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

April 30, 2018

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 employment by about 629 jobs, but about 500 of these new jobs are held by individuals living outside the county. However, we find that the earnings distribution shifts to the right for both local residents and workers who reside outside the county. Furthermore, we find some positive employment spillovers across industries but these new jobs appear to go to nonlocal residents as well.

JEL Codes: J21, J61, Q33, Q40

Keywords: Local Labor Markets, Worker Mobility, Resource Boom

1 Introduction

'Boomtowns' are a trope that can be found across history, literature, 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 prices sway in such a way that inhibits further drilling or the local economy cools down once the resource has been depleted. Moreover, the shifting of employment and compensation towards the mining sector and supporting industries has the potential to disturb employment in sectors producing goods that are not directly tied to resource extraction.

Although the impact 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 is 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). Ours is among the first studies to explain employment migration across borders as a result of a resource boom.

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 or 'Dutch disease'. 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. For example, Jacobsen and Parker (2016) show that the gains in employment in the manufacturing sector during the oil boom of the 1970's were small compared to the negative impact that that the oil bust of the 1980's had in the manufacturing sector. 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. Murshed and Serino (2011) consider poor industry diversification and export structures as the causal mechanism. 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 the longer that states specialized in oil and gas production they had higher crime rates, lower per capita income, and lower educational attainment. Gourley and Madonia (2016) also show resource development is associated with higher rates of both violent and property crime.

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) finds that a large resource endowment has been, on net, positive for long-term economic growth. Boyce and Herbert Emery (2011) explores 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. We find that increases in the value of new production increases employment in the oil and gas sector, but that there are similar job gains seen in other sectors as well suggesting that there are contemporaneous, positive employment spillovers.

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. 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.

Other researchers have also focused on the drilling increases 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 50 miles away. Our IV estimates indicate that a one standard deviation increase in the value of new drilling per capita implies an increase of about 629 jobs in a county-year but a large majority of those jobs (502) are held by workers living outside the county. However, average monthly earnings increases in the county and this appears to represent a rightward shift in the earnings distributions for both workers who live locally and workers living outside the county.

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

[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.

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 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

[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.

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. 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

Note that the earnings bins are predefined by LEHD and there is no additional data provided to compute statistics for other earnings bins. 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
- 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 and (3) counties that are more than 50 miles away.[4]

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 three distance categories. For estimation, the job count statistics from the OD, WAC and RAC files will be expresses as rates where the denominator is either total county population or total employment.

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

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

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	119.4
	(150.7)	(23.9)	(445.9)
Ohio	7.1	1.8	67.2
	(88.5)	(4.8)	(307.6)
Pennsylvania	32.7	1.4	180.7
	(207.7)	(3.5)	(578.5)
West Virginia	30.5	23.1	128.3
	(145.4)	(42.5)	(445.9)

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.

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 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 the reference year.

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 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 1 shows a map of counties that are exposed to either play. 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). Across all counties and years the value of new production was about $21.4 per person, $7.1 per person in Ohio, $32.7 per person in Pennsylvania and $30.5 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 2 displays the monthly average prices of gas and oil. While gas and oil prices both peaked around 2008, gas prices have largely declined since and oil prices recovered and then declined again significantly after 2014.[5] Despite the decline in gas prices, the total value of production in the area has increased substantially. Figures 3 and 4 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

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

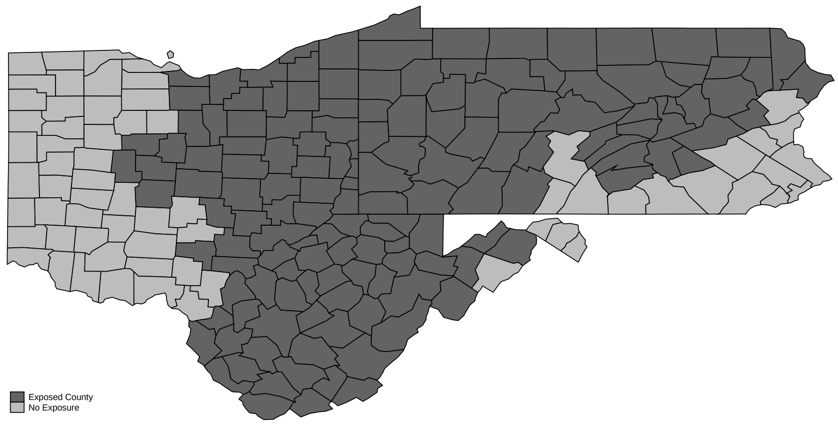

Figure 1: Figure shows county exposure to either Marcellus or Utica plays in Ohio, West Virginia and Pennsylvania.

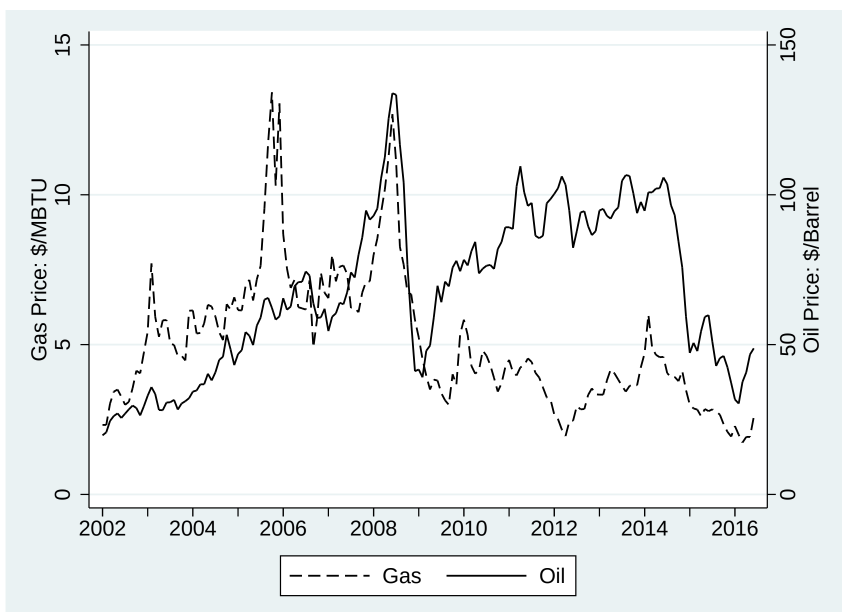

Figure 2: Monthly Oil and Gas Prices.

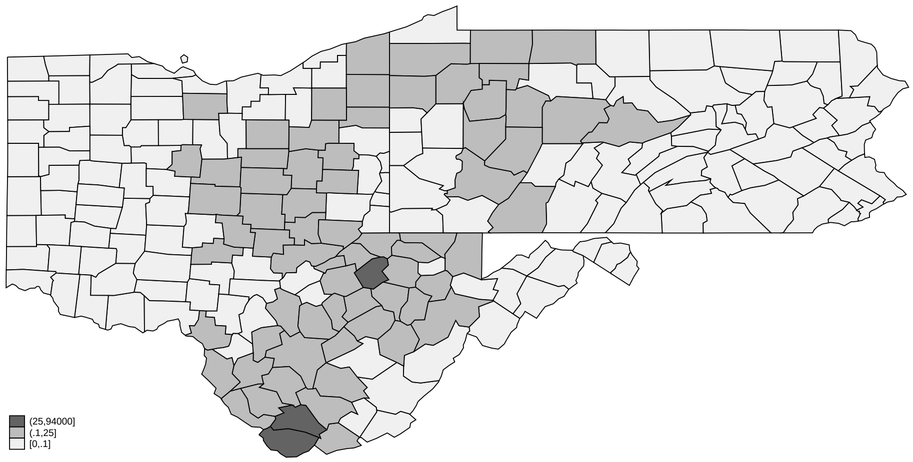

Figure 3: Bins reflect oil and gas production (in $) per capita in 2002 by county.

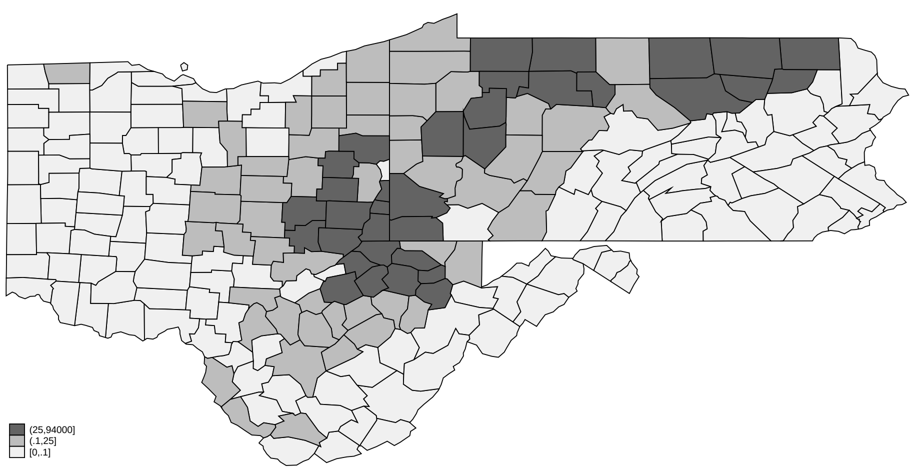

Figure 4: Bins reflect oil and gas production (in $) per capita in 2014 by county.

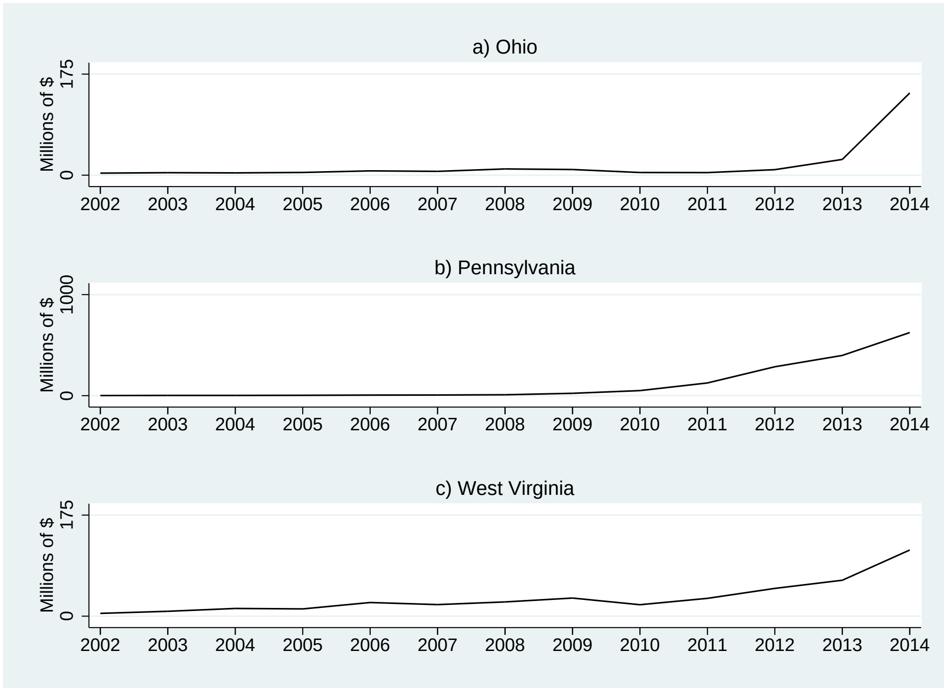

Figure 5: Total Value of New Production by State. Trailing 12 month value of new well production as of April 1 in each year.

the distribution per capita production (broken into thirds) over the entire sample period. While production is essentially zero in counties unexposed to Marcellus or Utica, within the exposed counties production varies significantly both over time and across space.

Figure 5 shows a time-series of the total value of new production by state. As indicated in panel A of Table 1, Ohio has less exposure to the plays and produced less than Pennsylvania and West Virginia. Pennsylvania has substantially greater production than either of the other two states. Also seen in Figure 5 is the difference in timing of new production that has taken place across states. Ohio's extraction began to increase much later than Pennsylvania or West Virginia. These differences in production across time and location also aid in identification of the model.

4 Empirical Model and Identification

We aim to estimate the impact of the value of new oil and gas production 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 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.

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 Miles)	0.140	(0.049)
Total County Population	123506.9	(212197.2)

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

4.1 Empirical Specification

Our benchmark specification is:

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

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 recession effects. The LODES are annual but reflect outcomes as of the second quarter of each year. 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. We follow Feyrer et al. (2017) and Fetzer (2014) to construct instruments using a two step process that addresses this potential endogeneity problem. 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 as in Equation (3):

$$New\widehat{Value}_{ct} = \frac{(e^{\hat{\mu}_c + \hat{\kappa}_{pt}} - 1)}{pop_{ct}} \tag{3}$$

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 of the causal effect than least squares.

5 Results

Table 3 presents the main results where 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 computation in Equation (3). Also shown in each panel is the implied change in jobs due to a one standard deviation increase in $NewValuePerCapita_{ct}$.[6] The implied change in jobs is only computed when point estimates are statistically different from zero.

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. A one-thousand dollar increase in the value of new production per capita increases average monthly earnings by 13%. 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

[6] 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 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.

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)	0.0697***	0.0175***	0.0134	-0.0326*	0.0326*
	(0.0137)	(0.0070)	(0.0102)	(0.0176)	(0.0176)
Implied Change in Jobs	–	+320	–	-208	+208
N	2730	2730	2730	2730	2730
R^2	0.934	0.969	0.812	0.909	0.909
			Panel B: IV		
New Value per Capita (1000s)	0.1303***	0.0345***	0.0198	-0.0564***	0.0564***
	(0.0574)	(0.0112)	(0.0204)	(0.0216)	(0.0216)
Implied Change in Jobs	–	+629	–	-359	+359
N	2730	2730	2730	2730	2730
R^2	0.933	0.968	0.812	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. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.025.

should sum to zero. The 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 629 jobs,[7] but more than half the increase in jobs is reallocated to workers living outside the county.

[7]For example, to compute the value 629, $\hat{\beta}$=0.0345, $SD(NewValuePerCapita_{ct}(1000s))$=0.181, and $P\bar{O}P$ = 100808.13, where the mean and standard deviation is computed only for counties exposed to the shale plays.

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. In this model we again find evidence that there are stronger gains in non-local employment since we see an increase of approximately 347 jobs in all other sectors in the workplace employment category but no statistically discernible gains in residential employment. Moreover, the point estimates indicate slightly greater workplace employment growth in sectors other than the oil and gas sector indicating relatively strong spillover effects from the new production.

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)	0.0084***	0.0091**	0.0054***	0.0080
	(0.0033)	(0.0041)	(0.0014)	(0.0109)
Implied Change in Jobs	+153	+167	+198	–
N	2730	2730	2730	2730
R^2	0.907	0.974	0.914	0.848
	Panel B: IV			
New Value per Capita (1000s)	0.0155***	0.0190*	0.0149***	0.0049
	(0.0055)	(0.0098)	(0.0028)	(0.0188)
Implied Change in Jobs	+282	+347	+271	–
N	2730	2730	2730	2730
R^2	0.904	0.974	0.899	0.848

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. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.025.

The Residential Area Characteristics (RAC) and Workplace Area Characteristics (WAC) files from the LODES data provide greater industry detail than the Origin-Destination (OD) files but 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 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.[8] 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.

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. The dependent variable in column (2) is the fraction of jobs held by workers who live outside the county but within 25 miles. The dependent variable in columns (3) and (4) are constructed similarly. Column (5) of Table 3 indicates that, of the jobs reallocated outside the county (column (4) of Table 3), most of the jobs are held by individuals living in counties at least 25 miles away from the county where the job is located. Indeed, more than 500 new jobs are held by employees that reside 25 miles or more from the county the new job is held. This suggests that few of the new nonlocal employees live in areas that closely border the county where the job is located, and that these workers are commuting a fairly long distance.

[8] Note that these six variables do not sum to one 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.

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)	0.0031***	0.0012**	0.0013	0.0096***	0.0022	0.0002
	(0.0011)	(0.0006)	(0.0008)	(0.0041)	(0.0016)	(0.0019)
Implied Change in Jobs	+57	+22	–	+175	–	–
N	2730	2730	2730	2730	2730	2730
R^2	0.946	0.936	0.979	0.895	0.909	0.955
Panel B: IV						
New Value per Capita (1000s)	0.0046	-0.0030	0.0019	0.0195***	0.0113***	0.0001
	(0.0030)	(0.0020)	(0.0044)	(0.0068)	(0.0039)	(0.0072)
Implied Change in Jobs	–	–	–	+355	+206	–
N	2730	2730	2730	2730	2730	2730
R^2	0.946	0.934	0.979	0.893	0.905	0.955

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. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value< 0.025.

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

	(1) Fraction of Jobs Held Locally	(2) Fraction of Jobs Held within 25 Miles	(3) Fraction of Jobs Held within 25 to 50 Miles	(4) Fraction of Jobs Held Greater Than 50 Miles
	Panel A: OLS			
New Value per Capita (1000s)	-0.0326*	0.0091	0.0097	0.0138**
	(0.0176)	(0.0108)	(0.0073)	(0.0064)
Implied Change in Jobs	-208	–	–	+252
N	2730	2730	2730	2730
R^2	0.909	0.973	0.933	0.836
	Panel B: IV			
New Value per Capita (1000s)	-0.0564***	-0.0225	0.0362*	0.0427***
	(0.0216)	(0.0172)	(0.0205)	(0.0187)
Implied Change in Jobs	-359	–	+230	+272
N	2730	2730	2730	2730
R^2	0.908	0.972	0.930	0.831

Note: Dependent variables are indicated in the columns. Models include county fixed effects, year fixed effects and county-specific recession effects. The dependent variables measure the fraction of the local jobs held by residents in different locations. The dependent variable in Column (1) is the fraction of local jobs held by residents who live within the county. The dependent variable in Column (2) is the fraction of local jobs held by residents who live outside the county but within 25 miles. The dependent variables in Columns (3) and (4) are constructed similarly. 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. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value<0.025.

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. However, the shift in earnings increases manifests itself different 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)). Further, the point estimate in column (6), the highest earnings category, is also positive for nonlocals. 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)	-0.0249***	-0.0116	0.0038	0.0048	0.0067	0.0211***
	(0.0062)	(0.0076)	(0.0083)	(0.0071)	(0.0079)	(0.0089)
Implied Change in Jobs	-159	–	–	–	–	+135
N	2730	2730	2730	2730	2730	2730
R^2	0.871	0.915	0.899	0.835	0.847	0.917
			Panel B: IV			
New Value per Capita (1000s)	-0.0533***	0.0200	-0.0231	-0.0189	0.0534*	0.0219
	(0.0163)	(0.0139)	(0.0196)	(0.0170)	(0.0283)	(0.0189)
Implied Change in Jobs	-339	–	–	–	+340	–
N	2730	2730	2730	2730	2730	2730
R^2	0.866	0.911	0.893	0.828	0.834	0.917

Note: Dependent variables are indicated in the columns. Models include county fixed effects, year fixed effects and county-specific recession effects. The dependent variable in Column (1) is the fraction of local jobs held by residents who live within the county and make less than $1250/month. The dependent variable in Column (2) is the fraction of local jobs held by residents who live within the county and earn between $1250/month and $3333/month. The dependent variable in Column (6) is the fraction of local jobs held by workers who live outside the county and earn more than $3333/month. The dependent variables in other columns are constructed similarly. 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. Standard errors are clustered by county. * p-value<0.10, ** p-value<0.05, *** p-value<0.025.

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 more than 600 jobs but local residential aggregate employment is unaffected. Furthermore, when disaggregating employment by sector, workplace employment increases in all sectors but local residential employment increases are only found in the oil and gas sector. 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.

These 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.

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*

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*

Jacobsen, G. D. and Parker, D. P. (2016). The Economic Aftermath of Resource Booms: Evidence from Boomtowns in the American West, *The Economic Journal* **126**(593): 1092–1128.
URL: *http://onlinelibrary.wiley.com/doi/10.1111/ecoj.12173/abstract*

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. 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*

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*

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*

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*

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*