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

Shahrzad Ghourchian
*Siena College*

## Motivation and Contribution

- 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.

## Methodology

- The difference-in-difference (DiD) method: 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.
- I define the treatment group to be counties that experienced shale extraction and the treatment to be the shale development boom that happened in 2008.
- To avoid endogeneity, shale counties were selected based on their geographic location, as it is common in the literature.
- To avoid spillover effect, I construct and report results for a second sample that excludes non-shale counties that share a border with a shale county.
- Moreover, 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 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. To address this issue, I 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.
- 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:

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

  The dependent variable, $ln(Outcome_c t)$, represents outcomes of interest: housing price index, Gini coefficient, and housing affordability index. $\beta_3$ 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.
- I 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:

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

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

## Results

| | Dependent Variable | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Housing Price Index | | | Gini Coefficient | | | Housing Affordability Index | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Post 2008 | 0.00329<br>(0.01) | 0.00479<br>(0.01) | 0.00291<br>(0.01) | 0.382***<br>(0.05) | 0.382***<br>(0.05) | 0.313***<br>(0.04) | 0.158***<br>(0.02) | 0.157***<br>(0.02) | 0.160***<br>(0.02) |
| Shale County | 0.103***<br>(0.01) | 0.0243<br>(0.02) | 0.0340<br>(0.02) | 0.363***<br>(0.04) | 0.00889<br>(0.06) | 0.00652<br>(0.06) | -0.0248<br>(0.02) | -0.0139<br>(0.02) | -0.0257<br>(0.02) |
| Shale x Post 2008 | 0.0554***<br>(0.01) | 0.0553***<br>(0.01) | 0.0580***<br>(0.01) | 0.0549<br>(0.05) | 0.0549<br>(0.05) | 0.0481<br>(0.05) | -0.0546**<br>(0.02) | -0.0547**<br>(0.02) | -0.0633***<br>(0.02) |
| Constant | 4.178***<br>(0.01) | 4.126***<br>(0.02) | 4.211***<br>(0.13) | -2.627***<br>(0.04) | -2.184***<br>(0.05) | -2.460***<br>(0.36) | 6.636***<br>(0.01) | 6.673***<br>(0.02) | 6.390***<br>(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 |
| County Set of Control Variables | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Neighboring County Set of Control Variables | No | No | Yes | No | No | Yes | No | No | 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 |

Standard errors are in parentheses
$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

## References

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

## Data

- For housing values: county-level annual housing price index (HPI) from the Federal Housing Finance Agency (FHFA)
- For inequality: Internal Revenue Service (IRS) data for the Gini index
- For affordability: the median household income from the IRS data and the real housing price index from FHFA for the frequently used housing affordability index, the income-to-housing price ratio
- For counties' characteristics: Census 2010 data (income per capita, population density, and housing density)

## Conclusion

- The results suggest that the shale boom was associated with appreciation in housing values for 5.5% and a decrease of %6 in affordability in the state of Oklahoma for shale counties compared to non-shale counties.
- Although previous literature provides evidence for higher employment and income in shale counties due to the boom, the estimation fails to find any statistically significant effect on inequality.
- The results are consistent across three different samples, with or without covariates for shale counties or their neighbors.