Who Really Benefits from a Resource Boom? Evidence from the Marcellus and Utica Shale Plays*

R. Kaj Gittings
Department of Economics
Texas Tech University
kaj.gittings@ttu.edu

Travis Roach
Department of Economics
University of Central Oklahoma
troach2@uco.edu

February 19, 2019

Abstract

We analyze the impact of the recent energy boom in the Marcellus and Utica shale formations on local labor markets in Ohio, Pennsylvania and West Virginia. Unique to our analysis is the use of origin-destination files (LODES) from the U.S. Census Bureau which provide employment statistics at the census block level based on where jobs are located and where workers live who hold those jobs. The richness of this data enables us to identify cross border mobility of workers as labor demand increases due to greater resource extraction. We find that increases in the value of new oil and gas production significantly increases local workplace employment and average earnings in the county, but that job gains go mostly to workers who reside outside of the county. Specifically, a one standard deviation increase in the value of new oil and gas production per capita in a county-year increases workplace employment by 283 jobs, decreases the fraction of jobs held locally, and increases the flow of workers from counties 25-200 miles away. Furthermore, we find some positive employment spillovers across industries but these new jobs appear to go to non-local residents as well. We do find evidence that the earnings distribution shifts to the right for both local residents and workers who reside outside the county.

JEL Codes: J21, J61, Q33, Q40

Keywords: Local Labor Markets, Worker Mobility, Resource Boom

*We are grateful for helpful comments and suggestions from Ian Schmutte, seminar participants at the University of Georgia and session attendees at Western Economic Association International.

1 Introduction

'Boomtowns' are a trope that can be found across history in literature and popular media, as well as the economics literature. These so-called resource booms cause a surge of economic activity that injects money into the economy and transforms local labor markets almost immediately. The results of such breakaway activity, though, are not always positive as the companion term, 'resource curse' connotes. In some cases, prices sway in such a way that inhibits further drilling or the local economy cools down once the resource has been depleted. In other cases, the shifting of employment and compensation towards the mining sector and supporting industries has the potential to disturb employment and entrepreneurship in sectors producing goods that are not directly tied to resource extraction. Even unexpected outcomes have occurred like a growth in crime, and changes in marriage and birth rates. In this paper, we examine a new angle of the job creation and labor dynamics that are associated with resource booms (or curses) by examining the origin of workers and the amount of jobs that are created or destroyed for both local, pre-resource boom residents, and non-local, post-resource boom workers.

There is a rich history in the economics literature devoted to studying the effects of resource booms and mineral extraction in particular. These studies vary in their scope (international or sub-national) and in the resource that is examined. However, a few 'stylized facts' and theoretical predictions have come to the surface that are relevant to the nascent hydraulic fracturing boom that is unlocking tight oil and gas formations in areas that previously saw relatively little drilling activity. One vein of the resource boom literature investigates what is known as the resource curse. These studies highlight the somewhat paradoxical finding that countries with large resource endowments witness slower long-run economic growth (Kim and Lin (2015); Sachs and Warner (2001)). One reason behind the resource curse may be that as the demand for labor in the mining sector increases along with wages, there becomes an incentive for workers in other sectors to move into mining. In order to keep employment from exiting towards the mining sector, other sectors must also offer higher wages (Sala-i Martin and Subramanian (2013)). Two potential outcomes occur in the non-mining sectors if these goods are traded outside of the resource rich area: These industries pass-on higher labor costs and lose market share by having higher prices, or output falls as they are unable to fulfill their demand for labor when keeping wages constant because employees leave for other industries. There are, of course, competing stories on the exact mechanism behind why areas that are resource abundant witness lower than average economic growth. Blanco and Grier (2012) show that educational attainment and human capital development may be lower in areas that are resource dependent, and Rickman et al. (2017) shows that this is taking place in shale rich regions. Speaking on changes in entrepreneurship, Betz et al. (2015) and Tsvetkova and Partridge (2017) both show that long-run growth may be hurt by resource dependence and a move away from entrepreneurial ventures outside of the mining sector. Murshed and Serino (2011) considers poor industry diversification and export structures as the causal mechanism for lower growth. Similarly, Ginn and Roach (2015) show that the Texas economy was more resilient to oil price shocks after NAFTA because the state's industries were more diversified. Haggerty et al. (2014) take a long-term look at resource dependence and specialization and provide evidence that resource riches may hamper economic activity. Looking at the boom of the 1980s these authors determine that areas that specialized in oil and gas production for longer periods of time had higher crime rates, lower per capita income, and lower educational attainment. Gourley and Madonia (2016) also show that resource development is associated with higher rates of both violent and property crime. Specific to the recent shale boom, James and Smith (2017) and Komarek (2018) find evidence that crime has increased in areas that have had new drilling take place. Interestingly, Street (2018) is able to differentiate people that migrated to the North Dakota region from local residents and shows that crime rates actually fell among those residents that lived in the area prior to the resource boom. Other negative unintended consequences that are caused by resource booms have been studied including the finding that cases of gonorrhea have increased in the Marcellus region (Komarek and Cseh (2017)), and that marital and non-marital birth rates have increased while marriage rates have stayed the same (Kearney and Wilson (2018)).

Evidence of a resource curse is not wholly conclusive, though, as a number of papers find that resource rich areas actually drive employment and income gains. For instance, there is a chance that resource extraction leads to positive spillover effects in surrounding industries due to an increase in population, and hence demand, for non-mining goods (e.g. food and accommodations). Also, demand may increase due to sudden income gains from royalty payments. Brown et al. (2016) find that for every one dollar increase in royalty payments an additional $0.50 in local income was created. Michaels (2011) shows that counties in the south that had a relatively large oil endowment have had higher population and per capita income growth. Indeed, Alexeev and Conrad (2009) find that a large resource endowment has been, on net, positive for long-term economic growth. Boyce and Herbert Emery (2011) explore resource abundance and growth rates and finds a negative effect on growth, yet also finds a positive effect of resource abundance on income levels. In direct contrast to the theory that a boom in mining sector employment crowds out investment in non-mining sectors (especially tradable goods), Weber (2014) finds that mining and non-mining employment are positively correlated, while Allcott and Keniston (2014) show that mineral resource extraction benefits manufacturing employment. Munasib and Rickman (2015) use synthetic control methods to show that many of these effects are idiosyncratic to the area with large gains found in North Dakota, and zero for Pennsylvania.

Although the myriad impacts of resource booms have been studied extensively, our research makes an important next step in describing how a sudden change in mineral wealth affects local labor markets. Our analysis employs the LEHD Origin-Destination Employment Statistics (LODES) data produced by the Longitudinal Employer-Household Dynamics Program within the U.S. Census Bureau. The primary benefit of these data is that they provide employment statistics at the census block level based on where workers live (origin) and where jobs are located (destination). Therefore, we are able to determine not only the effect on employment and wages across sectors as has been done in prior work, but also how the employment and wage benefits are distributed across workers who live locally and workers who live outside the drilling area. Hence, we can track employment and earnings shifts across geopolitical boundaries. Identification of this sort of worker mobility has not previously been discussed because these statistics are not obtainable with typical employment data on the location of employers (such as the Quarterly Census of Employment and Wages) or with employment data based on where the worker lives (such as the Current Population Survey) which prior work has relied on. Additionally, the Marcellus-Utica shale area is able to serve as a laboratory to study a sudden resource boom in modern, developed economies. Prior to the shale boom, practically no drilling or oil and gas production occurred in the region and the invention of hydraulic fracturing created un-tapped reserves overnight. In this way, the Marcellus-Utica experience is akin to a natural experiment on the effects of a resource boom on job-migration and spillovers. Ours is among the first studies to explain employment migration as a result of a resource boom. We find that the majority of the gains of a resource boom go to those that live outside of the area where drilling is taking place. Specifically, we find that the fraction of jobs held locally broadly declines as the value of new oil and gas production per capita increases. Moreover, we see that in some cases workers are traveling distances from counties that are 25-200 miles away. These results are robust to changes in how we define local employment from a county-based measure to Labor Market Areas (LMA) which are aggregated by commuting zone. We also find evidence of an increase in earnings for both workers that live locally and workers that live outside the county.

2 The Shale Revolution and Local Labor Markets

We are not the first to study the effect of the recent oil and gas boom on local labor markets as the development of hydraulic fracturing has also led to several studies specifically regarding shale plays. Munasib and Rickman (2015) examine regional economic impacts of the recent energy booms in Arkansas, North Dakota, and Pennsylvania. They find that North Dakota witnessed an increase in employment in nearly all sectors following the boom, but that the same could not be said for Arkansas or Pennsylvania. Tsvetkova and Partridge (2016) find that the shale boom has had positive spillover effects in the non-tradeable goods sector and small or negative impacts on traded goods. However, their results suggest that an equivalently sized shock to the rest of the economy produces more jobs on average than oil or gas shocks. Agerton et al. (2017) link rig counts with job creation and are able to determine the short-run and long-run impacts of upstream oil and gas investment. These authors find that one additional rig results in 31 new jobs contemporaneously, and up to 315 jobs over a longer time horizon. Miljkovic and Ripplinger (2016) estimate the effect of the boom on employment within the mining sector and further determine that there has not been spillover effects on agricultural wages.

Most closely related to our work in estimating the effect of energy booms on local and nonlocal employment is Feyrer et al. (2017). They use the Quarterly Census of Employment and Wages (QCEW) and the IRS Statistics of Income data to measure wages and employment at the county level. They find that new oil and gas production increases income within a county, and further find that there are geographic spillovers with wages increased for areas within 100 miles of a well. Data from the QCEW is based on the location of the employer and data from the IRS is based on the location of the tax filer. In this sense, they have wage measures for residents who live locally in the county and wage measures based on jobs located in the county. However, workers in the QCEW may live outside the county and tax filers in the IRS data may work outside the county where they file. The benefit of the LODES data that we utilize in our study is that we can identify jobs in a county that belong to residents who also live in the county versus jobs in the county that are held by individuals living outside the county. Their work is also greater in scope as it analyzes all counties and resource plays in the United States while our Origin-Destination based analysis focuses on the Marcellus and Utica formations in Ohio, Pennsylvania and West Virginia. Also closely related to our study is work by Green et al. (2017) and Wilson (2016). These authors have also addressed the effects of resource booms on migration and the incentive to travel large distances due to the recent fracking boom. Green et al. (2017) show that the resource boom had much larger spillover effects in Canada in part due to commuters. Wilson (2016) uses exogenous variation in pre-fracking newspaper articles and shows that there was a 3.8% increase in in-migration rates in North Dakota due to the resource boom. Additionally, Wilson (2016) makes use of LODES data (as we do here) to show the impact that a change in earnings has on the count of long distance commuters. The author finds that increased earnings were especially effective at incentivizing long-distance commuting (more than 50 miles) into North Dakota, but that the effect was more muted in other parts of the country. Our work ties newly created jobs directly to where the job holder lives, and we find that a substantial fraction of new jobs are held by workers living outside the county which confirms that there has been geographic spillover as in Feyrer et al. (2017).

Other researchers have also focused on the drilling changes in the Marcellus and Utica formations, including Komarek (2016) and Paredes et al. (2015). Komarek (2016) finds some spillover benefits in non oil and gas industries while Paredes et al. (2015) does not. Both studies use employment and wage data from the Bureau of Economic Analysis that is based on place of work only. Komarek (2016) designates counties as either 'high fracking' or 'low fracking' based on the amount of unconventional wells that are drilled. Komarek (2016) finds that in 'high fracking' areas non-traded industries were helped by shale gas development with employment increasing by 3-6% and wages increasing by 8-12%, but that there was no statistically significant effect on manufacturing. Conversely, Paredes et al. (2015) find that there are no spillover effects when using propensity score matching methods. When they use panel data methods they find statistically significant effects on employment, but not for income, concluding that the spillover effects from shale gas development may be minimal. Our work is differentiated from theirs in several ways. First, their primary measures of the intensity of resource extraction are based on well counts while we measure the total value of new oil and gas production in each period. Using well counts rather than production value conflates the size of the shock since some drilled wells may be relatively dry while other wells may be highly productive. Second, we adopt an instrumental variables technique similar to the method used in Feyrer et al. (2017) that accounts for the potential endogeneity of resource endowment and extraction. Finally, our use of the LODES data can differentiate the impact of new drilling on jobs provided in the county, jobs held in the county by local residents, and the cross border mobility of workers who travel from outside the county. Our findings support Komarek (2016) in that there is a large increase in jobs in the Marcellus and Utica shale regions. However, we further find that the majority of these jobs are held by people that reside in counties that are more than 25 miles away. Our IV estimates indicate that a one standard deviation increase in the value of new production per capita decreases the fraction of jobs held locally. We do find evidence that a standard deviation change in the value of new production per capita increases average monthly earnings in the county by about 14%. This appears to represent a rightward shift in the earnings distributions for both workers who live locally and workers living outside the county which manifests as a decrease in the fraction of locally held jobs that pay $1,250 per month or less and an increase in the amount of non-locally held jobs that pay $1,250-3,333 per month.

3 Data

The Marcellus and Utica plays have a significant presence in Ohio, Pennsylvania and West Virginia which form the geographical reference area for our analysis.[1] Our primary data source for local labor market outcomes are the LEHD Origin-Destination Employment Statistics (LODES) produced by the Longitudinal Employer-Household Dynamics Program within the U.S. Census Bureau. The public use files are available with geographic data available at the Census block level which we aggregate to the county level. Of interest to us is that the LODES data provide statistics based on both where the job is located and where workers live who hold those jobs. This allows us to identify the effect of production shocks on the employment outcomes of both the local population and those who work in a particular county but are nonresidents. Data on drilling activity and well production were obtained through a confidentiality agreement with DrillingInfo.org. These data have geographic identifiers for each well that allows us to aggregate well production to the county level and merge with the LODES data. The LODES data also provide job counts conditional on being in particular bins of monthly earnings but do not provide earnings statistics directly. Therefore, we also utilize average monthly earnings data from the Census Bureau's Quarterly Workforce Indicators (QWI).[2]

3.1 LODES Data

The LODES data we use span the years 2002-2014.[3] Statistics are released at the census block level in three main sets of files: OD (Origin-Destination data), RAC (Residence Area

[1]The Marcellus formation also covers part of New York but temporary fracking moratoriums have been in place since at least 2010. Our drilling data indicate that only a handful of wells were drilled in the Marcellus region of New York prior to 2010.

[2]LODES data, QWI data and the corresponding documentation can be found at http://lehd.ces.census.gov/data/.

[3]We use LODES version 7. These data are not updated annually and are on an irregular production schedule.

Characteristics data) and WAC (Workplace Area Characteristics data). While the data are annual, the reference period for each set of files is a snapshot pertaining the second quarter of the year where jobs in the snapshot are presumed to be held as of April 1. Employer location is determined by the physical address reported by employers in the Quarterly Census of Employment and Wages (QCEW). Therefore, job counts provided in the WAC are allocated to the location of the business address. In some cases, this may differ from the worksite where an employee reports, such as construction jobs. The residence location in the RAC files is determined from federal administrative records that identify a single residence location for every worker in each year.

The ability to distinguish between residential and workplace employment and having access to the Origin-Destination files makes the LODES data ideal for understanding job flows across political boundaries (e.g., county borders). Workplace employment (WAC) allocates a job to a location based on the employer address and that may or may not be close to where the worker lives. This is the norm for most official labor market statistics based on administrative data (QCEW, the Census Bureau's Quarterly Workforce Indicators (QWI), and BLS data outside survey products like the Current Population Survey and American Time Use Survey). The residential employment data (RAC) allocates a job based on the worker address. This is more rare (CPS survey data, American Community Survey estimates, decennial census) especially with regard to administrative based data. While employment is reported by establishments in the underlying administrative data, the residential address for workers is obtained by matching the employees reported by establishments with worker tax returns to get a residential address. This is done internally at the Census Bureau before the LODES public use product is released. This distinction between WAC and RAC and the combined OD files is one of the primary reasons why we adopted the LODES for this project.

Earnings are measured from all sources of income from a worker's primary employer during the period. A "primary job" is identified as the job a person held with the most earnings. All statistics in the LODES data in our analysis refer to primary jobs held in the private sector. The OD, RAC and WAC files provide block level job counts for a number of characteristics. From the RAC and WAC files we incorporate the following block level statistics:

- Total Number of Jobs
- Total Number of Jobs Earning $1250/month or Less
- Total Number of Jobs Earning $1251/month to $3333/month
- Total Number of Jobs Earning $3333/month or More
- Total Number of Jobs in Each NAICS Sector[4]

Note that the earnings bins are predefined by LEHD and there is no additional data provided to compute statistics for other earnings ranges. The block level WAC and RAC statistics are aggregated to the county level by simple summation across all the blocks within a county. The block counts in the OD files identify the residential block (origin) and the workplace block (destination). The origin location can cross state boundaries so that, once blocks are aggregated, the job counts for a particular Origin-Destination identify commuting patterns within a county, from outside a county and even across state borders. From the OD files, we incorporate the following statistics:

- Total Number of Jobs
- Total Number of Jobs Earning $1250/month or Less
- Total Number of Jobs Earning $1251/month to $3333/month
- Total Number of Jobs Earning $3333/month or More
- Total Number of Jobs in Goods Producing Industry Sectors

[4]The lowest level of industry disaggregation that the LODES public use data provides is at the 2 digit NAICS sector level.

- Total Number of Jobs in Trade, Transportation, and Utility Industry Sectors
- Total Number of Jobs in All Other Service Industry Sectors

Industry job counts in the OD files are aggregated to industry groups rather than individual NAICS sectors to preserve the confidentiality of the underlying microdata. To aggregate the OD statistics to the county level, we use the NBER County Distance Database to identify origin counties that are (1) within 25 miles of the destination county, (2) counties that are within 25 and 50 miles of the destination county (3) counties between 50 and 125 miles, (4) counties between 125 and 200 miles, and (5) counties that are more than 200 miles away.[5]

Therefore the county-level aggregates from the OD files yield total jobs counts in the categories above that are held by workers locally, held by workers outside the county (nonlocal), and held by workers outside the county in one of the five distance categories. For estimation, the job count statistics from the OD, WAC and RAC files will be expressed as rates where the denominator is either total county population or total employment.[6]

3.2 Well Production Data

Data on drilling activity in the three states we study were obtained from DrillingInfo.org. The DrillingInfo database is a national database that contains detailed information about each well that has received a permit for drilling, has been "spudded" (drilling initiated), and completed. Each well that has been completed is identified with a date that the well was completed, a geographic location of the well, as well as measures of oil and gas production since the completion date. Using this information we compute the total value of new oil and gas production for each county for each month. Average monthly prices per unit of gas

[5] The county distance files can be obtained at http://www.nber.org/data/county-distance-database.html.

[6] To avoid measurement issues associated with a potentially noisy year-to-year local population estimates, we use county population in Census year 2000 as the denominator.

Table 1: Distribution of Counties and Value of New Production per Capita

Panel A: County Exposure to Plays

County	Total	Exposure to Marcellus or Utica	Fully Exposed to Marcellus or Utica
Ohio	88	41	27
Pennsylvania	67	52	40
West Virginia	55	51	40

Panel B: Distribution of Value of New Production per Capita

	Full Sample	2008	2014
All Counties	21.4	7.3	118.7
	(151.3)	(24.4)	(447.3)
Ohio	6.9	1.9	65.8
	(86.3)	(4.8)	(300.1)
Pennsylvania	32.6	1.4	178.6
	(205.2)	(3.6)	(569.0)
West Virginia	30.8	23.2	130.3
	(153.7)	(43.7)	(473.7)

Note: Panel A shows the number of counties in each category. The second column displays the number of counties that have at least some geographical overlap with either play. The third column displays the number of counties which have borders that are fully contained in either play. Panel B shows summary statistics for the value of new production per capita. Entries are means with standard deviations in parentheses.

and oil are obtained from the Henry Hub gas prices and Cushing (West Texas Intermediate) oil prices. As stated above, the LODES data are annual with a reference period of the second quarter of each year. Therefore, we compute an annualized total value of new oil and gas production in the prior twelve months as of April 1 of each year. For the analysis, we measure total value of new production per capita by dividing the total value by total county population in year 2000.

Not all counties in these states are exposed to the Marcellus and Utica plays. Panel A of Table 1 shows the extent to which counties in these states are exposed. The first column

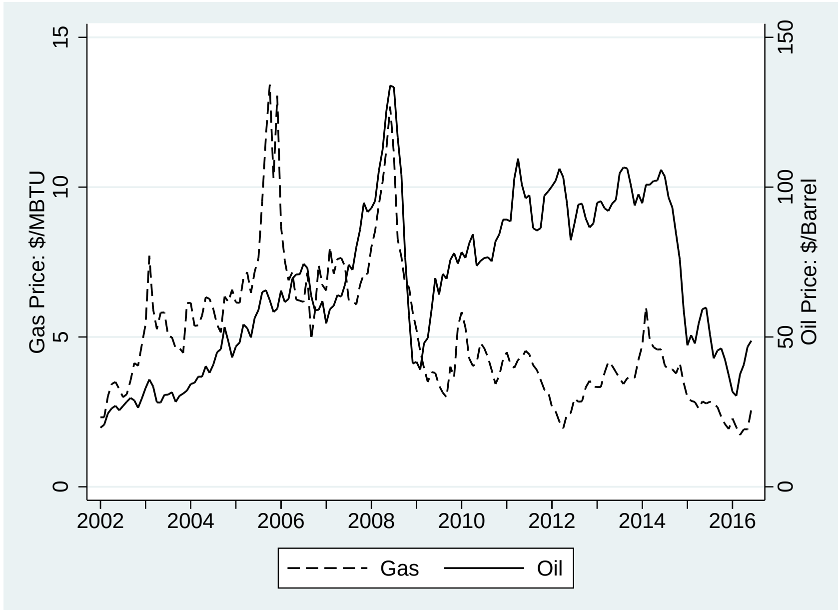

Figure 1: Monthly Oil and Gas Prices.

displays the total number of counties in each state while the next two columns indicate how many counties in each state have some exposure to either play (column 2) or are fully exposed to a play. Exposure in this sense means that a county's borders contain a play. Most of Pennsylvania and West Virginia have some exposure to one of the plays but less than half the counties in Ohio are exposed. The Marcellus and Utica formations are large and a substantial number of counties fully cover at least one of the plays. Figure A1 in the appendix shows a map of counties that are exposed to either play. Figure A2 shows how production increased across geography at different points in the sample period. This geographic variation in exposure to the plays combined with variation in production within the plays is crucial to our identification strategy.

Panel B of Table 1 shows the distribution of the value of new production per capita (our main variable of interest) by state for the full sample and at two different points in the sample (2008 and 2014).[7] Across all counties and years the value of new production was about

[7]Figure A2 in the Appendix shows this data graphically and in more detail.

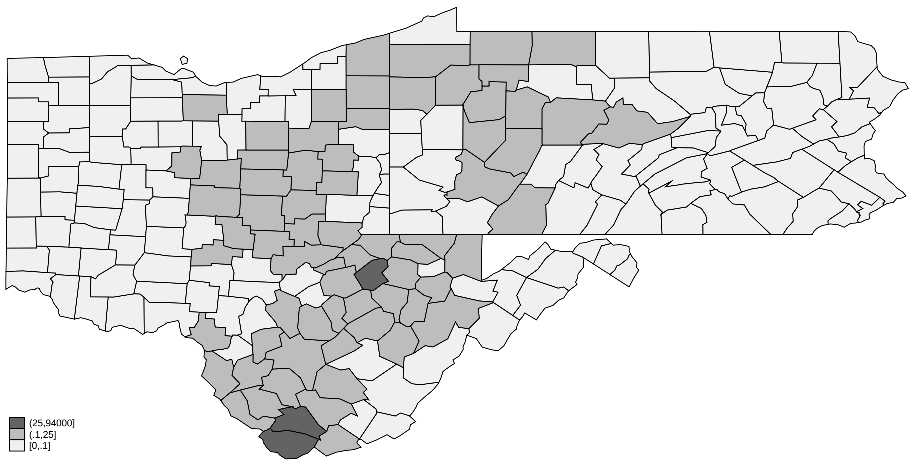

Figure 2: Bins reflect oil and gas production (in $) per capita in 2002 by county.

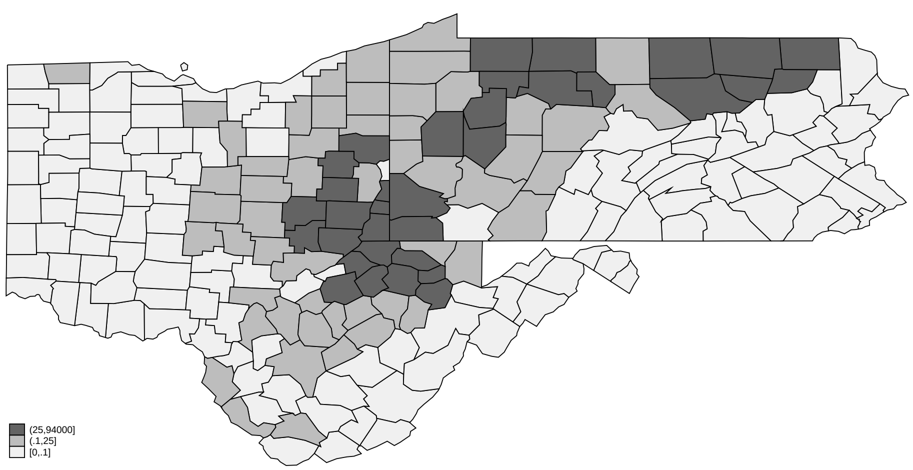

Figure 3: Bins reflect oil and gas production (in $) per capita in 2014 by county.

$21.4 per person, $6.9 per person in Ohio, $32.6 per person in Pennsylvania and $30.8 per person in West Virginia. Oil and gas production in all three states increased substantially since 2008, and standard deviations are large relative to the mean illustrating a substantial amount of variation in our variable of interest.

The Marcellus and Utica plays produce a substantial amount of gas relative to oil production. Figure 1 displays the monthly average prices of gas and oil. While gas and oil prices both peaked around 2008, gas prices have largely declined since while oil prices recovered and then declined again significantly after 2014.[8] Despite the decline in gas prices, the total value of production in the area has increased substantially. Figures 2 and 3 display the geographic change in total value of new production at the beginning and end of our sample period, respectively. The three bins producing the shaded areas in both figures are determined by the distribution per capita production over the entire sample period. Those intervals were chosen to reflect three groups: zero or almost zero production (white), production up to about the mean (light grey), and those counties producing above the mean (dark grey). Figure 3 (2014 production) suggests that most counties produce nothing and the remaining counties are nearly evenly split between producing a positive amount but below the mean and those producing above the mean.

4 Empirical Model and Identification

We aim to estimate the impact of new oil and gas production shocks from the Marcellus and Utica plays on earnings and several measures of employment. The LODES data allows us to distinguish jobs in a county that are held by local residents and jobs that are held by workers who live outside the county. Table 2 provides summary statistics for variables created from the LODES data that we use as outcomes in the empirical model. The table

[8]Our analysis ends in 2014 and is therefore not confounded with the the post-2014 collapse in oil prices.

also shows summary statistics for the log of average monthly earnings from the QWI data and total county population.

As we expect, workplace area and residential area employment statistics roughly measure the same total employment in a county in the aggregate. On average, workplace employment has a higher proportion of low earnings jobs and lower proportion of high earnings jobs than residential area employment. Both employment measures show a relatively small proportion of jobs in the NAICS 21 Sector (Mining, Quarrying, Oil and Gas Exploration). More than half of the destination jobs are in the Goods Producing industry group or the Trade, Transportation and Utilities group. Perhaps most surprising is that only about half of the workplace jobs are held by residents who live in the same county. Of the jobs that belong to workers living outside the county, 18% live within 25 miles of the county where the job is located, 17% live between 25 and 50 miles away and 14% live more than 50 miles away. The average log monthly earnings from the QWI data is 7.9 or about $2630/month.

4.1 Empirical Specification

Our benchmark specification is:

$$y_{ct} = \beta NewValuePerCapita_{ct} + \mu_c + \lambda_t + \pi^c RECESSION_t + \varepsilon_{ct} \tag{1}$$

where y_{ct} is a particular outcome in county c and year t. The main variable of interest is $NewValuePerCapita_{ct}$ which measures the total value of new oil and gas production per capita over the prior twelve months. County fixed effects are represented by μ_c and year fixed effects are represented by λ_t. Our data span the years 2002-2014 which includes the financial crisis and resulting recession. $\pi^c RECESSION_t$ are a set of indicators for county-specific

Table 2: Summary Statistics

	Mean	Std Dev
Log Avg. Monthly Earnings	7.874	(0.199)
Workplace Area Employment	43381.7	(92132.1)
Residential Area Employment	43707.7	(77117.4)
% Workplace Employment <$1250/month	0.289	(0.062)
% Workplace Employment $1250/month to $3333/month	0.431	(0.059)
% Workplace Employment >$3333/month	0.280	(0.092)
% Residential Employment <$1250/month	0.273	(0.045)
%Residential Employment $1250/month to $3333/month	0.429	(0.048)
% Residential Employment >$3333/month	0.298	(0.076)
% Workplace Employment NAICS 21	0.033	(0.070)
% Residential Employment NAICS 21	0.024	(0.041)
% Workplace Employment Goods Producing Industries	0.305	(0.123)
% Workplace Employment Trade Industries	0.234	(0.055)
% Workplace Employment Other Services	0.461	(0.111)
% Workplace Employment Held Locally	0.510	(0.111)
% Workplace Employment Not Held Locally	0.490	(0.111)
% Workplace Employment Not Local (<25 Miles)	0.180	(0.116)
% Workplace Employment Not Local (25 to 50 Miles)	0.169	(0.072)
% Workplace Employment Not Local (50 to 125 Miles)	0.106	(0.041)
% Workplace Employment Not Local (125 to 200 Miles)	0.026	(0.015)
% Workplace Employment Not Local (> 200 Miles)	0.006	(0.006)
Total County Population (year 2000)	121639.7	(211676.7)

Note: N=2730 for all variables. The unit of observation is a county-year. NAICS 21 includes mining, quarrying and oil and gas exploration.

recession effects.[9] Therefore, the $RECESSION_t$ indicator equals one in 2008 and 2009 but the estimated parameters, π^c, pertain to each county individually. In all regression models the error term (ε_{ct}) is clustered by county.

While geological formations and the underlying resources are exogenous to labor market outcomes, an area's ability to extract the oil and gas are not. For instance, there are ample county and state differences in regulations regarding new drilling operations, and, further, there are differences in existing infrastructure in a county that could (at least temporarily) house workers that arrive to extract the resource. We follow Feyrer et al. (2017) and Fetzer (2014) to construct instruments using a two step process that addresses this potential endogeneity problem.[10] The total value of new production is modeled as a function of county effects and a set of indicators for each play-year as in Equation (2):

$$\ln(NewValue_{ct} + 1) = \mu_c + \kappa_{pt} + \nu_{ct} \tag{2}$$

The parameters κ_{pt} capture annual changes in production associated with each play while μ_c accounts for systematic differences in production across counties. There are 121 counties that have at least some exposure to the Marcellus play, 117 that have at least some exposure to the Utica play and 94 counties have exposure to both. The instrument will be valid since any particular county within a play will contribute a small fraction of total production in the play as a whole. The instruments are constructed using the parameter estimates from Equation (2) and scaling the predicted values of $NewValue_{ct}$ by county population in the 2000 census as in Equation (3):

[9] County-specific recession indicators are constructed simply by interacting county indicators with a recession indicator. Recession dates are taken from the NBER. Since the LODES data are annual, if a recession occurred for part of the year we assign a value of one for the entire year.

[10] An endogeneity test rejects the null that the $NewValue_{ct}$ variable is exogenous and the null hypothesis that our instrument is weak is also rejected.

$$\widehat{NewValue}_{ct} = \frac{(e^{\hat{\mu}_c + \hat{\kappa}_{pt}} - 1)}{pop_{c,2000}} \quad (3)$$

4.2 Identification

Our benchmark specification include county fixed effects that control for unobservable time-invariant differences in labor market outcomes that are specific to each county. The county fixed effects also account for the fact that some counties already have access to promising geologic formations and pre-existing infrastructure to extract the resources. The inclusion of year fixed effects account for nonlinear aggregate shocks to labor markets and oil and gas production that are common across counties. These parameters will net out aggregate shocks to the demand for oil and gas that are correlated with production and employment. The 2008-2009 financial crisis resulted in a recession that impacted all local economies but likely had a differential effect across local labor markets, especially those experiencing the fracking boom. These differential effects are captured by the county-specific recession parameters.

In the models estimated by least squares, identification of the causal effect of new oil and gas production on labor market outcomes therefore relies on conditional exogeneity. That is, there are no unobserved innovations in ε_{ct}, conditional on county, year and county-specific recessionary effects, that are correlated with the oil and gas production and labor market outcomes in counties. As noted above, this assumption is likely violated since a county's ability to extract the resources is likely endogenous to changes in local labor markets, including how quickly a local area can adapt to new infrastructure demand and employment needs (local factors that vary year to year and are not captured by the county fixed effects). Therefore, results from the least squares specifications should be received with some skepticism. It is for this reason we prefer the instrumental variables model and believe these results to be more reliable estimates of the causal effect than least squares.

5 Results

Before discussing each set of results individually, an overview of how the various tables and analyses are linked is useful. We are interested in the impact of new drilling operations on various measures of labor market outcomes. We first measure how earnings and employment rates are impacted by the value of new oil and gas production using both WAC (workplace) and RAC (residential) measures. Table 3 shows that production shocks created new jobs as identified by workplace employment (WAC) but residential employment (RAC) was unaffected – a first indication that jobs have been created, but many of those jobs may have gone workers who do not live locally.

Table 4 disaggregates the WAC and RAC data by NAICS sector. Here we find that employment in the mining sector (NAICS 21) has increased for locals and non-locals because there is a positive change in jobs as measured by both the workplace (WAC) and residential (RAC) data. Following the outcomes of Tables 3 and 4 that job outcomes are different for those that live in the area that drilling has taken place, we delve into the origin-destination (OD) files of the LODES data. Table 5 shows that new oil and gas production leads to higher employment for non-locals in the goods-producing and the trade, transportation, and utilities sectors. Here we also find that new oil and gas production causes a small decrease in employment for locals in the trade, transportation, and utilities sector.

Beginning with Table 6, we are able to classify job holders that are local residents or nonlocals and how far away their home county is located. The dependent variables measure fractions of local workplace employment and sum to one. Hence, these results identify how employment is reallocated from local to nonlocal employment and how far away those workers live. In Table 7 we discuss how the fraction of local and non-local jobs within certain earnings ranges have changed. We find that locally held jobs paying less than $1,250 per month have decreased while the amount of nonlocal jobs have increased in the $1,230-$3,333 earnings bin.

5.1 Discussion of Results

Table 3 presents our primary results estimating the models outlined in Equations (1)-(3). The dependent variables are listed across the columns. Panel A shows the parameter estimates and standard errors pertaining to $NewValuePerCapita_{ct}$ (in 1000's of dollars) estimated by OLS from Equation (1). Panel B shows the corresponding IV estimates where the instrument for $NewValuePerCapita_{ct}$ is obtained from the model in Equation (2) and Equation (3). Also shown in each panel is the implied change in jobs due to a one standard deviation increase in $NewValuePerCapita_{ct}$.[11] Also included in the table is an estimate for the implied dollar value of new production to create one job and the mean of the dependent variables.

In all cases the point estimates from the IV model are larger in absolute value and more precisely estimated than least squares, but the signs and significance are consistent across the panels. An increase in the value of new production per capita increases average monthly earnings by 14%. We also find that new production increases workplace employment, but has no effect on residential employment. This is the first indication that while new oil and gas production increases the number of jobs being offered in the county (workplace employment), those individuals who live outside the county are holding many of the new jobs. Another way to measure this effect is through the fraction of county jobs that are held by local residents versus the fraction that are held by workers who live outside the county. The dependent variable in column (4) measures the fraction of workplace area jobs that are held by local residents. The dependent variable in column (5) measures the fraction of workplace area jobs that are held by workers who live outside the county. These two variables sum to one, so the point estimates on $NewValuePerCapita_{ct}$ in these models should sum to zero. The

[11]The implied change in jobs is computed as $\hat{\beta} * SD(NewValuePerCapita_{ct}) * \bar{POP}$ when the denominator of the dependent variable is county population in the 2000 census and is computed as $\hat{\beta} * SD(NewValuePerCapita_{ct}) * Workforce\bar{Area}Employment$ in all other cases. The standard deviation, mean population and mean workforce area employment are computed conditional on counties exposed to Marcellus or Utica rather than including counties that do not produce any oil or gas.

point estimates in columns (4) and (5) suggest that increases in new oil and gas production shifts the distribution of jobs to workers who live outside the county. Focusing on the IV estimates in Panel B, a one standard deviation in the value of new production per capita increases workplace area employment by about 283 jobs, but a meaningful portion of the workforce is also moved to workers living outside the county. Note that the implied $/job in column (2) is $2.13 million. This number is in line with but larger than Feyrer et al. (2017) who report that 0.85 jobs cost $1 million in new production. We believe the driving difference in these estimates is due to the fact that their analysis pertains to shale plays across the United States (including many areas heavily producing oil) while our analysis of the Marcellus and Utica plays largely pertains to increases in natural gas extraction which has different profitability margins than oil production.

Table 3: Earnings and Employment Outcomes

	(1) Log Earnings	(2) Workplace Emp/Pop	(3) Residential Emp/Pop	(4) Fraction of Jobs Held Locally	(5) Fraction of Jobs Held Outside County
		Panel A: OLS			
New Value per Capita (1000s)	0.0695***	0.0133**	0.0096	-0.0331*	0.0331*
	(0.0139)	(0.0058)	(0.0099)	(0.0176)	(0.0176)
Mean of Dependent Variable	7.86	0.264	0.311	0.514	0.486
Implied Change in Jobs	–	+244	+176	-211	+211
Implied $ per Job	–	$2.48M	$3.44M	-$1.0M	$1.0M
N	2730	2730	2730	2730	2730
R^2	0.934	0.964	0.822	0.909	0.909
		Panel B: IV			
New Value per Capita (1000s)	0.1416***	0.0155***	0.0042	-0.0630**	0.0630**
	(0.0625)	(0.0023)	(0.0056)	(0.0246)	(0.0246)
Mean of Dependent Variable	7.86	0.264	0.311	0.514	0.486
Implied Change in Jobs	–	+283	+76	-403	+403
Implied $ per Job	–	$2.13M	$7.93M	-$0.52M	$0.52M
N	2730	2730	2730	2730	2730
R^2	0.932	0.964	0.822	0.908	0.908

Note: Dependent variables are indicated in the columns. Models include county fixed effects, year fixed effects and county-specific recession effects. New Value per Capita is measured as the value of new oil and gas production per capita in the twelve months prior to the reference period (beginning of Q2 each year) measured in 1000s of dollars. Implied Change in Jobs measures the estimated impact of a one standard deviation increase in New Value per Capita in a county-year based on the point estimate in the model. Implied $ per job computes the monetary value of new production required to create a single job based on the point estimate in the model. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.01.

In Table 4, workplace area employment and residential area employment are decomposed into employment in the oil and gas sector (NAICS 21) and employment in all other sectors. Increases in the value of new production increases employment in the oil and gas sector for both workplace employment (jobs provided by employers in the county) and residential employment (jobs held by a worker who lives in the county) by about the same amount. The finding that workplace employment grows in the mining sector is expected given the nature of the oil and gas industry and the fact that the some specialized workers may travel to shale plays across the United States to fulfill specific job duties as they are required in the time-line of the drilling process (for example 'mudding' and 'well completion'). These workers may work one month on a job site and then return home for a month before being brought back to the same location or taken to another.

Tables 3 and 4 both show that workplace employment rises, but do not show who or what type of worker obtains those new jobs. While the Residential Area Characteristics (RAC) and Workplace Area Characteristics (WAC) files provide greater industry detail than the Origin-Destination (OD) files, the RAC and WAC files lack the detail to tie jobs to worker residence. The OD files identify both where jobs are located and where workers live, but with great aggregation across industries. We now delve deeper into the question of who exactly is impacted by the shale boom and use statistics from the OD files to compute the employment per capita held locally and non-locally for the following industry groups available in the LODES OD files: Goods Producing; Trade, Transportation and Utilities; and Other Services. These six variables are used as dependent variables in Table 5.[12] The least squares models in Panel A suggest small increases in employment across most industry groups for local workers and increased employment only in the goods producing sectors for non-local workers. However, the IV estimates in Panel B yield a different conclusion. The Goods Producing and the Trade, Transportation and Utilities sectors experience increased employment due to new oil and gas production, but again we find that these jobs are held by non-local workers.

[12] Note that these six variables do not sum to one since the denominator is population and therefore the points estimates will not add to zero. Also, using county population is appropriate for measures of local employment but less appropriate for non-local employment.

Our results to this point have shown a consistent pattern that employment growth from new oil and gas production largely goes to those living outside the county where the oil and gas is produced. In Table 6, we disaggregate employment held by non-local workers by the distance between the county where they work and the county where they live. The numerators of the dependent variables are the count of workplace area jobs based on the distance from the county where the worker lives, and the denominator in each case is the total local workplace area employment. Therefore, the dependent variables sum to one and the point estimates should sum to zero. Column (1) in Table 6 simply reproduces column (4) in Table 3 for comparison purposes, where the outcome is the fraction of employment that is held locally. Here we see that most of those local jobs are reallocated to workers living 25-50 miles away (column 3) or 125-200 miles away (column 5). It is possible that some of those nonlocal workers live in a neighboring county or in the same commuting zone. In the robustness section, we estimate models where the data has been aggregated to commuting zones (BLS definitions of Labor Market Areas).

Table 4: Employment Outcomes by Sector

	(1) Workplace Emp/Pop in NAICS 21	(2) Workplace Emp/Pop in All Other Sectors	(3) Residential Emp/Pop in NAICS 21	(4) Residential Emp/Pop in All Other Sectors
	Panel A: OLS			
New Value per Capita (1000s)	0.0080***	0.0054	0.0054***	0.0042
	(0.0033)	(0.0035)	(0.0014)	(0.0108)
Mean of Dependent Variable	0.009	0.254	0.009	0.302
Implied Change in Jobs	+146	+98	+99	+77
Implied $ per Job	$4.15M	$6.17M	$6.13M	$7.85M
N	2730	2730	2730	2730
R^2	0.910	0.968	0.920	0.853
	Panel B: IV			
New Value per Capita (1000s)	0.0156***	-0.0001	0.0148***	-0.0106
	(0.0052)	(0.0119)	(0.0029)	(0.0242)
Mean of Dependent Variable	0.009	0.254	0.009	0.302
Implied Change in Jobs	+286	-3	+270	-194
Implied $ per Job	$2.11M	-$2.43M	$2.23M	-$3.11M
N	2730	2730	2730	2730
R^2	0.906	0.968	0.905	0.852

Note: Dependent variables are indicated in the columns. NAICS 21 includes mining, quarrying and oil and gas exploration. Models include county fixed effects, year fixed effects and county-specific recession effects. New Value per Capita is measured as the value of new oil and gas production per capita in the twelve months prior to the reference period (beginning of Q2 each year) measured in 1000s of dollars. Implied Change in Jobs measures the estimated impact of a one standard deviation increase in New Value per Capita in a county-year based on the point estimate in the model. Implied $ per job computes the monetary value of new production required to create a single job based on the point estimate in the model. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.01.

Table 5: Employment Outcomes by Industry Group and Worker Residence

	(1) Local Emp/Pop In Goods Producing Industries	(2) Local Emp/Pop in Trade, Transportation & Utilities	(3) Local Emp/Pop in Other Services Industries	(4) Nonlocal Emp/Pop in Goods Producing Industries	(5) Nonlocal Emp/Pop in Trade, Transportation & Utilities	(6) Nonlocal Emp/Pop in Other Services Industries
Panel A: OLS						
New Value per Capita (1000s)	0.0028**	0.0010*	0.0006	0.0087**	0.0016	-0.0013
	(0.0011)	(0.0006)	(0.0008)	(0.0036)	(0.0014)	(0.0017)
Mean of Dependent Variable	0.039	0.027	0.070	0.037	0.035	0.057
Implied Change in Jobs	+51	+18	+11	+158	+29	-23
Implied $ per Job	$11.9M	$34.1M	$55.4M	$3.8M	$20.7M	-$26.4M
N	2730	2730	2730	2730	2730	2730
R^2	0.947	0.933	0.975	0.905	0.910	0.949
Panel B: IV						
New Value per Capita (1000s)	0.0034	-0.0043*	-0.0014	0.0163**	0.0082**	-0.0068
	(0.0032)	(0.0022)	(0.0053)	(0.0068)	(0.0036)	(0.0081)
Mean of Dependent Variable	0.039	0.027	0.070	0.037	0.035	0.057
Implied Change in Jobs	+62	-78	-25	+298	+149	-124
Implied $ per Job	$9.7M	-$7.8M	-$24.3M	$2.03M	$4.05M	-$4.9M
N	2730	2730	2730	2730	2730	2730
R^2	0.947	0.930	0.975	0.904	0.908	0.948

Note: Dependent variables are indicated in the columns. Local refers to jobs held locally by residents living inside the county. Nonlocal refers to jobs held by individuals who live outside the county. Models include county fixed effects, year fixed effects and county-specific recession effects. New Value per Capita is measured as the value of new oil and gas production per capita in the twelve months prior to the reference period (beginning of Q2 each year) measured in 1000s of dollars. Implied Change in Jobs measures the estimated impact of a one standard deviation increase in New Value per Capita in a county-year based on the point estimate in the model. Implied $ per job computes the monetary value of new production required to create a single job based on the point estimate in the model. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.01.

Table 6: Reallocation of Workplace Area Employment by Distance from County

	(1) % of Jobs Held Locally	(2) % of Jobs Held within 25 Miles	(3) % of Jobs Held within 25 to 50 Miles	(4) % of Jobs Held within 50 to 125 Miles	(5) % of Jobs Held within 125 to 200 Miles	(6) % of Jobs Held Greater Than 200 Miles
Panel A: OLS						
New Value per Capita (1000s)	-0.0331*	0.0103	0.0098	0.0093**	0.0024	0.0014
	(0.0175)	(0.0116)	(0.0074)	(0.0044)	(0.0021)	(0.0010)
Mean of Dependent Variable	0.516	0.166	0.178	0.110	0.025	0.005
Implied Change in Jobs	-211	+66	+62	+59	+15	+9
Implied $ per Job	-$1.0M	$3.2M	$3.4M	$3.6M	$13.9M	$24.1M
N	2730	2730	2730	2730	2730	2730
R^2	0.909	0.973	0.933	0.837	0.838	0.827
Panel B: IV						
New Value per Capita (1000s)	-0.0628**	-0.0163	0.0390*	0.0305	0.0113***	-0.0018
	(0.0249)	(0.0161)	(0.0228)	(0.0191)	(0.0041)	(0.0022)
Mean of Dependent Variable	0.516	0.166	0.178	0.110	0.025	0.005
Implied Change in Jobs	-400	-104	+249	+195	+72	-11
Implied $ per Job	-$0.53M	-$2.03M	$0.85M	$1.1M	$2.9M	-$18.9M
N	2730	2730	2730	2730	2730	2730
R^2	0.908	0.972	0.930	0.833	0.832	0.823

Note: Dependent variables are indicated in the columns. Models include county fixed effects, year fixed effects and county-specific recession effects. New Value per Capita is measured as the value of new oil and gas production per capita in the twelve months prior to the reference period (beginning of Q2 each year) measured in 1000s of dollars. Implied Change in Jobs measures the estimated impact of a one standard deviation increase in New Value per Capita in a county-year based on the point estimate in the model. Implied $ per job computes the monetary value of new production required to create a single job based on the point estimate in the model. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.01.

Finally, Table 7 investigates the impact of new oil and gas production on the distribution of jobs classified by monthly earnings and whether the worker lives locally or outside the county. Focusing on the IV estimates in Panel B, it appears that the distribution of jobs for both local and nonlocal workers shifts to higher paying jobs generally. However, the shift in earnings increases manifests itself differently for locals and nonlocals. For locals, average earnings increases appear to be due to a smaller fraction of lower paying jobs. Although we cannot statistically distinguish which category of jobs increases for locals as a separate result, the point estimate for "between $1250/month and $3333/month" is positive (column (2)), while the point estimate for the highest earning category is negative (column (3)). For nonlocals, we see a statistically significant increase in the fraction of jobs paying a mid-level salary (column (5)). These results combined provide some evidence that new, higher paying jobs are more likely going to nonlocals rather than local residents in the county where the production occurs.

Table 7: Reallocation of Jobs by Monthly Earnings and Worker Residence

	(1) Fraction of Local Jobs <$1250/month	(2) Fraction of Local Jobs $1250/month to $3333/month	(3) Fraction of Local Jobs >$3333/month	(4) Fraction of Nonlocal Jobs <$1250/month	(5) Fraction of Nonlocal Jobs $1250/month to $3333/month	(6) Fraction of Nonlocal Jobs >$3333/month
Panel A: OLS						
New Value per Capita (1000s)	-0.0251***	-0.0116	0.0035	0.0052	0.0071	0.0209**
	(0.0060)	(0.0075)	(0.0086)	(0.0071)	(0.0082)	(0.0087)
Mean of Dependent Variable	0.164	0.232	0.120	0.135	0.197	0.153
Implied Change in Jobs	-160	-74	+23	+33	+45	+133
Implied $ per Job	-$1.3M	-$2.9M	$9.3M	$6.4M	$4.7M	-$1.6M
N	2730	2730	2730	2730	2730	2730
R^2	0.871	0.915	0.898	0.833	0.846	0.917
Panel B: IV						
New Value per Capita (1000s)	-0.0598***	0.0221	-0.0251	-0.0174	0.0578*	0.0224
	(0.0219)	(0.0144)	(0.0197)	(0.0157)	(0.0324)	(0.0202)
Mean of Dependent Variable	0.164	0.232	0.120	0.135	0.197	0.153
Implied Change in Jobs	-381	+140	-160	-111	+369	+143
Implied $ per Job	-$0.55M	$1.5M	-$1.3M	-$1.9M	$0.57M	$1.5supM
N	2730	2730	2730	2730	2730	2730
R^2	0.864	0.911	0.892	0.827	0.831	0.917

Note: Dependent variables are indicated in the columns. Models include county fixed effects, year fixed effects and county-specific recession effects. New Value per Capita is measured as the value of new oil and gas production per capita in the twelve months prior to the reference period (beginning of Q2 each year) measured in 1000s of dollars. Implied Change in Jobs measures the estimated impact of a one standard deviation increase in New Value per Capita in a county-year based on the point estimate in the model. Implied $ per job computes the monetary value of new production required to create a single job based on the point estimate in the model. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.01.

5.2 Robustness

The geographical unit of our analyses thus far has been a county. Since we are considering worker migration and job spillovers, using a political boundary to define a local labor market may miss some ways in which bordering counties share workers and jobs. The Bureau of Labor Statistics uses commuting information from the American Community Survey to group counties into labor market areas (LMAs). We use the crosswalk provided by the BLS to aggregate our oil and gas production data and the LODES data to distinct LMAs.[13]

Our data has 210 counties across the three states which are combined into 157 LMAs. We use the same models outlined in Equations (1)-(3) to estimate a subset of our main results using the LMA-level data. These results are found in appendix Table A1. The same patterns emerge for LMAs as with counties: workplace employment increases, residential employment does not; employment in the oil and gas sections increases for both; there are no clear employment spillovers in other industries; and local workplace employment is allocated to individuals who do not live within the LMA.

Next, our preferred panel data specification includes county fixed-effects, year fixed-effects and county-specific recession effects. Table A2 in the appendix illustrates the sensitivity of our main results to the inclusion or exclusion of these indicators as well as the inclusion of county-specific linear trends. This table is structured differently than the others to more clearly display the model differences. The dependent variables are listed in the far left rows while each column represents an alternative specification. The entries are the point estimate for $NewValue$ and clustered standard errors in parentheses.

The county-specific trends largely wipe out the statistical significance and even changes the sign of the point estimate in some cases. This is a common result in the minimum wage literature regarding the determinants of employment rates and has generated substantial debate

[13] The crosswalk can be found at https://www.bls.gov/lau/lmadir2015.xlsx.

as to whether including trends is appropriate (e.g., Neumark et al. (2014) and Allegretto et al. (2017)). Employment rates tend to be relatively persistent and the debate centers around whether there is appropriate variation remaining after stripping out the trends to achieve identification. Our panel data models likely suffer from the same problem. Even without the county trends, our models are often able to explain well over 90% of the variation in the dependent variable with even less remaining variation with the inclusion of trends. Otherwise, the models appear to be robust to the omission of year and county-recession effects with the exception of the Workplace Emp/Pop coefficient in column (3) which loses some precision while maintaining a similar magnitude.

6 Conclusion

In this paper we find that the advent and spread of hydraulic fracturing in the Marcellus and Utica formations has had a strong, and in some ways surprising, impact on labor markets in Pennsylvania, Ohio, and West Virginia. While others have studied labor market impacts in the Marcellus and Utica regions, our analysis is unique in its ability to track worker flows across geopolitical boundaries. Our primary data for local labor market outcomes is the LEHD Origin-Destination Employment Statistics (LODES) data produced by the Longitudinal Employer-Household Dynamics Program within the U.S. Census Bureau. The LODES data provide employment statistics at the census block level based on where workers live (origin) and where jobs are located (destination). Therefore, we can compute employment statistics based on place of work, place of residence, and employment statistics derived from information on both where the job is located and where the worker lives.

To achieve identification and study local labor market impacts and flows we aggregate the block level employment statistics from LODES to the county level and merge on oil and gas production data in each county. Using monthly average prices of oil and gas, we compute the value of new production per capita for each reference period. Using this measure of resource extraction, we estimate panel data models using both least squares and instrumental variables to identify the causal effect of new oil and gas production on local labor market outcomes in counties for both local residents and workers who live outside the county.

We find that increases in the value of new oil and gas production in a county increase average monthly earnings and aggregate employment, but nearly all of these gains go to individuals who work in the county but reside outside the county. Specifically, a one standard deviation increase in new production per capita increases workplace employment in a county-year by 283 jobs, but local 'residential' aggregate employment is unaffected. Among jobs created, new oil and gas production significantly shifts new jobs to workers who live outside the county where the drilling takes place, and most of the shift in employment is toward workers who live in a county more than 25 miles away, not in bordering counties. This finding is robust to how we define local versus non-local employment. When we account for counties that may be connected via typical commuting patterns (Labor Market Areas) we still find that increases in the value of new production per capita shifts employment patterns such that a larger fraction of all employment in the area is from non-local residents. While the evidence points to the fact that most gains from the resource boom have gone to non-residents, we do find evidence that local residents experience earnings increases and also share in the new jobs created in the oil and gas industry. This appears to represent a rightward shift in the earnings distributions for both workers who live locally and workers living outside the county as we see a decrease in the fraction of locally held jobs that pay $1,250 per month or less and an increase in the amount of non-locally held jobs that pay $1,250-3,333 per month.

The generalizability of our results to other resource rich areas is not immediately apparent. We consider the Marcellus-Utica region to be a great laboratory to study a sudden resource boom in modern economies. Prior to the shale boom, practically no oil and gas production occurred in the region. With the invention of hydraulic fracturing un-tapped reserves were, for all intents and purposes, created overnight. In this way, the Marcellus-Utica experience is akin to a natural experiment on the effects of a resource boom in a modern, developed economy. To the extent that an area does not have prior experience or a developed infrastructure of capital in extracting or producing the resource that is 'booming' (physical or human), we believe our results hold interesting insights. As Munasib and Rickman (2015) show, the shale boom has had differential effects across the United States. However, using a national sample Feyrer et al. (2017) find that new oil and gas production has had positive wage impacts, and further find that these changes are significant in a 100 mile radius. Our study refines their findings using a data source that is able to accurately match where new jobs are located with those that hold the new jobs. Using this origin-destination data we, too, find that earnings have increased and that jobs have been created as a result of new drilling activity, and that there are still positive spillover effects more than 100 miles away. Moreover, our results are supported by Wilson (2016) who finds that in-migration occurred in North Dakota as a result of the shale boom.

Our results are important for policy makers at all levels of governance as this new surge of economic activity has the potential to effect local tax receipts, local economic development, and employment shifts across industries and geopolitical boundaries. Our findings should temper expectations regarding local job growth through potential tax exceptions or subsidies provided by local economic development authorities intended to increase oil and gas operations. While we do find that an increase in oil and gas operations increases aggregate earnings and employment in a county, we also find that the majority of this boon goes to non-local (out of county) workers. While tax exemptions should be given a second thought, so too should the effect of banning or restricting hydraulic fracturing. An example in which cross-border differences in regulation may come to fruition is along the New York and Pennsylvania border. A moratorium on hydraulic fracturing has been imposed by the state of New York limiting the areas exposed to the Marcellus formation to extract those resources.

Given our results, there is likely a non-trivial flow of workers from New York into Pennsylvania not only in the oil and gas industry but in to other industries as well. Additionally, prior research has shown that in-migration of new workers has lead to changes in crime rates and even sexually transmitted disease (James and Smith (2017); Street (2018); Komarek and Cseh (2017)). Our findings that most new jobs are going to those that live outside of the county and are travelling more than 25 miles point to a potential discrepancy in local income tax receipts and the ability to afford greater policing or public health infrastructure that may be needed due to these negative externalities of increased drilling activities.

A Appendix

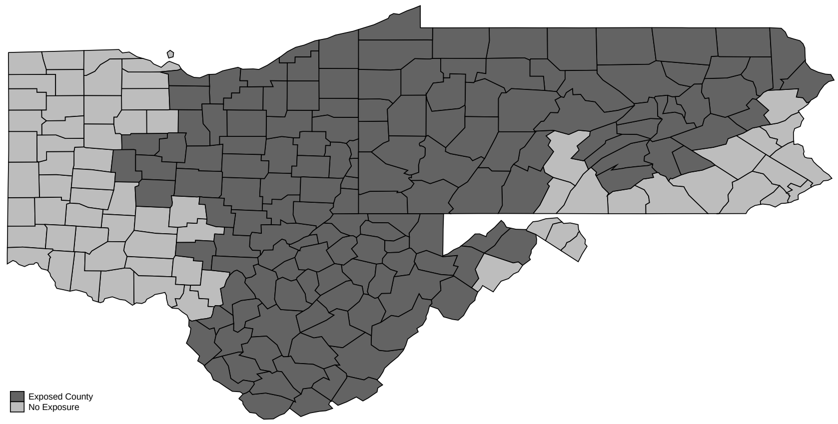

Figure A1: Figure shows county exposure to either Marcellus or Utica plays in Ohio, West Virginia and Pennsylvania.

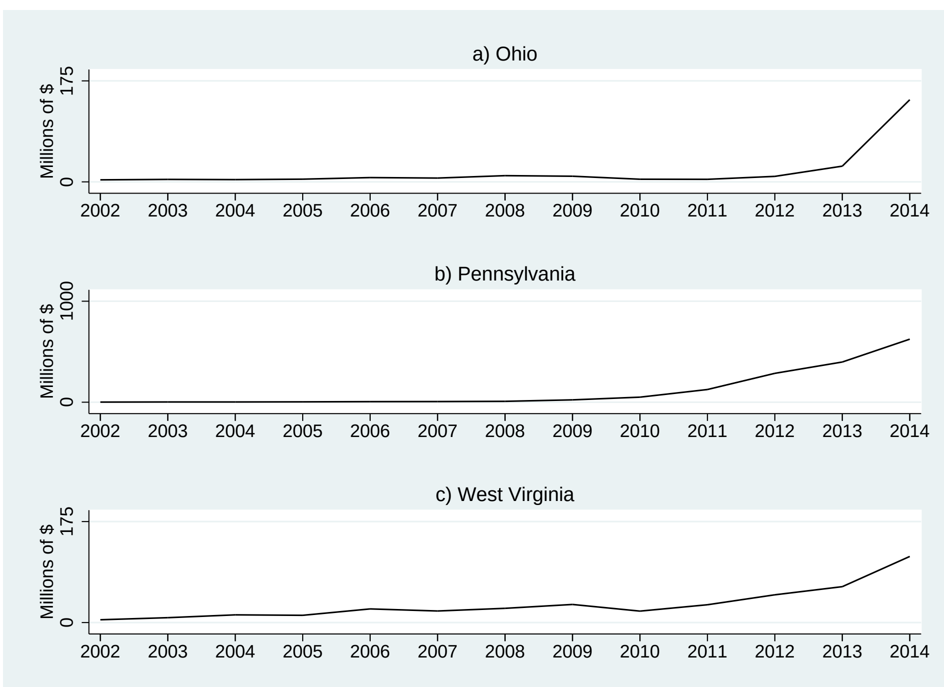

Figure A2: Total Value of New Production by State. Trailing 12 month value of new well production as of April 1 in each year.

Table A1: Main Results using Labor Market Areas

	Workplace Emp/Pop	Residential Emp/Pop	Workplace Emp/Pop NAICS 21	Workplace Emp/Pop not NAICS 21	Residential Emp/Pop NAICS 21	Residential Emp/Pop not NAICS 21	Fraction Jobs Held Locally	Fraction Jobs Held Nonlocally
Panel A: OLS								
New Value per Capita	0.0175**	0.0204***	0.0092**	0.0083*	0.0055***	0.0148*	-0.0191	0.0191
(1000s)	(0.0085)	(0.0066)	(0.0042)	(0.0050)	(0.0020)	(0.0081)	(0.0137)	(0.0137)
Mean of Dep Var	0.260	0.302	0.010	0.251	0.009	0.293	0.530	0.470
Implied $\Delta in Jobs$	+500	+140	+226	+203	+136	+364	-163	+163
Implied $ per Job	$1.9M	$1.6M	$3.6M	$4.0M	$6.0M	$2.2M	-$1.7M	$1.7M
N	2041	2041	2041	2041	2041	2041	2041	2041
R^2	0.955	0.802	0.907	0.961	0.912	0.837	0.898	0.898
Panel A: IV								
Value per Capita	0.0151**	0.0005	0.0128***	0.0023	0.0143***	-0.0138	-0.0560***	0.0560***
(1000s)	(0.0090)	(0.0209)	(0.0038)	(0.0091)	(0.0020)	(0.0200)	(0.0206)	(0.0206)
Mean of Dep Var	0.260	0.302	0.010	0.251	0.009	0.293	0.530	0.470
Implied $\Delta in Jobs$	+429	+500	+226	+203	+136	+364	-163	+163
Implied $ per Job	$1.9M	$1.6M	$3.6M	$4.0M	$6.0M	$2.2M	-$1.7M	$1.7M
N	2041	2041	2041	2041	2041	2041	2041	2041
R^2	0.955	0.801	0.906	0.961	0.899	0.834	0.895	0.895

Note: Dependent variables are indicated in the columns. Data are aggregated from counties to labor market areas (LMAs) according to the BLS crosswalk. Models include LMA fixed effects, year fixed effects and LMA-specific recession effects. New Value per Capita is measured as the value of new oil and gas production per capita in the twelve months prior to the reference period (beginning of Q2 each year) measured in 1000s of dollars. Implied Change in Jobs measures the estimated impact of a one standard deviation increase in New Value per Capita in a county-year based on the point estimate in the model. Implied $ per job computes the monetary value of new production required to create a single job based on the point estimate in the model. Standard errors are clustered by LMA. * p-value<0.10, ** p-value<0.05, *** p-value< 0.01.

Table A2: Sensitivity Analysis

	(1)	(2)	(3)	(4)	(5)
Workplace Emp/Pop	0.0155***	0.0266**	0.0169	-0.0096	-0.0065
	(0.0023)	(0.0124)	(0.0141)	(0.0216)	(0.0117)
Residential Emp/Pop	0.0042	0.0382	0.0082	0.0007	-0.0130
	(0.0256)	(0.0300)	(0.0301)	(0.0266)	(0.0265)
Workplace NAICS 21	0.0156***	0.0202***	0.0176**	0.0147	0.0096*
	(0.0052)	(0.0070)	(0.0076)	(0.0165)	(0.0050)
Workplace not NAICS 21	-0.0001	0.0064	-0.0007	-0.0243	-0.0161
	(0.0119)	(0.0099)	(0.0128)	(0.0308)	(0.0126)
Residential NAICS 21	0.0148***	0.0195***	0.0173***	0.0154	0.0084***
	(0.0029)	(0.0050)	(0.0052)	(0.0121)	(0.0022)
Residential not NAICS 21	-0.0106	0.0187	-0.0091	-0.0147	-0.0214
	(0.0242)	(0.0271)	(0.0278)	(0.0272)	(0.0257)
Fraction Jobs Local	-0.0628**	-0.1970***	-0.0545*	0.1022	0.0267
	(0.0249)	(0.0511)	(0.0282)	(0.1301)	(0.0288)
Fraction Jobs not Local	0.0628**	0.1970***	0.0545*	-0.1022	-0.0267
	(0.0249)	(0.0511)	(0.0282)	(0.1301)	(0.0288)
County FE	Y	Y	Y	Y	Y
Year FE	Y	N	Y	Y	Y
County-Recession FE	Y	N	N	N	Y
County-Specific Trends	N	N	N	Y	Y
N	2730	2730	2730	2730	2730

Note: Dependent variables in the rows of the first column. Each column (1)-(5) pertains to a different model specification for unobservable factors. Entries in column (1) reflect the benchmark specification and results from earlier tables are reproduced here for comparison. Entries are the point estimate and standard error for New Value per Capita ($1000's) from the IV estimation outlined in the paper. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.01.

References

Agerton, M., Hartley, P. R., Medlock, K. B. and Temzelides, T. (2017). Employment impacts of upstream oil and gas investment in the United States, *Energy Economics* **62**(Supplement C): 171–180.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988316303589*

Alexeev, M. and Conrad, R. (2009). The elusive curse of oil, *Review of Economics and Statistics* **91**(3): 586–98.

Allcott, H. and Keniston, D. (2014). Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America, *Working Paper 20508*, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w20508*

Allegretto, S., Dube, A., Reich, M. and Zipperer, B. (2017). Credible research designs for minimum wage studies: A response to neumark, salas, and wascher, *Industrial and Labor Relations Review* **70**(3).

Betz, M. R., Partridge, M. D., Farren, M. and Lobao, L. (2015). Coal mining, economic development, and the natural resources curse, *Energy Economics* **50**: 105–116.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988315001279*

Blanco, L. and Grier, R. (2012). Natural resource dependence and the accumulation of physical and human capital in Latin America, *Resources Policy* **37**(3): 281–295.
URL: *http://www.sciencedirect.com/science/article/pii/S0301420712000062*

Boyce, J. R. and Herbert Emery, J. C. (2011). Is a negative correlation between resource abundance and growth sufficient evidence that there is a resource curse?, *Resources Policy* **36**(1): 1–13.
URL: *http://www.sciencedirect.com/science/article/pii/S0301420710000450*

Brown, J., Fitzgerald, T. and Weber, J. (2016). Asset Ownership, Windfalls, and Income: Evidence from Oil and Gas Royalties, *The Federal Reserve Bank of Kansas City Research Working Papers* .
URL: *https://www.kansascityfed.org/publications/research/rwp/articles/2016/asset-ownership-windfalls-income-oil-gas-royalties*

Fetzer, T. (2014). Fracking growth, *CEP Discussion Paper*.
URL: *http://cep.lse.ac.uk/pubs/download/dp1278.pdf*

Feyrer, J., Mansur, E. T. and Sacerdote, B. (2017). Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution, *American Economic Review* **107**(5).
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20151326*

Ginn, V. and Roach, T. (2015). An Oil-Producing States Ability to Cope after a Regional Free Trade AgreementThe Case of Texas and NAFTA, *The International Trade Journal*

29(4): 309–336.
URL: *http://dx.doi.org/10.1080/08853908.2015.1054968*

Gourley, P. and Madonia, G. (2016). *Resource Booms and Crime.* DOI: 10.13140/RG.2.2.19586.68805.

Green, D., Morissette, R. and Sand, B. M. (2017). Economy Wide Spillovers From Booms: Long Distance Commuting and the Spread of Wage Effects, *Technical report*, Vancouver School of Economics.

Haggerty, J., Gude, P. H., Delorey, M. and Rasker, R. (2014). Long-term effects of income specialization in oil and gas extraction: The U.S. West, 19802011, *Energy Economics* **45**(Supplement C): 186–195.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988314001534*

James, A. and Smith, B. (2017). There will be blood: Crime rates in shale-rich U.S. counties, *Journal of Environmental Economics and Management* **84**: 125–152.
URL: *http://www.sciencedirect.com/science/article/pii/S0095069616305459*

Kearney, M. S. and Wilson, R. (2018). Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom, *The Review of Economics and Statistics* **100**(4): 678–690.
URL: *https://doi.org/10.1162/rest*$_{a0}$0739

Kim, D.-H. and Lin, S.-C. (2015). Natural Resources and Economic Development: New Panel Evidence, *Environmental and Resource Economics* pp. 1–29.
URL: *http://link.springer.com/article/10.1007/s10640-015-9954-5*

Komarek, T. and Cseh, A. (2017). Fracking and public health: Evidence from gonorrheaincidence in the Marcellus Shale region, *Journal of Public Health Policy* **38**(4): 464–481.
URL: *https://doi.org/10.1057/s41271-017-0089-5*

Komarek, T. M. (2016). Labor market dynamics and the unconventional natural gas boom: Evidence from the Marcellus region, *Resource and Energy Economics* **45**: 1–17.
URL: *http://www.sciencedirect.com/science/article/pii/S092876551630077X*

Komarek, T. M. (2018). Crime and natural resource booms: evidence from unconventional natural gas production, *The Annals of Regional Science* **61**(1): 113–137.
URL: *https://doi.org/10.1007/s00168-018-0861-x*

Michaels, G. (2011). The Long Term Consequences of Resource-Based Specialisation*, *The Economic Journal* **121**(551): 31–57.
URL: *http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0297.2010.02402.x/abstract*

Miljkovic, D. and Ripplinger, D. (2016). Labor market impacts of U.S. tight oil development: The case of the Bakken, *Energy Economics* **60**(Supplement C): 306–312.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988316302882*

Munasib, A. and Rickman, D. S. (2015). Regional economic impacts of the shale gas and tight

oil boom: A synthetic control analysis, *Regional Science and Urban Economics* **50**: 1–17.
URL: *http://www.sciencedirect.com/science/article/pii/S0166046214001094*

Murshed, S. M. and Serino, L. A. (2011). The pattern of specialization and economic growth: The resource curse hypothesis revisited, *Structural Change and Economic Dynamics* **22**(2): 151–161.
URL: *http://www.sciencedirect.com/science/article/pii/S0954349X10000858*

Neumark, D., Salas, I. and Wascher, W. (2014). Revisiting the minimum wageemployment debate: Throwing out the baby with the bathwater?, *Industrial and Labor Relations Review* **67**(3).

Paredes, D., Komarek, T. and Loveridge, S. (2015). Income and employment effects of shale gas extraction windfalls: Evidence from the Marcellus region, *Energy Economics* **47**: 112–120.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988314002540*

Rickman, D. S., Wang, H. and Winters, J. V. (2017). Is shale development drilling holes in the human capital pipeline?, *Energy Economics* **62**(Supplement C): 283–290.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988316303590*

Sachs, J. D. and Warner, A. M. (2001). The curse of natural resources, *European Economic Review* **45**(46): 827–838.
URL: *http://www.sciencedirect.com/science/article/pii/S0014292101001258*

Sala-i Martin, X. and Subramanian, A. (2013). Addressing the Natural Resource Curse: An Illustration from Nigeria, *Journal of African Economies* **22**(4): 570–615.
URL: *http://jae.oxfordjournals.org/content/22/4/570*

Street, B. (2018). The Impact of Economic Opportunity on Criminal Behavior: Evidence from the Fracking Boom, *Working Paper* .

Tsvetkova, A. and Partridge, M. (2017). The shale revolution and entrepreneurship: An assessment of the relationship between energy sector expansion and small business entrepreneurship in US counties, *Energy* **141**: 423–434.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544217316201*

Tsvetkova, A. and Partridge, M. D. (2016). Economics of modern energy boomtowns: Do oil and gas shocks differ from shocks in the rest of the economy?, *Energy Economics* **59**: 81–95.
URL: *http://www.sciencedirect.com/science/article/pii/S0140988316301876*

Weber, J. G. (2014). A decade of natural gas development: The makings of a resource curse?, *Resource and Energy Economics* **37**: 168–183.
URL: *http://www.sciencedirect.com/science/article/pii/S0928765513000882*

Wilson, R. (2016). Moving to Economic Opportunity: The Migration Response to the Fracking Boom, *SSRN Scholarly Paper ID 2814147*, Social Science Research Network,

Rochester, NY.
URL: *https://papers.ssrn.com/abstract=2814147*