, 2016, pp. 167–186.

[31] Norman Loayza, Alfredo Mier Y Teran, and Jamele Rigolini. *Poverty, inequality, and the local natural resource curse*. The World Bank, 2013.

[32] Alejandro López-Feldman, Jorge Mora, and J Edward Taylor. “Does natural resource extraction mitigate poverty and inequality? Evidence from rural Mexico and a Lacandona Rainforest Community”. In: *Environment and Development Economics* 12.2 (2007), pp. 251–269.

[33] Joseph Marchand. “The distributional impacts of an energy boom in Western Canada”. In: *Canadian Journal of Economics/Revue canadienne d’économique* 48.2 (2015), pp. 714–735.

[34] Robert K Merton. “Social structure and anomie”. In: *American sociological review* 3.5 (1938), pp. 672–682.

[35] Guy Michaels. “The long term consequences of resource-based specialisation”. In: *The Economic Journal* 121.551 (2010), pp. 31–57.

[36] Lucija Muehlenbachs, Elisheba Spiller, and Christopher Timmins. “Shale gas development and the costs of groundwater contamination risk”. In: *Resources for the Future Discussion paper* (2013).

[37] Lucija Muehlenbachs, Elisheba Spiller, and Christopher Timmins. “The housing market impacts of shale gas development”. In: *American Economic Review* 105.12 (2015), pp. 3633–59.

[38] Abdul Munasib and Dan S Rickman. “Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis”. In: *Regional Science and Urban Economics* 50 (2015), pp. 1–17.

[39] Jacquelyn Pless. “Regulating hydraulic fracturing: states take action”. In: National Conference of State Legislatures Denver (CO). 2010.

[40] Nathan Ratledge and Laura Zachary. “Impacts of Unconventional Oil and Gas Booms on Public Education”. In: (2017).

[41] Nicholas J Sanders et al. *Toxic assets: How the housing market responds to environmental information shocks*. Tech. rep. 2012.

[42] Kai A Schafft et al. "Local impacts of unconventional gas development within Pennsylvania's Marcellus shale region: Gauging boomtown development through the perspectives of educational administrators". In: *Society & Natural Resources* 27.4 (2014), pp. 389–404.

[43] Thomas Tunstall. "Recent economic and community impact of unconventional oil and gas exploration and production on South Texas counties in the Eagle Ford Shale area". In: *Journal of Regional Analysis and Policy* 45.1100-2016-90147 (2015), p. 82.

[44] Juan M Villa. "diff: Simplifying the estimation of difference-in-differences treatment effects". In: *The Stata Journal* 16.1 (2016), pp. 52–71.

[45] Jeremy G Weber. "The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming". In: *Energy Economics* 34.5 (2012), pp. 1580–1588.

[46] Jeremy G Weber, J Wesley Burnett, and Irene M Xiarchos. "Broadening benefits from natural resource extraction: Housing values and taxation of natural gas wells as property". In: *Journal of Policy Analysis and Management* 35.3 (2016), pp. 587–614.

[47] Jeremy G Weber and Claudia Hitaj. "What can we learn about shale gas development from land values? Opportunities, challenges, and evidence from Texas and Pennsylvania". In: *Agricultural and Resource Economics Review* 44.2 (2015), pp. 40–58.

[48] Minyu Zhou et al. "The Coal Mining Industry and Income Inequality Across the United States and Appalachia". In: *Proceedings Environmental Considerations in Energy Production* (2015), pp. 183–194.

Tables

First Sample	All 47 Counties		Shale Counties		Non-Shale Counties	
	Mean	SD	Mean	SD	Mean	SD
Housing Price Index	65.12	6.47	64.49	5.13	65.37	6.93
Gini Coefficient	0.19	0.07	0.20	0.08	0.19	0.07
Housing Affordability Index	836.72	121.27	849.20	113.37	831.73	124.04
Median Households Income	38509.37	6051.36	40821.83	6523.47	37734.75	5718.75
Population Density	99.75	205.07	60.47	36.41	115.46	239.82
Housing Density	43.95	90.52	26.73	14.44	50.84	105.96
Second Sample	All 32 Counties		Shale Counties		Non-Shale Counties	
	Mean	SD	Mean	SD	Mean	SD
Housing Price Index	65.04	6.19	64.51	5.21	65.41	6.77
Gini Coefficient	0.19	0.07	0.20	0.08	0.18	0.07
Housing Affordability Index	849.11	128.03	853.51	114.22	846.10	136.81
Median Households Income	38826.48	6496.09	40821.83	6523.47	37461.24	6124.50
Population Density	63.46	81.20	64.13	35.21	63.00	101.36
Housing Density	27.60	33.20	28.35	13.69	27.08	41.60

Table 1: Summary Statistics of data from 2004 to 2017. All variables are county level. The upper part of this table shows mean and standard deviation for the first sample of 47 counties of Oklahoma, the bottom part shows mean and standard deviation for the second sample of 32 counties.

First Sample	Housing Price Index	Gini Coefficient	Housing Price Affordability	Income Per Capita	Housing Density	Population Density
Shale County	-0.0130	0.0299	0.0279	2711.2	-24.11	-54.99
	(-0.80)	(0.55)	(1.20)	(1.35)	(-1.29)	(-1.29)
Constant	4.165***	-2.202***	6.617***	37734.8***	50.84**	115.5**
	(446.03)	(-58.58)	(504.01)	(38.28)	(2.78)	(2.79)
Observations	196	196	196	196	196	196
R^2	0.125	0.699	0.071	0.041	0.015	0.015
Second Sample	**Housing Price Index**	**Gini Coefficient**	**Housing Price Affordability**	**Income Per Capita**	**Housing Density**	**Population Density**
Shale County	-0.00909	0.0746	0.0202	3360.6	1.275	1.126
	(-0.47)	(1.10)	(0.73)	(1.42)	(0.12)	(0.04)
Constant	4.163***	-2.250***	6.628***	37461.2***	27.08**	63.01*
	(332.48)	(-50.90)	(342.96)	(25.87)	(2.75)	(2.63)
Observations	128	128	128	128	128	128
R^2	0.142	0.696	0.063	0.065	0.000	0.000

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 2: Results for examining preexisting differences in counties' characteristics using a cross-sectional estimator. The upper part of this table shows the results for the first sample of 47 counties of Oklahoma and the bottom part shows the results for the second sample of 32 counties.

	Dependent Variable								
	Housing Price Index			Gini Coefficient			Housing Affordability Index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post 2008	0.00329	0.00479	0.00291	0.382***	0.382***	0.313***	0.158***	0.157***	0.160***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.04)	(0.02)	(0.02)	(0.02)
Shale County	0.103***	0.0243	0.0340	0.363***	0.00889	0.00652	-0.0248	-0.0139	-0.0257
	(0.01)	(0.02)	(0.02)	(0.04)	(0.06)	(0.06)	(0.02)	(0.02)	(0.02)
Shale x Post 2008	0.0554***	0.0553***	0.0580***	0.0549	0.0549	0.0481	-0.0546**	-0.0547**	-0.0633***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)
Income Per Capita		0.0167	0.0204		-0.00638	-0.0157		-0.0276*	-0.0373**
		(0.01)	(0.01)		(0.02)	(0.03)		(0.01)	(0.01)
Population Density		-0.00161	-0.00137		0.00164	0.000680		0.00161	0.00120
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
Housing Density		0.00371	0.00325		-0.00383	-0.00184		-0.00380	-0.00298
		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)	(0.00)
Population Density (of neighboring counties)			-0.00138			0.00103			-0.00197
			(0.00)			(0.00)			(0.00)
Housing Density (of neighboring counties)			0.00323			-0.00244			0.00424
			(0.00)			(0.01)			(0.01)
Median Income (of neighboring counties)			-0.00000248			0.00000790			0.00000799
			(0.00)			(0.00)			(0.00)
Constant	4.178***	4.126***	4.211***	-2.627***	-2.184***	-2.460***	6.636***	6.673***	6.390***
	(0.01)	(0.02)	(0.13)	(0.04)	(0.05)	(0.36)	(0.01)	(0.02)	(0.17)
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	684	684	628	686	686	630	684	684	628
Adj.R^2	0.834	0.323	0.352	0.682	0.562	0.575	0.392	0.434	0.477

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3: Regression results of the shale development's impact on the housing price index, the Gini coefficient, and the housing affordability index for shale counties compared to non-shale before and after the shale boom (using the first sample).

	Dependent Variable								
	Housing Price Index			Gini Coefficient			Housing Affordability Index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post 2008	-0.0673***	-0.0673***	-0.0244**	0.518***	0.518***	0.316***	0.227***	0.227***	0.125***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.04)	(0.02)	(0.02)	(0.02)
Shale County	0.169***	0.0323	0.0452	0.335***	-0.0269	-0.0166	-0.0411	-0.0277	-0.0436
	(0.01)	(0.02)	(0.02)	(0.05)	(0.07)	(0.08)	(0.03)	(0.03)	(0.03)
Shale x Post 2008	0.0616***	0.0616***	0.0672***	0.0949	0.0949	0.0927	-0.0502*	-0.0502*	-0.0636*
	(0.02)	(0.02)	(0.02)	(0.07)	(0.07)	(0.07)	(0.02)	(0.02)	(0.02)
Income Per Capita		0.0261	0.0238		0.000278	-0.00693		-0.0333*	-0.0397*
		(0.01)	(0.01)		(0.03)	(0.03)		(0.01)	(0.01)
Population Density		-0.00154	-0.00117		-0.00210	-0.00233		-0.000831	-0.00114
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
Housing Density		0.00424	0.00341		0.00600	0.00585		0.00204	0.00267
		(0.00)	(0.00)		(0.01)	(0.01)		(0.01)	(0.01)
Population Density (of neighboring counties)			-0.00311			-0.00128			-0.000641
			(0.00)			(0.00)			(0.00)
Housing Density (of neighboring counties)			0.00697			0.00345			0.00137
			(0.00)			(0.01)			(0.01)
Median Income (of neighboring counties)			0.00000180			0.00000732			0.00000498
			(0.00)			(0.00)			(0.00)
Constant	4.179***	4.090***	4.023***	-2.642***	-2.240***	-2.526***	6.655***	6.690***	6.512***
	(0.01)	(0.02)	(0.13)	(0.04)	(0.07)	(0.40)	(0.02)	(0.03)	(0.19)
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	448	448	392	448	448	392	448	448	392
$Adj.R^2$	0.828	0.381	0.440	0.673	0.561	0.578	0.381	0.427	0.472

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4: Regression results of the shale development's impact on the housing price index, the Gini coefficient, and the housing affordability index for shale counties compared to non-shale before and after the shale boom (using the second sample).

	Dependent Variable								
	Housing Price Index			Gini Coefficient			Housing Affordability Index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post 2007	0.0264*	0.0285*	0.0269*	1.049***	1.049***	0.313***	0.153***	0.153***	0.154***
	(0.01)	(0.01)	(0.01)	(0.06)	(0.06)	(0.04)	(0.02)	(0.02)	(0.02)
Shale County	0.103***	0.0243	0.0340	0.363***	0.00889	0.00652	-0.0248	-0.0139	-0.0257
	(0.01)	(0.02)	(0.02)	(0.04)	(0.06)	(0.06)	(0.02)	(0.02)	(0.02)
Shale x Post 2007	0.0554***	0.0553***	0.0580***	0.0549	0.0549	0.0481	-0.0546**	-0.0547**	-0.0633***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)
Income Per Capita		0.0167	0.0204		-0.00638	-0.0157		-0.0276*	-0.0373**
		(0.01)	(0.01)		(0.02)	(0.03)		(0.01)	(0.01)
Population Density		-0.00161	-0.00137		0.00164	0.000680		0.00161	0.00120
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
Housing Density		0.00371	0.00325		-0.00383	-0.00184		-0.00380	-0.00298
		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)	(0.00)
Population Density (of neighboring counties)			-0.00138			0.00103			-0.00197
			(0.00)			(0.00)			(0.00)
Housing Density (of neighboring counties)			0.00323			-0.00244			0.00424
			(0.00)			(0.01)			(0.01)
Median Income (of neighboring counties)			-0.00000248			0.00000790			0.00000799
			(0.00)			(0.00)			(0.00)
Constant	4.178***	4.126***	4.211***	-2.627***	-2.184***	-2.460***	6.636***	6.673***	6.390***
	(0.01)	(0.02)	(0.13)	(0.04)	(0.05)	(0.36)	(0.01)	(0.02)	(0.17)
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	684	684	628	686	686	630	684	684	628
Adj.R^2	0.834	0.323	0.352	0.682	0.562	0.575	0.392	0.434	0.477

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5: Regression results of the falsification test: the shale development's impact on the housing price index, the Gini coefficient, and the housing affordability index for shale counties compared to non-shale before and after shale the boom using 2007 as the year that the shale boom happened (using the first sample).

	Dependent Variable								
	Housing Price Index			Gini Coefficient			Housing Affordability Index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post 2009	0.00329	0.0285*	0.0269*	0.382***	0.382***	0.452***	0.158***	0.153***	0.154***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)
Shale County	0.103***	0.0243	0.0340	0.363***	0.00889	0.00652	-0.0248	-0.0139	-0.0257
	(0.01)	(0.02)	(0.02)	(0.04)	(0.06)	(0.06)	(0.02)	(0.02)	(0.02)
Shale x Post 2009	0.0554***	0.0553***	0.0580***	0.0549	0.0549	0.0481	-0.0546**	-0.0547**	-0.0633***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)
Income Per Capita		0.0167	0.0204		-0.00638	-0.0157		-0.0276*	-0.0373**
		(0.01)	(0.01)		(0.02)	(0.03)		(0.01)	(0.01)
Population Density		-0.00161	-0.00137		0.00164	0.000680		0.00161	0.00120
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
Housing Density		0.00371	0.00325		-0.00383	-0.00184		-0.00380	-0.00298
		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)	(0.00)
Population Density (of neighboring counties)			-0.00138			0.00103			-0.00197
			(0.00)			(0.00)			(0.00)
Housing Density (of neighboring counties)			0.00323			-0.00244			0.00424
			(0.00)			(0.01)			(0.01)
Median Income (of neighboring counties)			-0.00000248			0.00000790			0.00000799
			(0.00)			(0.00)			(0.00)
Constant	4.178***	4.126***	4.211***	-2.627***	-2.184***	-2.460***	6.636***	6.673***	6.390***
	(0.01)	(0.02)	(0.13)	(0.04)	(0.05)	(0.36)	(0.01)	(0.02)	(0.17)
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	684	684	628	686	686	630	684	684	628
Adj.R^2	0.834	0.323	0.352	0.682	0.562	0.575	0.392	0.434	0.477

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 6: Regression results ofthe falsification test: the shale development's impact on the housing price index, the Gini coefficient, and the housing affordability index for shale counties compared to non-shale before and after shale the boom using 2009 as the year that the shale boom happened (using the first sample).

	Dependent Variable								
	Housing Price Index			Gini Coefficient			Housing Affordability Index		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post 2008	-0.273**	-0.280**	-0.285**	-0.606**	-0.606**	-0.584*	0.202	0.214	0.241
	(0.09)	(0.09)	(0.10)	(0.22)	(0.22)	(0.23)	(0.20)	(0.20)	(0.22)
Shale County	0.174***	0.0243	0.0340	0.363***	0.00889	0.00652	-0.159***	-0.0139	-0.0257
	(0.01)	(0.02)	(0.02)	(0.04)	(0.06)	(0.06)	(0.01)	(0.02)	(0.02)
Shale x Post 2008	0.0554***	0.0553***	0.0580***	0.0549	0.0549	0.0481	-0.0549**	-0.0547**	-0.0633***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)
Income Per Capita		0.0167	0.0204		-0.00638	-0.0157		-0.0276*	-0.0373**
		(0.01)	(0.01)		(0.02)	(0.03)		(0.01)	(0.01)
Population Density		-0.00161	-0.00137		0.00164	0.000680		0.00161	0.00120
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
Housing Density		0.00371	0.00325		-0.00383	-0.00184		-0.00380	-0.00298
		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)	(0.00)
Population Density (of neighboring counties)			-0.00138			0.00103			-0.00197
			(0.00)			(0.00)			(0.00)
Housing Density (of neighboring counties)			0.00323			-0.00244			0.00424
			(0.00)			(0.01)			(0.01)
Median Income (of neighboring counties)			-0.00000248			0.00000790			0.00000799
			(0.00)			(0.00)			(0.00)
Constant	-42.04**	-43.47**	-43.94**	-167.7***	-167.2***	-162.1***	13.93	16.09	19.81
	(14.75)	(15.41)	(15.64)	(35.55)	(34.34)	(36.92)	(32.00)	(31.42)	(34.37)
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	684	684	628	686	686	630	684	684	628
Adj.R^2	0.834	0.323	0.352	0.682	0.562	0.575	0.741	0.434	0.477

t statistics in parentheses

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 7: Regression results including quadratic term: the shale development's impact on the housing price index, the Gini coefficient, and the housing affordability index for shale counties compared to non-shale before and after the shale boom using 2008 as the year that shale boom happened (using the first sample).

7 Appendix

A Compromise Gini Coefficient

The publicly available IRS data are in group form. Therefore, I rely upon the studies of Cowell and Mehta (1982), Cowell (2000), and Frank (2009) to construct a compromise Gini coefficient. Accordingly, as Frank (2009) explains, the lower limit of the Gini coefficient can be derived based on the assumption that all individuals in a group receive exactly the mean income of the group:

$$G_L = \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}\frac{n_i n_j}{n\mu}|\mu_i - \mu_j| \tag{5}$$

where n is the number of individuals, μ is mean income, and subscripts i and j denote within-group values. The upper limit of the Gini coefficient can be derived based on the assumption that individuals within the group receive income equal to either the lower or the upper bound of the group interval:

$$G_U = G_L + \sum_{i=1}^{k}\frac{n_i^2(a_{i+1} - \mu_i)(\mu_i - a_i)}{n^2\mu(a_{i+1} - a_i)} \tag{6}$$

The compromise Gini coefficient proposed by Cowell and Mehta (1982) is then $2/3G_U + 1/3G_L$.

B List of Counties

Here is the list of 49 counties included in the first sample:
Adair, Beckham, Blaine, Bryan, Caddo, Canadian, Carter, Cherokee, Cleveland, Comanche, Craig, Creek, Custer, Garfield, Garvin, Grady, Jackson, Kay, Kingfisher, Le Flore, Lincoln, Logan, McCurtain, McIntosh, Marshall, Mayes, Murray, Muskogee, Noble, Nowata, Oklahoma, Okmulgee, Osage, Ottawa, Pawnee, Payne, Pittsburg, Pontotoc, Pottawatomie, Rogers, Seminole, Sequoyah, Stephens, Texas, Tulsa, Wagoner, Washington, Woods and Woodward.

The list of 32 counties included in the second sample:
Adair, Beckham, Bryan, Canadian, Carter, Cleveland, Garfield, Garvin, Grady, Jackson, Kingfisher, Le Flore, Logan, Marshall, Mayes, McCurtain, Noble, Nowata, Okmulgee, Osage, Ottawa, Payne, Pittsburg, Pottawatomie, Seminole, Sequoyah, Stephens, Texas, Wagoner, Washington, Woods, Woodward.