Labour Economics 82 (2023) 102362

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Oil & gas induced economic fluctuations and self-employment[☆]

Bulent Unel[a,*], Gregory B. Upton Jr.[b]

[a] Department of Economics, Louisiana State University, Baton Rouge, LA 70803, United States
[b] Center for Energy Studies, Louisiana State University, Baton Rouge, LA 70803, United States

ARTICLE INFO

JEL classification:
J24
L26
M13
Q33
Q35

Keywords:
Entrepreneurship
Oil and natural gas
Hydraulic fracturing
Incorporated self-employment
Unincorporated self-employment

ABSTRACT

This paper exploits oil and gas activity generated by recent technological advancements to understand the effect of localized boom and busts on self-employment. We find a positive contemporaneous impact on self-employment, mainly driven by self-employment in non-mining industries. We also find that self-employment is pro-cyclical, meaning that self-employment increases during oil and gas booms and contracts during the bust. Finally, results suggest that self-employment explains an economically meaningful share of the employment adjustment; specifically we estimate that about 11% of the employment adjustment can be explained by self-employed workers, a group which makes up about 9% of total employment.

1. Introduction

Entrepreneurs, in particular small business owners, are commonly viewed as the engine of economic growth and prosperity. They play a disproportionately large role in job creation and destruction, and hold a substantial portion of the U.S. wealth. For instance, Fairlie et al. (2019) find that startups create approximately 3 million jobs per year with 2.9 million of these jobs persisting five years later. Further, De Nardi et al. (2007) report that more than half of self-employed business owners fall in the top wealth quantile. Therefore, understanding the nature of entrepreneurship and identifying factors that affect it has always been of interest to governments, policymakers, and researchers.

Self-employment has been considered as the “simplest kind of entrepreneurship” (Blanchflower and Oswald, 1998), and a large body of research documents that individuals’ decision to engage in entrepreneurship changes with economic conditions.[1] In this paper, we contribute to this literature by investigating how localized boom and bust cycles, induced by oil and gas activity, impacted self-employment in the U.S. over the past two decades. This time period provides a unique opportunity to study self-employment dynamics, as it was dominated by plausibly exogenous fluctuations in oil and gas production from shale geological formations. These “boom towns” experienced boom and bust cycles induced by a combination of external factors including oil and natural gas prices, resource availability, and differences in the oil to gas ratio of hydrocarbons produced that naturally varies spatially within individual basins and across basins. As a growing recent literature documents, these areas experienced employment growth throughout a broad group of sectors of the economy. Although our analysis will focus on the entire U.S. (not just these shale areas), we will show that much of the fluctuation during the time frame of our analysis (2005–2019) was driven by activity in “shale boom” areas.

We utilize data on self-employment from the U.S. Census American Community Surveys (ACS), and further distinguish between incorporated and unincorporated self-employed workers, as these have been shown to have starkly different traits and income profiles (Levine and Rubinstein, 2017). We aggregate self-employment to commuting zones (CZs), which represent clusters of counties that have strong commuting

☆ Bulent Unel gratefully acknowledges Jill and Roger Jenkins Energy Research Grant from the LSU Foundation. We thank Ozkan Eren, Barton Willage and especially Dan Keniston for their helpful comments and suggestions, and Mark Agerton for sharing computer code to compile oil & gas data with the research community. We also thank participants of the 5th Annual CEBRA Workshop for Commodities and Macroeconomics hosted by the Federal Reserve Board, especially Tony Cookson for his comments and suggestions. Finally, we thank the editor Wilbert van der Klaauw and two anonymous referees for their constructive comments and suggestions.

* Corresponding author.
E-mail addresses: bunel@lsu.edu (B. Unel), gupton3@lsu.edu (G.B. Upton Jr.).

[1] Some studies identify entrepreneurship by the number of employer establishments (e.g., Haltiwanger et al. (2013)). But employer-based counts represent a small share of entrepreneurship (Fairlie, 2014). Individual business owners cover both employer and non-employer firms.

https://doi.org/10.1016/j.labeco.2023.102362
Received 14 October 2021; Received in revised form 14 March 2023; Accepted 20 March 2023
Available online 5 April 2023
0927-5371/

ties between workers and businesses (Tolbert and Sizer, 1996). To identify the causal effect of oil and gas activity on self-employment, we use detail well-level oil and natural gas production data for more than one million wells in the United States. For each county and month, we calculate the value of production from new wells (i.e., wells began production within the preceding 12 months). These monthly county values are then aggregated into yearly CZs.

Our main findings can be summarized as follows. First, we find that oil and gas activity has a positive contemporaneous impact on self-employment. Lagged effects of oil and gas activity on self-employment are negative and imply that once the boom subsides, self-employment adjusts closer to the pre-boom level. Together, this suggests that self-employment is pro-cyclical, i.e. it increases during a business cycle expansion and reduces during a business cycle contraction. Second, we estimate the sensitivity of total employment to new oil and gas production and then decompose the share of the employment adjustment that stems from self-employment. Point estimates suggest that about 11% of the employment adjustment can be explained by self-employed workers, a group which makes up about 9% of total employment. Finally, our industry-level analysis indicate that the effects on self-employment are driven primarily by individuals in non-mining industries. About two thirds of the adjustment in self-employment comes from service sectors. Notably, prior studies have found that overall employment effects from the shale boom have been concentrated in mining, construction and transportation (Feyrer et al., 2017; Marchand, 2012), thus this research reveals that the industry of adjustment for self-employed workers might be very different than the economy as a whole.

This paper first contributes to a large literature that examines the role of external factors affecting the choice to become entrepreneur. The literature has considered factors including access to credit (Asiedu et al., 2012; Cagetti and De Nardi, 2006; Hurst and Lusardi, 2004), inheritance and gifts (Blanchflower and Oswald, 1998), globalization (Dinopoulos and Unel, 2015; Eren et al., 2019), immigration (Fairlie and Meyer, 2003; Pekkala Kerr and Kerr, 2020), and government policies (Beland and Unel, 2019; Cullen and Gordon, 2007), among others. A few papers empirically investigated the nexus between self-employment and the economy wide business cycle (Fossen, 2020; Levine and Rubinstein, 2018).

This paper also contributes to a growing body of work that quantifies the economic effects of localized natural resource based booms.[2] While this literature began before the specific shale oil and gas booms of this past decade (Allcott and Keniston, 2018; Black et al., 2005), this new era of shale has created a significant resurgence in part because of the clean empirical identification afforded by the nature of the shock. Previous studies have investigated the impact of shale boom on various outcomes, and paper is more closely related to studies that have investigated the effects of resource booms on labor-market outcomes. Agerton et al. (2017); Feyrer et al. (2017); Green et al. (2019); Marchand (2012) find a positive effect of oil and gas activity on employment and wages.[3] Our analysis complements and extends these studies by considering self-employed workers who are not included in establishment based labor-market data.

Our paper relates most closely to two recent studies at the intersection of oil & gas booms and entrepreneurship activity: Tsvetkova and Partridge (2017) and Bellon et al. (2021). The former examines the impact of employment growth in oil and gas industry on self-employment growth in the U.S. over the 2001–2013 period. They find that oil and gas sector expansion crowds out self-employment, which stems from a large reduction in self-employment in non-mining sectors. Our analysis differs from theirs in several aspects. Most notably, we use oil and gas production from *new* wells, which depends on preexisting geology and the recent drilling technology. We refer to "new wells" as those that began production within the preceding 12 months.[4] These monthly county values are then aggregated into yearly CZs. Unlike employment growth, new production is plausibly exogenous to other shocks occurring in the region, and more appropriate for identification. Results of our analysis do not corroborate conclusions from theirs.

Using data on unexpected payments to individuals from the shale boom in Texas, Bellon et al. (2021) investigate the effect of wealth windfalls on self-employment decisions. They find that individuals who received large wealth shocks have higher self-employment rates relative to others who received smaller or no wealth shocks. Our study documents the net effect of the shale boom on entrepreneurship, and does not focus on individual decisions based on specific royalty payments. We therefore view Bellon et al. (2021) as complementary to our work.

The rest of this paper is organized as follows. Section 2 introduces the data on labor markets and oil and gas production along with measurement and construction of key variables. Section 3 describes our estimation strategies. Section 4 presents our results and discuss their implications. Section 5 investigates the robustness of our findings, and Section 6 concludes the paper.

2. Data and descriptive statistics

This section discusses the sources and construction of key variables used in our analysis. We first discuss the data on labor markets, in particular, on self-employed individuals and their characteristics. We then introduce and discuss the data on oil and gas production.

2.1. Labor market data

Data on self-employment and wage workers are from the American Community Survey (ACS). We utilize publicly available data from the Integrated Public Use Micro Samples (IPUMS) website (Flood et al., 2020) for years 2005 to 2019. Microdata is acquired at the Public Use Microdata Area (PUMA) geographic granularity.[5] Counties are the smallest geographic unit identified in data files, however, due to small samples especially in rural counties, for approximately forty percent of individuals county identifiers are not available for disclosure reasons.[6]

Our analysis is at the Commuting Zone (CZ) level. We choose CZs as the level of geographic granularity as they represent clusters of counties with strong commuting ties between employers and employees as many workers commute across counties for work (Tolbert and Sizer, 1996). CZs therefore represent a more cohesive labor market than counties, and have been extensively used in recent research (Autor and Dorn, 2013, Feyrer et al., 2017, Autor et al. 2019, among others). Using crosswalk files from Autor and Dorn (2013), we map PUMAs to 741 commuting zones that cover the entire area of the U.S. over the period 2005–2019.[7]

[2] This paper contributes to the literature investigating short-term boom and busts induced by resource extraction, in contrast to the large literature on resource endowments and long run economic growth (Alexeev and Conrad, 2009; Michaels, 2010; Oliver and Upton, 2022; van der Ploeg, 2011; Venables, 2016).

[3] Other studies corroborate the positive impact of the shale boom on local-labor markets (Bartik et al., 2019; Decker et al., 2021; Jacobsen, 2019; Komarek, 2016; Maniloff and Mastromonaco, 2017; Marchand and Weber, 2018; McCollum and Upton, 2018; Weber, 2012). It should be noted that due to the oil and natural gas price declines of 2014, there is an emerging literature on the "bust" side of the cycle as well. For instance, Brown (2015) finds that elimination of each active rig eliminates 28 jobs in the first month and this increases to 171 jobs eliminated in the long-run.

[4] For example, if a well reports production for the first time in July of 2010, production from this well will be included in the data until June of 2011.

[5] 2005 is the earliest year for which yearly data is continuously available at the PUMA level.

[6] Metropolitan statistical area (MSA) is another geographic unit identified in the ACS. However, like counties, many MSAs are incompletely identified, and using them does not cover the entire area of the U.S.

[7] Crosswalk files are available at David Dorn's website: https://www.ddorn.net/data.htm.

Table 1
Summary statistics on worker class, 2005–2019.

	Self-employed Individuals			Wage
	All	Incorp.	Unincorp.	Workers
	1	2	3	4
Female (%)	36.9	30.0	41.0	49.1
Age	45.8	46.8	45.1	39.7
	(11.0)	(10.3)	(11.4)	(12.7)
White (%)	81.7	83.7	80.4	75.4
Some College (%)	64.7	73.6	59.5	64.5
Hours Worked	40.5	44.7	37.9	39.4
	(15.7)	(14.3)	(15.9)	(11.2)
Annual Income (1,000)	50.7	66.9	41.1	46.4
	(60.3)	(63.9)	(55.8)	(40.9)
Mining (%)	0.2	0.3	0.2	0.6
Construction (%)	18.2	16.8	19.0	6.2
Manufacturing (%)	4.1	5.5	3.3	12.8
Transport (%)	4.8	4.6	4.9	4.7
Service (%)	72.7	72.8	72.6	75.6
Sample Share (%)	9.4	3.5	5.9	90.6

Notes: The data draw on the ACS Files from IPUMS (Flood et al. 2020), and cover all individuals who are 16 years or older and working in the non-agricultural private sector. Some College represents all individuals who have at least some college education. The average annual income are in thousands of the 2015 US dollars. Numbers in parentheses are standard deviations, and the census weights are used in all calculations. Sample includes about 16.9 million observations.

Utilizing CZs not only identifies a more cohesive labor market, but also circumvents dropping the prior-mentioned forty percent of individuals with missing county identifiers. This is especially important in this context, as many shale counties are rural.

The ACS covers more than two million households per year, and provides information about demography (i.e., gender, age, race), educational level, work (i.e., employment status, worker class, industry worked, occupation, etc.), health insurance, and migration. Our analysis considers individuals 16 years and older who work in non-agricultural private sector, and we exclude all observations with imputed/missing employment status, worker class, and industry. The survey classifies workers as self-employed or wage and salary workers, with the former further classified as incorporated and unincorporated.[8]

Previous studies have broadly referred to all self-employed individuals as entrepreneurs (Blanchflower and Oswald, 1998; Borjas and Bronars, 1989; Fairlie, 2014). However, Levine and Rubinstein (2017) show that incorporated and unincorporated self-employed individuals differ in their cognitive and non-cognitive traits.[9] For example, incorporated self-employed are more educated and earn more than salaried workers. By contrast, unincorporated self-employed individuals are generally in low-skill intensive occupations. While there are differences in nomenclature across the literature, we specifically examine self-employment, and further distinguish between incorporated and unincorporated self-employed workers.

Table 1 reports descriptive statistics on different classes of workers; numbers in parentheses are standard deviations. Column 1 in Table 1 presents the statistics related to all self-employed individuals, whereas columns 2 and 3 disaggregate self-employed workers as incorporated or unincorporated business owners, respectively. The last column reports the corresponding statistics for wage and salary workers. The sample covers about 16.9 million observations, 9.4 percent of which is self-employed (see the last row). Columns 1–3 reveal that white and male workers are a higher proportion of self-employed compared to wage workers.

Summary statistics show that incorporated and unincorporated self-employed workers differ in a few important ways. While incorporated self-employed workers work more than wage and salary workers, unincorporated self-employed workers work fewer hours. Further, the average annual earned income of incorporated self-employed workers is about 40 percent higher than wage workers, and is 61 percent higher than unincorporated self-employed. Thus, consistent with Levine and Rubinstein (2017), these summary statistics show that incorporated self-employed individuals are more educated, work longer hours, and have higher annual income than unincorporated self-employed and wage workers.[10] Both incorporated and unincorporated self-employed workers represent higher shares in the construction and service sectors and lower shares in mining, manufacturing and transportation.

Fig. 1 shows time trends of self-employed individuals as a percent of total employment in our sample. The share of self-employed individuals is about 10 percent in 2005, and declines to about 8.3 percent by 2013, remaining relatively constant thereafter. Note that the share of incorporated and unincorporated self-employed individuals follow a similar pattern. The time trend of each group, however, vary substantially across commuting zones.

2.2. Oil and gas data

We use detailed well-level production estimates for more than one million wells in the United States as compiled by Enverus (formerly DrillingInfo). Enverus collects data from state agencies such as the Railroad Commission of Texas, the Department of Natural Resources in Louisiana, and North Dakota Industrial Commission. In different states, oil and gas production is reported at different levels of aggregation, which typically include leases, units, or wells.[11] *Enverus* compiles the data across states and estimates well-level monthly production estimates of oil and natural gas.[12]

We utilize the timing and intensity of oil and gas activity alongside geological data on reservoirs thousands of feet below the earth's surface. Fig. 2(a) shows the seven largest shale areas in the United States.[13] The shale plays are geographically dispersed throughout the country, and their placement is determined by geological formations thousands of feet below the earth's surface. Naturally, each of these formations has different compositions of oil and natural gas. For example, the Haynesville and Appalacia areas are overwhelmingly "dry" natural gas, as opposed to Bakken which is overwhelmingly oil. Eagle Ford has a mix of both oil and natural gas, and the ratio of oil and natural gas naturally changes geographically across the play.

[8] Abraham et al. (2019) show a growing discrepancy between self-employment rates as measured in household surveys (e.g., CPS-ASEC) and self-employment rates as measured in tax data. Hwoever, in their analysis the discrepancy is mainly stem from the surge in passenger driver self-employment. Our results are not driven by the transport sector, and thus their point does not pose any serious problems to our analysis.

[9] In addition, incorporation involves creating a legal entity that is distinct from the owner. This allows the corporation to take out debts and be involved in legal disputes separate from its owner.

[10] Numbers reported in Table 1 are in thousands of 2015 dollars, using the CPI retrieved from FRED, Federal Reserve Bank of St. Louis. The bottom 1 percent and the top 1 percent of income data are trimmed from the sample.

[11] Unitization is when several tracts of land with different ownership are pooled together for purposes of sharing royalties. For instance, a company cannot typically drill on a one acre plot of land and associate all of the production to the surface land owner, as the oil and natural gas is being pulled from adjacent land with different owners. Individual states have different processes for addressing this common issue. A detailed description of the laws surrounding oil and gas drilling with a focused comparison between Louisiana and Texas can be found in (Martin and Yeates, 1992).

[12] In many instances, monthly production is reported by well. In other instances, monthly oil and gas production is reported for a group of wells. Thus, *Enverus* allocates reported production to individual wells to provide a well level estimate.

[13] These shale plays are from the U.S. geological survey (EIA, 2011). For illustrative purposes we utilize the seven geographic areas highlighted in the U.S. Energy Information Administration's Drilling Productivity Reports.

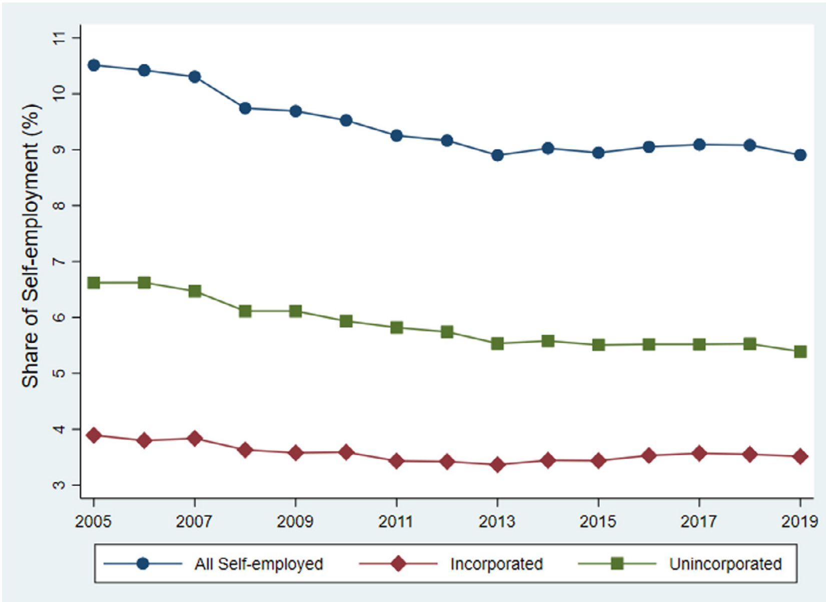

Fig. 1. Share of self-employed individuals in total employment (%).

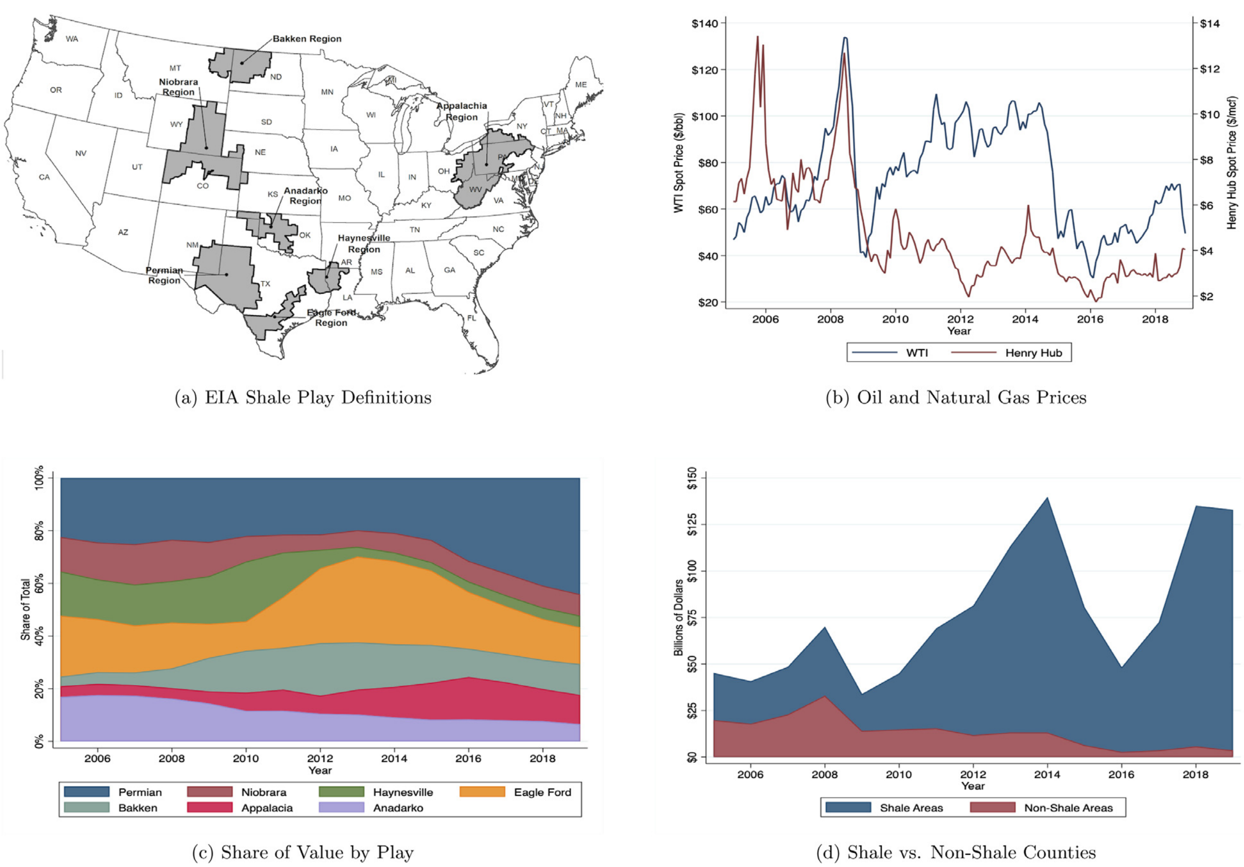

Fig. 2. Oil and natural gas data.

Between 2004 and 2019, there were about 492 thousand wells were drilled nation-wide, with more than 60 percent of these wells drilled in the seven shale areas identified in EIA's Drilling Productivity Reports. Appendix Table A1 shows the number of wells drilled that eventually reported economic quantities of oil and/or natural gas across regions. The Appalachia has the largest number of wells reporting new production over this period, with about 75 thousand wells. The smallest wells drilled reporting production is in the Haynesville with 18 thousand wells.

For our analysis, oil and gas production from the "new" wells will be used to identify the timing and intensity for which economic activity was occurring in these regions. A number of papers focusing on different geographic regions have generally found that oil and natural gas production from existing wells do not respond to changes in commodity prices, as drilling is the most important margin of adjustment (Anderson et al., 2018; Newell and Prest, 2019). This is because when a well is initially drilled and completed it will, within a month or so, reach its maximum production. The industry generally refers to this as a well's "initial production." Once production commences, the well can continue producing for many years, in some instances decades. Thus, the goal of this research is to utilize the economic activity generated in areas as a plausibly exogenous shock to the local labor market, in the same vein as Bartik et al. (2019); Feyrer et al. (2017).

For each month, oil and natural gas production from all wells that reported first production within the past 12 months across county equivalents in the United States is summed. For example, a well that reports first production in June of 2010, will have oil and/or gas production included in all months from June of 2010 to May of 2011.[14] These monthly production numbers are then multiplied by West Texas Intermediate oil price and Henry Hub natural gas price sourced from the U.S. Energy Information Administration (EIA) to calculate the estimated total value of oil and gas production from new wells in each county.[15] These county estimates are then aggregated into Commuting Zones, and all values are expressed in 2015 millions of dollars using the CPI.[16]

Fig. 2 (b) shows the variation in the timing of oil and natural gas price changes. In the early part of the sample, oil and natural gas prices move in tandem with one another. But in 2009, the price of oil and natural gas began to diverge. This divergence continued until 2014, when prices converged once again with the oil price drop. The timing of oil and natural gas price shocks impacted different plays in very different ways. As shown in Fig. 2(c), in the early sample period, the Haynesville shale accounted for a relatively large share of value of new production, but by 2014 this had attenuated significantly. Compare this to Bakken that had a small share of value of new production at the beginning of the sample, and this peaked in 2013 before the oil price crash in 2014.

Finally, Fig. 2d shows the value of production from new wells (inclusive of oil and natural gas) in these largest seven shale plays compared to all other counties in the United States. This figure illustrates that the value of production in the early sample period for shale and non-shale regions were approximately parallel. But beginning around 2009, the two diverge considerably, with the lion's share of value of new production coming from shale counties. The 2014 price bust is observed and then a rebound through 2019. Altogether, Fig. 2 illustrates the spatial and temporal variation that will be utilized in identifying the effect of oil and gas induced localized business cycles on self-employment.

Appendix Table A2 shows summary statistics on the distribution of the new value of oil and gas produced by commuting zone throughout our sample period. Of the 741 total commuting zones, 337 had some value of oil and gas production. The remaining 404 CZs with no new oil and gas production are, in essence, the control group in this analysis. Thus, approximately 45 percent of CZs had some new oil and gas production over the sample period were treated but at very different levels of treatment.

Focusing now on the CZs with some value of production, the median CZ had $2.7 billion in new production over the sample period, or approximately $194 million per year over the fourteen year sample. But, this distribution is right skewed, with the average CZ having almost $37 billion in new value of production, or approximately $2.6 billion per year. The value of production from within the first year of a well's production (i.e. "new wells") (NV_t) changes substantially over the sample period. The CZ with the largest value of new production was in West Texas with over $841 billion in new production, or an average of over $60 billion in new value of production per year over the sample period. Further, the correlation between NV_t and NV_{t-1} varies considerably across CZs with a mean and standard deviation of 0.52 and 0.29, respectively.

3. Empirical methodology

Oil and gas activity can affect self-employment by creating new business opportunities in the mining sector and other industries that support the mining sector through input-output linkages. Further, local landowners who receive bonus and royalty checks for oil and gas production that occurs beneath their land may use this capital to start their own business (Brown et al., 2019). However, a business cycle expansion can increase the opportunity cost of becoming self-employed due to an increase in market wages. Therefore the net impact of these oil and gas shocks on self-employment is a priori ambiguous, and addressing it is ultimately an empirical question.

In estimating the impact of oil and gas production on self-employment, we follow an econometric framework similar to Feyrer et al. (2017). Let E_{zt} denote the number of self-employed individuals working in the non-agricultural private sector in commuting zone (CZ) z in year t,[17] and NV_{zt} denote the total market value of oil and gas production (in 2015 millions of USD) from *new* wells in the commuting zone normalized by the 2005 CZ employment.[18] We then estimate

$$\triangle E_{zt} = \beta NV_{z,t} + \beta_1 NV_{z,t-1} + \beta_2 NV_{z,t-2} + \eta_z + \eta_t + \varepsilon_{zt}, \tag{1}$$

where $\triangle E_{zt} = (E_{zt} - E_{z,t-1})/L_{z,2005}$, i.e. the change in the number of self-employment normalized by the 2005 CZ employment. CZ fixed effects η_z are included to control for time-invariant factors that can affect entrepreneurship dynamics across commuting zones, and time fixed effects η_t to control for common macroeconomics shocks and trends. Finally, ε_{zt} is the error term.[19]

We consider only production from wells within the first 12 months of reported production (i.e. "new wells") because the output (determined by the pre-existing geology and new drilling technology) is plausibly exogenous to other shocks in the economy (see footnote 14). In addition, previous studies have generally found that oil and natural gas production from *existing* wells do not respond to changes in commodity prices because drilling is the most important margin of adjustment (Anderson et al., 2018). In Eq. (1), we also include $NV_{z,t-\ell}$ for $\ell = 1, 2$ to control for the delayed effects of oil and gas production.[20] Thus, parameters β and β_ℓ measure the contemporaneous and lagged effects

[14] In this example, the production from this well over June 2010-December 2010 is included in calendar year 2010. Similarly, the production from that same well from January 2011-May 2011 is included in the 2011 calendar year. Thus, an individual well can contribute to the value of new value of production variable in two years. For this well, any oil and gas produced after May 2011 is excluded.

[15] Oil prices vary significantly across locations, especially during the peak of the shale boom. During the time of this writing, similar large wellhead price discounts are observed in natural gas markets. In this way, the value of production is likely over-stated and therefore point estimates are likely understated.

[16] CPI retrieved from FRED, Federal Reserve Bank of St. Louis: https://fred.stlouisfed.org/series/CPI.

[17] If we consider a particular type of self-employment, say incorporated, E_{zt} denotes the number of incorporated self-employed business owners. For each CZ and year, we calculate E_{zt} by multiplying the share of self-employment in CZ total employment (calculated from the ACS data) with the CZ employment numbers (from the Bureau of Labor Statistics). The latter data are available at the county-level since 1990, and we aggregated these counties into CZs using Dorn's crosswalk file.

[18] The normalization is done to control for the cross-zone differences in population. We chose the 2005 CZ employment (instead of the previous year's employment) to reduce concerns that population grew due to the shale boom. However, we later show robustness checks where we normalize by the previous year's employment.

[19] Equation (1) does not include any CZ-level controls (e.g., proportion of the CZ labor force who are male, white, proportion of the labor force with some college education, etc.). These are potentially outcome variables, and considered bad controls (Angrist and Pischke, 2009). However, including them in the model does not have significant effects on estimates, as we shall show later.

[20] Including additional lags does not meaningfully impact results and coefficients on further lags are insignificant.

of new oil and gas production, respectively. Since the change in self-employment and new production values are normalized by the 2005 CZ employment, β has a simple interpretation: one million dollars of production increases the number of self-employed individuals by β.

Estimation of Eq. (1) using the Ordinary Least Square (OLS) approach can provide unbiased coefficient estimates of βs if the value of production is essentially randomly assigned to commuting zones based on pre-existing geology. While the availability of a resource is clearly exogenous, a firm's decision to extract in a particular location may not be. For instance, a rural area with relatively low land value might be more likely to be utilized for oil and gas drilling while urban areas are inherently difficult to drill in due to physical constraints.

We use the instrumental variable (IV) approach to address the potential endogeneity problem. In constructing the instrument, it is assumed that the value of new production in a county varies across time and across the shale plays, following Feyrer et al. (2017). However, estimating new values based on temporal and spatial variables must be done with caution because the dependent variable (i.e. value of oil and gas production from new wells) contains many zeros. As we explain in more detail below, we estimate the following model:

$$NV_{ct} = \exp(\alpha_c + \lambda_{st}) + v_{ct}, \tag{2}$$

where $\exp(\cdot)$ denotes the exponential function, α_c is a dummy for each county, λ_{st} represents a set of dummies for each play-year combination, and v_{ct} is error term. Observe that the predictions from equation (2) incorporate the timing of new production from the play dummies while controlling for the idiosyncratic level of production in each county. The predicted values for new production in the county ($\widehat{NV}_{ct}$) are therefore based on the timing of new production for all counties in a given shale play. Adding these predicted values across counties in each commuting zone, we obtain predicted values for the new production in the CZ ($\widehat{NV}_{zt}$). We use $\widehat{NV}_{zt}$ (normalized by the 2005 CZ employment) as an instrument for the oil and gas production NV_{zt} in Eq. (1). Similarly, we use $\widehat{NV}_{zt-\ell}$ as an instrument for $NV_{z,t-\ell}$, for $\ell = 1, 2$.

We estimate Eq. (2) using the Poisson pseudo maximum likelihood (PPML) estimator of Gourieroux et al. (1984), popularized by Silva and Tenreyro (2006). This method is especially appealing in the present context because the majority of counties do not have any new production in a given year.[21] Silva and Tenreyro (2011) and Correia et al. (2020) show that the PPML estimator behaves well when the dependent variable has a large portion of zeros. In addition, in the presence of heterokedasticity, the log-linearized models estimated by OLS leads to biased estimates (Silva and Tenreyro, 2006).

We will also estimate equation (1) utilizing total employment. Although point estimates for β and β_ℓ will be presented, the focus of the analysis is the share of the total employment effect that can be explained by self-employed workers. We heed caution in the interpretation of β in a specific context, as the value of production is based on prices of oil and natural gas at trading hubs.[22] Thus, the actual value of production in the specific location is likely less, therefore β and β_ℓ are underestimated relative to the value at the wellhead.

4. Results

This section presents the results of our empirical analysis based on Eq. (1). We then conduct our analysis at industry level and discuss its implications. Finally, we present a rich set of sensitivity checks to investigate robustness of the results.

Table 2
Impact of new oil & gas production on self-employment in the U.S.

	A. OLS Estimates			B. IV Estimates		
Variable	All 1	Incorp 2	Unincorp 3	All 4	Incorp 5	Unincorp 6
NV_t	0.066*	0.004	0.062**	0.060	−0.015	0.074*
	(0.035)	(0.015)	(0.026)	(0.040)	(0.014)	(0.035)
NV_{t-1}	−0.032	0.021	−0.053*	−0.037	0.043***	−0.081
	(0.043)	(0.018)	(0.031)	(0.058)	(0.017)	(0.056)
NV_{t-2}	−0.022	−0.022*	0.008	−0.029	−0.037***	0.009
	(0.028)	(0.013)	(0.019)	(0.042)	(0.013)	(0.035)

Notes: The sample size in each panel is 10,374 observations from 741 U.S. commuting zones over the 2006–2019 period (includes 2005 with lagged value). All regressions include CZ-fixed and year-fixed effects and observations are weighted by 2005 CZ employment. Standard errors are clustered at the CZ level, and ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively. For columns 4–6, standard errors are obtained from bootstrapping with 999 repetitions.

4.1. Main results

Table 2 reports the impact of new oil and gas production on the change in self-employment using equation (1). Here "All" represent total self-employment, i.e. both incorporated and unincorporated. The sample size in each column is 10,374 observations from 741 commuting zones over the 2006–2019 period. All regressions include commuting zone and year fixed effects, and numbers in parentheses are robust standard errors clustered at the CZ-level. Regressions are weighted by the 2005 CZ employment.

Panel A reports the OLS results and Panel B the IV estimates, and our focus is the IV results.[23] Since the instrument is predicted from the data, standard errors obtained from bootstrapping (with 999 replications) clustered at the CZ level. Further, the first-stage (Kleibergen-Paap) F-statistic for IV regressions is about 31.4, which is well above 10—the conventional cutoff used in the literature. Comparing OLS and IV estimates, note that the magnitude of the contemporary IV coefficients is mostly larger than estimates obtained with OLS. Estimates in Column 4 indicate that the contemporaneous effect of new production on self-employment is positive but statistically insignificant (p-val ≈ 0.13). Ignoring the precision, the point estimate implies that one million dollars of oil and gas production from new wells increases the number of self-employed individuals by 0.06; or equivalently, one billion dollars of new production increases self-employment by 60. The coefficients on lagged values (i.e., $NV_{t-\ell}$) are negative and statistically insignificant.[24] However, their joint impact is statistically significant at the 5% level, resulting in a reduction of 66 self-employed workers after a hypothetical one period shock within about two years after the shock subsides. The cumulative effect (i.e., the sum of all coefficients) is close to zero and statistically insignificant.[25] Thus, oil and gas extraction from new wells has a positive and significant impact on self-employment, but self employment adjusts closer to pre-boom levels after the boom subsides. In other words, this result suggests that self-employment is pro-cyclical.

Columns 5 reports the impact of new production on incorporated self-employment. The contemporaneous effect of new production on in-

[21] Feyrer et al. (2017) estimate $\ln(NV_{ct} + 1) = \alpha_c + \lambda_{st} + v_{ct}$, another common way to handle zeros in the dependent variable. They measure NV_{ct} in millions of dollar, and thus adding one effectively means adding one million dollars of new oil and gas production. In reconstructing their instrument identically, we find that the results are sensitive to how new values are measured (e.g., in dollars or million dollars), which is ultimately an arbitrary choice. Having said that, we obtain larger point estimates with large standard errors when we use the Feyrer et al. approach.

[22] This was especially the case during the shale boom, when prices diverged considerably in different locations across the U.S. (Agerton and Upton, 2019)

[23] Note that although specific point estimates and standard errors vary across the OLS and IV specifications, examination of both reveal the same overarching conclusion.

[24] Including only one year-lagged value of new production, the estimated coefficients on NV_t and NV_{t-1} are 0.072(0.030) and -0.067(0.029), respectively. Both coefficients are statistically significant at the 5% level.

[25] The ACS does not follow individuals over time, and thus we cannot directly study entry and exit into self-employment.

corporated self-employment is negative but statistically insignificant. However, the lagged effects are large and statistically highly significant. Estimates suggest that a one billion dollars of new oil and gas production increases the incorporated business owners by 43 one year after the production, but again self employment adjusts downward after the boom subsides. Insignificant contemporaneous effect may stem from the fact that creating an incorporated business may take longer time due to coordination and capital requirements (Levine and Rubinstein 2018).

Columns 6 reports results for unincorporated self-employment. As in column 4, the contemporaneous effect of new production on incorporated self-employment is positive and statistically significant. The first-lagged effect is large, negative, and statistically insignificant, and the second-lagged effect is positive, small, and statistically insignificant. Estimates suggest that a one billion dollars of new oil and gas production increases unincorporated self-employment by 74, but the positive contemporaneous effect is wiped out one year after the production.[26] The sum of coefficients in this column is almost zero, indicating that in response to a one period shock there would be no net significant effect on self-employment after two years.

In sum, results in Table 2 indicate that oil and gas production from the recent shale-boom revolution has had a positive contemporaneous effect on self-employment, but once the boom subsides self-employment adjusts closer to the pre-boom level. A plausible explanation for why oil & gas shocks induce this boom-bust pattern is that while the production from an oil and/or gas well can continue for decades once drilled, wells typically experience a relatively quick decline from the initial production levels; this is especially true in shale formations. Newell et al. (2016), for example, show that the median well is producing about 40 percent of the initial peak one year after the start of production, less than 25 percent two years of after the production. Further, as discussed earlier, several studies have generally found that drilling constitutes the most important margin of production and economic activity. Once the drilling process is completed, the demand for labor and other inputs will naturally subside.

Results in Context We now turn to evaluate the economic significance of our findings. Consider, for example, 2012 in which the average value of new oil and gas production was about 282 million dollars across CZs. Our estimates imply that this activity increases self-employment by 17, which is a very small contribution to self-employment given that the average number of self-employed individuals in that year is about 16,800. However, it should be noted that the areas where shale boom is taking place tend to be small rural counties. Consider, for example, the CZ consisting of Loving, Reeves, Ward, and Winkler counties in Texas. In 2012, the total employment was 13,124 and the value of new production was about 2.12 billion dollars. Consequently, extrapolating point estimates, the shale-boom activity is estimated to increase self-employment by 127 ($\approx 2.12 \times 60$), which is about 15% increase in self-employment in the county.

We next present a more comprehensive way to assess the magnitude of our findings in context. To do so, we estimate equation (1) utilizing total employment (inclusive of self-employment and wage & salaried workers). Column 1 in Table 3 reports the results, which indicate that one billion dollars of new oil and gas production increases total employment by 538. The coefficients on lagged values are negative and large, and their joint impact is statistically significant at the 5% level. For comparison, column 2 reproduces the benchmark results from Table 2, and the last column reports the share of the total employment adjustment that stems from self-employment.

Since effects of the lagged values are imprecisely estimated, we re-estimate Eq. (1) with only NV_{zt} and NV_{zt-1}. In this case, the estimate on NV_{zt-1} represents the cumulative effect of lagged production values. We

Table 3
Self-employment as share of labor market adjustment .

Variable	Employment 1	Self-employment 2	Contribution 3
A. Estimates with two lags			
NV_t	0.538**	0.060	11.2%
	(0.266)	(0.040)	
NV_{t-1}	–0.186	–0.037	19.9%
	(0.339)	(0.058)	
NV_{t-2}	–0.369	–0.029	7.9%
	(0.333)	(0.042)	
B. Estimates with one lag			
NV_t	0.700***	0.072**	10.3%
	(0.204)	(0.030)	
NV_{t-1}	–0.571***	–0.067**	11.7%
	(0.159)	(0.029)	

Notes: The first column uses the total employment as the dependent variable, including self-employment and wage workers. Column 2 represents the benchmark results from column 4 in Table 2. All regressions include CZ-fixed and year-fixed effects and observations are weighted by 2005 CZ employment. Standard errors obtained from bootstrapping (with 9999 repetitions) are clustered at the CZ level, and ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

now find that one billion dollars of new oil and gas production increases employment by 700. Viewing the lagged coefficient, the point estimate implies a reduction in 570 employment the following year. A joint test of significance for the sum of the contemporaneous and lagged effect yields a statistically significant increase in 130 employment, suggesting that oil and gas production has a more persistent impact on wage workers.[27]

Column 2 of Panel B present results for self-employment. One billion dollars of new oil and gas production increases self-employment by 72, which is about 10.3% of the increase in total employment, as shown in the last column. The estimate on NV_{zt-1} indicates that lagged effect can explain about 11.7% of the employment adjustment ($11.7\% \approx 0.067/0.571 \times 100$). These findings suggest that about 11% of the total employment adjustment comes from self-employed individuals, a group that makes up around 9% of employment over our sample period. Thus, we find that self-employment has made up an economically significant share of the employment adjustment. Theses results suggest that employment fluctuations in establishment based employment data therefore likely underestimates the total fluctuation in employment.

Finally, estimates in Table 2 can be interpreted in the context of a hypothetical one-period shock. But in practice, shocks to do not persist just one year; in fact, when firms begin activity in an area, that activity typically persists for many years. Fig. 3 shows impulse response functions (IRFs) corresponding to each IV regression in Table 2. Each IRF begins with a baseline self-employment indexed to 100 in 2005, and (consistent with prior empirical tests) utilizes the value of production from new wells aggregated up to calendar years (i.e., NV_t) over 2004–2019.[28] Observe that self-employment is clearly pro-cyclical, meaning that it increases during oil and gas booms and contracts during the bust. Thus, this corroborates the prior-discussion based on positive contemporaneous coefficients and negative lagged coefficients. IRFs utilizing the OLS coefficients from Table 2 are presented in Appendix Fig. A1.

[26] Including only one year-lagged value of new production, the estimated coefficients on NV_t and NV_{t-1} are 0.071(0.031) and -0.071(0.033), respectively. Both coefficients are statistically significant at the 5% level.

[27] Feyrer et al. (2017) find that each *million* dollars of new oil and gas production in a county is associated with a contemporaneous increase in 0.85 new jobs within the county. Using their replication files, we find that each *million* dollars of new production in a CZ is associated with a contemporaneous increase in 0.96 new jobs within the CZ. Our estimate, 0.70, is lower than theirs, and the difference *mainly* stems from using a different instrument (see footnote [21]). If we constructed the instrument using their log-linear approach, the corresponding point estimate would be about 1.05.

[28] For example, we construct Fig. 3a first calculating $\triangle E_t = 0.060NV_t - 0.037NV_{t-1} - 0.029NV_{t-2}$ using the point estimates in column 4 in Table 2. Setting $E_{05} = 100$, we then calculate $E_t = E_{t-1} + \triangle E_t$, for $t = 2006, \ldots, 2019$.

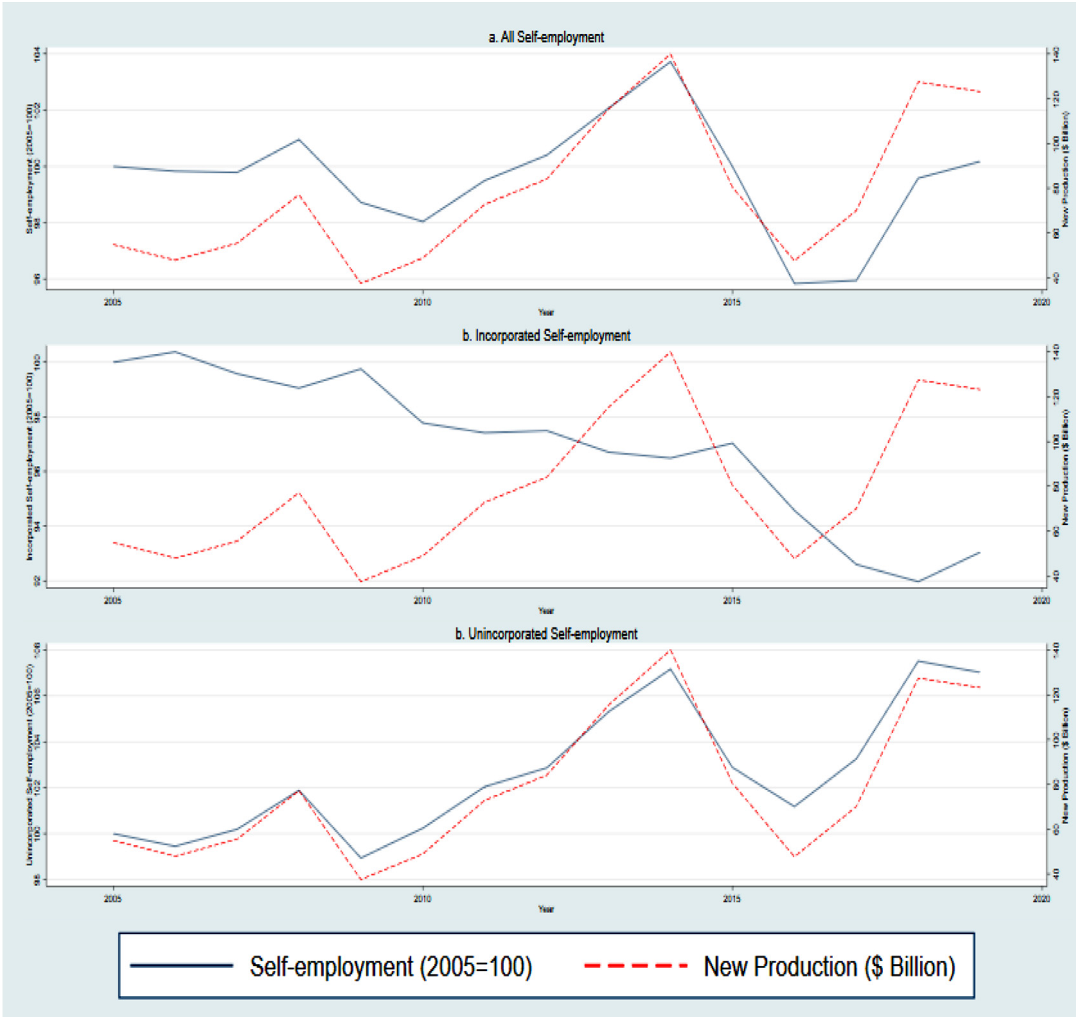

Fig. 3. Impulse response functions based on the IV estimates.

4.2. Industry-level analysis

In this section, we divide our sample into five broad industries: mining, construction, transportation, manufacturing, and other services; and separately estimate the impact of new production on self-employment in each industry. IV results are given in Table 4. All estimates include CZ and year fixed effects and standard errors obtained from bootstrapping (999 replications) are clustered at the CZ level, and observations are weighted by the 2005 CZ employment.

For comparison, Panel A reproduces the results in columns 4–6 of Table 2. Note that the sum of self-employed individuals across industries in each commuting zone gives the total self-employment there, i.e. $\triangle E_{zt} = \sum_k \triangle E_{zt}^k$ where k indexes industry. It then follows from equation (1) that the sum of estimated coefficients on NV_t (or $NV_{t-\ell}$) across industries for any type of self-employment equals the corresponding point estimate on NV_t ($NV_{t-\ell}$) in Panel A.

Panel B reports the impact of new oil and gas production on self-employment in the mining sector. Note that point estimates are small and mostly insignificant, suggesting that the impact of new production on self-employment stems mostly from non-mining sectors. This finding is not surprising because a small share of self-employed individuals works in this sector (see Table 1). Estimates in Panel C indicate that new oil and gas production has not had any effects on self-employment in the transport sector. The new production has a small and insignificant effect on self-employment in construction (see Panel D). Notably, a number of prior students have found that the largest employment effects of the shale boom were observed in the mining, transportation and construction sectors (Feyrer et al., 2017; Marchand, 2012; Upton and Yu, 2021), thus results from this research suggest that industry level effects on self employment are different than employment effects for wage and salaried workers. We also do not find self-employment effects in manufacturing (Panel E).

Finally, Panel F of Table 4 shows that the impact on self-employment mainly stems from the "Other Services" sector. Specifically, one billion dollars of new production increases self-employment by 48, with negative lagged effects. We further disaggregate "Other Services" sector into four sub-sectors: wholesale and retail trade, business and repair services, personal and entertainment services, and professional services (which include education, finance, health, utilities, and other professional services). Our analysis shows that the signs of estimated coefficients in each sub-sector are similar to those in Panel F (see Appendix Table A3). Results reported in this table indicate that effects are more substantially observed in Wholesale & Retail sectors, Personal and Professional Services.

In sum, sectoral level analysis reveals that the largest effects on self employment stem from the services sectors. Thus, although employment effects for wage and salaried workers have been concentrated in the mining, transportation and construction sectors, we find that self-

Table 4
Impact of new oil & gas production on self-employment by industry .

Variable	All 1	Incorp 2	Unincorp 3	All 4	Incorp 5	Unincorp 6
	A. All			B. Mining		
NV_t	0.060	–0.015	0.074*	0.009	0.002	0.007*
	(0.040)	(0.014)	(0.038)	(0.007)	(0.003)	(0.004)
NV_{t-1}	–0.037	0.043***	–0.081	–0.013	–0.000	–0.013
	(0.058)	(0.017)	(0.056)	(0.012)	(0.005)	(0.008)
NV_{t-2}	–0.029	–0.037***	0.009	0.003	–0.001	0.004
	(0.042)	(0.013)	(0.035)	(0.006)	(0.003)	(0.004)
	C. Transport			D. Construction		
NV_t	0.002	0.001	0.002	–0.006	–0.012**	0.006
	(0.006)	(0.005)	(0.006)	(0.015)	(0.005)	(0.014)
NV_{t-1}	–0.007	–0.003	–0.003	0.014	0.017*	–0.003
	(0.010)	(0.008)	(0.012)	(0.025)	(0.010)	(0.022)
NV_{t-2}	0.000	0.002	–0.001	–0.009	–0.009	0.001
	(0.007)	(0.004)	(0.008)	(0.014)	(0.006)	(0.014)
	E. Manufacturing			F. Other Services		
NV_t	0.007	0.002	0.005	0.048	–0.007	0.055**
	(0.010)	(0.005)	(0.005)	(0.030)	(0.012)	(0.028)
NV_{t-1}	–0.010	–0.001	–0.009	–0.022	0.030	–0.052
	(0.018)	(0.009)	(0.010)	(0.043)	(0.019)	(0.041)
NV_{t-2}	0.004	–0.003	0.007	–0.028	–0.027*	–0.002
	(0.009)	(0.005)	(0.006)	(0.031)	(0.015)	(0.023)

Notes: The sample size in each panel is 10,374 observations. All regressions include CZ-fixed and year-fixed effects and observations are weighted by 2005 CZ employment. Industry aggregations based on 1990 Census Bureau industry classification system as reported in IPUMS-CPS. Other services includes wholesale and retail trade, finance and insurance, business and repair services, personal services, entertainment and recreation services, and professional and related services. Standard errors obtained from bootstrapping (with 9999 repetitions) are clustered at the CZ level, and ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5
Impact of new oil & gas production on self-employment: robustness .

Variable	All 1	Incorp 2	Unincorp 3	All 4	Incorp 5	Unincorp 6
	A. Benchmark			B. More Controls		
NV_t	0.060	–0.015	0.074*	0.063	–0.013	0.076*
	(0.040)	(0.014)	(0.038)	(0.041)	(0.014)	(0.039)
NV_{t-1}	–0.037	0.043***	–0.081	–0.041	0.041**	–0.082
	(0.058)	(0.017)	(0.056)	(0.060)	(0.017)	(0.057)
NV_{t-2}	-0.029	–0.037***	0.009	–0.021	–0.033**	0.012
	(0.042)	(0.013)	(0.035)	(0.044)	(0.013)	(0.036)
	C. Spillovers			D. Controlling Migration		
NV_t	0.055**	–0.016	0.071**	0.050	–0.014	0.064*
	(0.025)	(0.014)	(0.028)	(0.033)	(0.012)	(0.031)
NV_{t-1}	–0.035	0.042***	–0.078*	–0.030	0.033**	–0.063
	(0.039)	(0.015)	(0.042)	(0.047)	(0.014)	(0.044)
NV_{t-2}	–0.030	–0.038***	0.008	–0.028	–0.028**	–0.000
	(0.027)	(0.010)	(0.025)	(0.035)	(0.013)	(0.028)
	E. Population Trends			F. Leads		
NV_t	0.039	–0.022	0.061**	0.009	–0.057	0.066
	(0.027)	(0.015)	(0.028)	(0.060)	(0.040)	(0.059)
NV_{t-1}	0.066	0.077*	–0.011	–0.006	0.068*	–0.075
	(0.098)	(0.044)	(0.064)	(0.063)	(0.035)	(0.070)
NV_{t-2}	–0.096	–0.060*	–0.036	–0.040	–0.047***	0.007
	(0.085)	(0.035)	(0.055)	(0.040)	(0.016)	(0.039)
NV_{t+1}				0.041	0.034	0.006
				(0.052)	(0.029)	(0.034)

Notes: The sample size is 10,374 in Panels A-E, and 9,633 in Panel F. All regressions include CZ-fixed and year-fixed effects and observations are weighted by 2005 CZ employment. Standard errors obtained from bootstrapping (with 999 repetitions) are clustered at the CZ level, and ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

employment effects are concentrated in other services sectors, specifically Wholesale & Retail trade and Personal and Professional Services. This the sectors of adjustment for self-employment are different than the sectors of the overall employment adjustment.

4.3. Discussion of channels

We now turn to discuss possible channels through which an oil and gas boom can affect self-employment. First, local landowners receive bonus and royalty checks for oil and gas production that occurs beneath their land. A bonus check is given to the landowner at the time that a lease is signed as a lump sum payment. But also, once production begins landowners receive royalty payments that is some share of the value of the oil and gas produced (typically 20–25%).[29] These royalty payments may continue for a short time if the well is relatively unsuccessful, or can continue for years and even decades as the well continues along its tail of production. Thus, when local residents receive payments this might provide resources needed to start a business. This mechanism is explored in Bellon et al. (2021) who find that individuals who receive wealth shocks of $50,000 or greater have about 60% greater self-employment rate relative to individuals who receive small or no wealth shocks. In the present context, royalty payments might reduce liquidity constraints and spur entrepreneurship (Levine and Rubinstein, 2018).

The second channel through which oil and gas operators can stimulate a local economy is through the broader economic activity generated. First, when landowners receive bonus and royalty payments, they will likely spend some share of these in the local economy, providing stimulus. Second, the drilling activities themselves will also boost employment and earnings. In the case of the shale plays, the operating company typically contracts out a service company to both drill the well and complete the hydraulic fracturing needed to stimulate the well to begin production. These workers will earn income directly, and then spend some share of these earnings in the local economy.

We are unable to distinguish between these two channels, and both channels could impact self-employment across sectors. This is a general limitation of the broader literature focusing on the economic implications of localized oil and gas activity.

5. Robustness

In this section, we present sensitivity checks to investigate the robustness of our finding to additional controls, spatial spillovers, migration, among others. Robustness checks to our main specification are presented in Table 5, and results from industry-level analyses are qualitatively the same (and available upon request). To facilitate comparison, Panel A reports the benchmark results.

Additional controls

Panel B of Table 5 reports results with additional CZ-level time varying controls included in our specification shown in equation (1). Control variables include the lagged CZ demographic composition (shares of whites, the foreign-born, the college educated, and those ages 16–24, 25–39, and 40–64 in the adult population) and the lagged CZ employment composition (share of employment among women and unemployment rate). As mentioned earlier, these are not included in the main

[29] The surface owner of the land where the actual well is drilled typically receives a "rental" payment that is the value of renting the surface area needed to drill and produce. Most landowners, though, receive a bonus and royalty payments even though no actually drilling activity occurred on their land.

specification because they are potentially endogenous variables. Their inclusion, however, does not have any meaningful impact on estimates.

Spatial spillovers

We test the sensitivity of our results to potential spatial spillovers. Specifically, CZs that are in states with shale activity but that themselves do not overlap with EIA's seven major shale plays highlighted in the *Drilling Productivity Reports* are removed from the sample. In addition, states that directly border counties with shale activity were removed.[30] Prior studies have also found that significant midstream and downstream investments occurred in response to the shale boom (Dismukes and Upton, 2022).

Because these spatial spillovers created similar "boom town" effects in areas without shale production, inclusion of these ares has the potential to bias point estimates downward. For example, Lake Charles Louisiana was the fastest employment growth MSA in the country from 2013 to 2018, and had no shale activity. But, the MSA underwent billions of dollars of capital expenditure, in chemical manufacturing and the export of natural gas in the form of liquefied natural gas (LNG) that was made possible by oil and gas extraction growth in shale regions (Scott and Upton, 2019). Panel C report results, which are very similar to those in Panel A.

Labor migration

Areas experiencing economic booms generally induce in-migration, which naturally changes population level and demographic composition. Self-employment changes can come from both self-employed workers migrating into and out of boom areas, or due to individuals living in these areas becoming self-employed. This is perhaps one way in which a regional boom bust cycle can differ from a nation-wide business cycle.

To partially address whether results might be driven by migration, we next conduct an additional robustness check dropping recent movers from the sample. The American Community Survey (ACS) is a repeated cross sample and therefore we cannot track individuals over time. However, the ACS ask individuals whether they had lived in the "same house" or a "different house" one year earlier. Those who had moved indicate whether they moved within state, between states, or were abroad one year ago. We investigate how the new oil and gas production has affected self-employment among non-movers, and the results are reported in Panel D of Table 5. Point estimates are similar to those in the benchmark results in Panel A, suggesting that effects are not driven by migration.

Population trends

As discussed in Section 3 the dependent variable of interest in the change in employment normalized to the employment in each CZ in 2005, defined as $\triangle E_{zt} = (E_{zt} - E_{z,t-1})/L_{z,2005}$. We choose to normalize by 2005 employment, as labor migration can cause population to increase in response to the boom. As an additional robustness, we normalize the change in self-employment to the employment in the prior year, $\triangle E_{zt} = (E_{zt} - E_{z,t-1})/L_{z,t-1}$. Results shown in Panel E remain qualitatively similar to those in the benchmark case, although the effects are less precisely estimated (especially for incorporated self-employment).

Including leads

We next perform a test by including a lead value of new oil and gas production. The value of production in $t+1$ should not have an impact on self-employment in year t. Panel F reports the results from this exercise, and note that the estimated coefficient on the lead value is statistically insignificant. Point estimates on other variables become larger for incorporated self-employment, and remains mostly the same for unincorporated ones. Overall, including the lead value does not change the main conclusion we have in the benchmark case.

Table 6
Impact of new production on self-employment - great recession robustness .

Variable	Benchmark Results			Impact of Recession		
	All 1	Incorp 2	Unincorp 3	All 4	Incorp 5	Unincorp 6
NV_t	0.060	–0.015	0.074**	0.064	–0.006	0.070**
	(0.040)	(0.014)	(0.038)	(0.042)	(0.018)	(0.034)
$R\times NV_t$				0.047	–0.021	0.068
				(0.136)	(0.065)	(0.085)
NV_{t-1}	–0.037	–0.043***	–0.081	–0.042	0.032	–0.074
	(0.058)	(0.017)	(0.056)	(0.066)	(0.024)	(0.051)
$R\times NV_{t-1}$				–0.055	0.059	–0.114
				(0.102)	(0.072)	(0.117)
NV_{t-2}	–0.029	–0.037***	0.009	–0.030	–0.032*	0.002
	(0.042)	(0.013)	(0.035)	(0.046)	(0.017)	(0.032)
$R\times NV_{t-2}$				–0.076	–0.070	–0.006
				(0.125)	(0.075)	(0.073)

Notes: The sample size in each panel is 10,374 observations from 741 U.S. commuting zones over the 2006–2019 period (includes 2005 with lagged value). R is a dummy variable that equals one during 2008–2010 recession, and zero otherwise. All regressions include CZ-fixed and year-fixed effects and observations are weighted by 2005 CZ employment. Standard errors obtained from bootstrapping (with 999 repetitions) are clustered at the CZ-level, and ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

The Great Recession

As a final robustness check, we test the sensitivity of our results to considering the effect of the Great Recession. During the Great Recession national unemployment increased sharply, investment fell substantially, and many businesses exited the market (Christiano et al., 2015). It was particularly hard for small businesses because financial institutions generally consider small business lending riskier than larger firms (Duygan-Bump et al., 2014). As shown in Fig. 1, both incorporated and unincorporated self-employment rates declined during the Great Recession.

Prior literature documenting labor market effects of the shale boom have largely focused on the initial shale boom coinciding with the Great Recession.[31] While these results are notable in their own right, this timing generally calls into question the applicability of results in calibrating models to understand relationships during more "normal" times. For instance, the national unemployment rate peaked at 10 percent during 2009, coinciding almost exactly with when U.S. oil production began to increase from its almost 40 year trough in 2008. Following a recession, the aggregate labor market is slack, and therefore had this shock occurred at a time with a tighter labor market, magnitudes of impacts across outcomes might be very different.

To examine the impact during the Great Recession, we interact the value of production with an indicator variable (denoted R for *recession*) that equals one for 2008–2010, and zero otherwise. These results are presented in Table 6. For easy comparison, the benchmark result from Table 2 is presented in columns 1–3 with results including recession interaction effects in columns 4–6. All coefficients for recession interactions are imprecisely estimated, and the estimates on other coefficients remain mostly the same. Thus, results of this robustness check

[30] After applying these criteria, the following non-shale boom states are included: AK, AZ, CA, CT, DE, FL, GA, HI, ID, IL, IA, ME, MI, MN, MS, MO, NV, NH, NJ, NC, OR, RI, SC, TN, VT, WA, and WI. For a more detailed description see McCollum and Upton (2018), Decker et al. (2021) and Upton and Yu (2021).

[31] For instance, even the more recent studies Bartik et al. (2019); Feyrer et al. (2017); Tsvetkova and Partridge (2017) utilize data until 2014 at the latest.

suggest that the benchmark result is not driven by the recession time period.

6. Conclusion

The U.S. shale boom has given rise to a large literature studying the economic effects of natural resource shocks. Although the literature has extensively investigated the effects on labor markets, limited attention has paid to how this revolution has affected entrepreneurship. The shale oil and gas boom is also of interest to economist broadly, as it provides a unique opportunity to assess the cyclicality of self employment across the business cycle.

Taking advantage of these plausibly exogenous shocks to local labor markets, we present empirical evidence that self-employment increases in response to oil and gas activity. Specifically, we find that new oil and gas production has a positive and significant contemporaneous impact on self-employment, and results are mainly driven by private service sectors. Second, we find that self-employment is pro-cyclical, i.e. it increases during oil and gas booms and contracts during the bust. Finally, our findings indicate that the impact on self-employment can explain a significant portion of the employment adjustment. In particular, about 11% percent of the employment adjustment comes from self-employed business owners.

Data availability

Data will be made available on request.

Appendix

Table A1
Number of wells opened in the United States, 2004–2019.

	Wells Count
Anadarko	28,673
Appalachia	74,149
Bakken	18,385
Eagle Ford	3592
Haynesville	18,187
Niobrara	70,222
Permian	66,626
Other Areas	184,359
Total	492,193

Notes: Data from Enverus and EIA’s Drilling Productivity Reports.

Table A2
Distribution of value of new production by commuting zone.

	Total Value	Annual Value
1st percentile	$.01	$0
5th percentile	$1.3	$0.1
10th percentile	$10.3	$0.7
25th percentile	$196	$14
50th percentile	$2,715	$194
75th percentile	$16,261	$1,162
90th percentile	$102,348	$7,311
95th percentile	$226,760	$16,197
99th percentile	$447,858	$31,910
Mean	$36,978	$2,642
Max	$841,362	$60,097

Notes: Values are in millions of 2015 dollars and represent the sum of all production from new wells over the sample period 2005–2019 by commuting zone. There are 337 CZs with some oil and gas production from wells that are less than one year of production.

Table A3
Impact of new oil & gas production on self-employment in service sectors .

Variable	All 1	Incorp 2	Unincorp 3	All 4	Incorp 5	Unincorp 6
	A. Wholesale & Retail			B. Business & Repair		
NV_t	0.013	–0.001	0.015	0.002	–0.005	0.007
	(0.011)	(0.006)	(0.010)	(0.012)	(0.005)	(0.011)
NV_{t-1}	–0.012	0.008	–0.020	0.010	0.009	0.002
	(0.015)	(0.010)	(0.015)	(0.018)	(0.008)	(0.015)
NV_{t-2}	0.000	–0.004	0.005	–0.020*	–0.006	–0.014
	(0.009)	(0.006)	(0.009)	(0.011)	(0.005)	(0.009)
	C. Personal Services			D. Professional Services		
NV_t	0.013	0.000	0.013	0.020	–0.001	0.021
	(0.009)	(0.005)	(0.009)	(0.016)	(0.010)	(0.014)
NV_{t-1}	–0.003	0.003	–0.007	–0.017	0.010	–0.027
	(0.015)	(0.006)	(0.015)	(0.025)	(0.017)	(0.024)
NV_{t-2}	–0.007	–0.005	–0.002	–0.002	–0.011	0.010
	(0.010)	(0.004)	(0.009)	(0.017)	(0.011)	(0.015)

Notes: The sample size in each panel is 10,374 observations. All regressions include CZ-fixed and year-fixed effects and observations are weighted by 2005 CZ employment. Standard errors obtained from bootstrapping (with 9999 repetitions) are clustered at the CZ-level, and ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

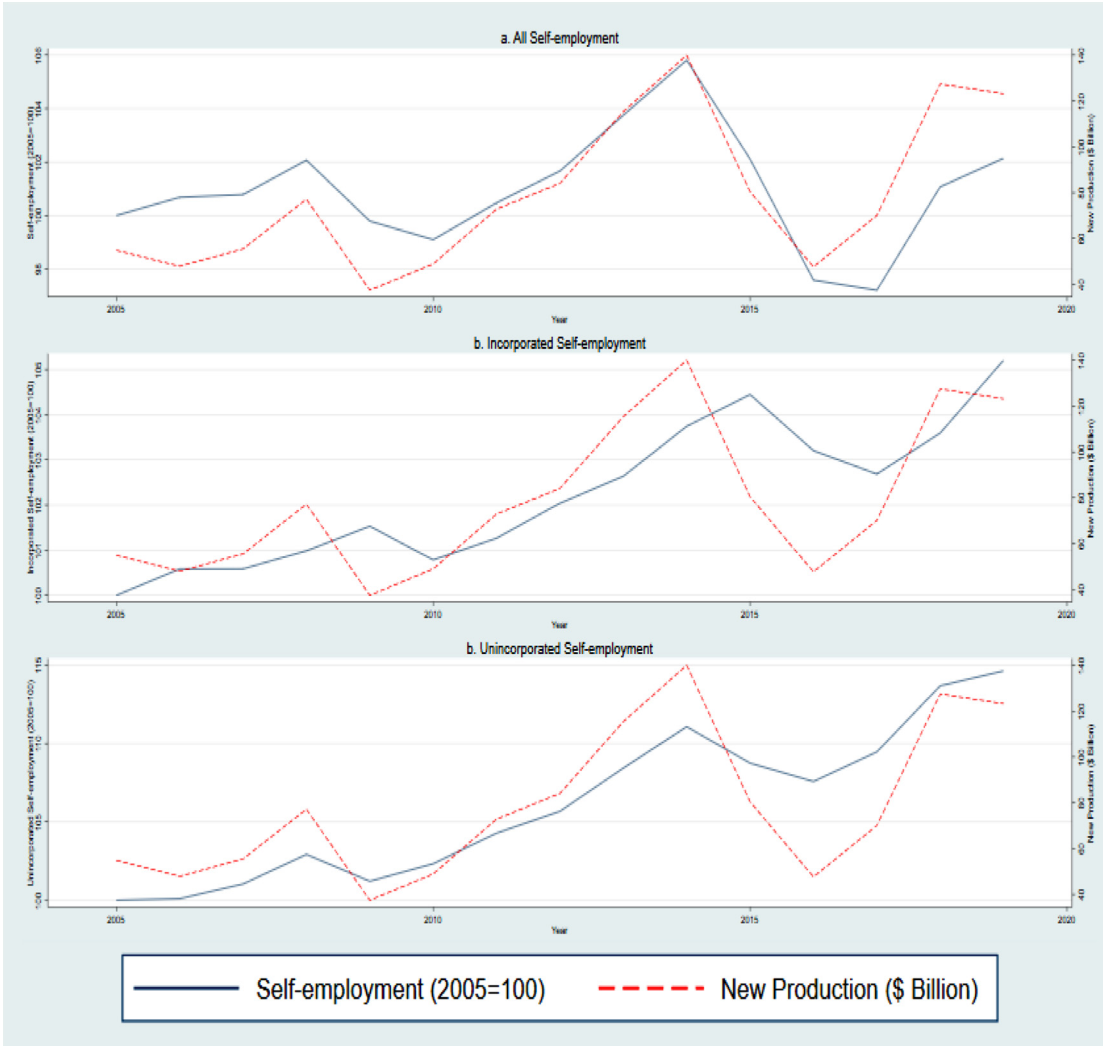

Fig. A1. Impulse response functions based on the OLS estimates.

References

Abraham, K.G., Haltiwanger, J., Sandusky, K., Spletzer, J., 2019. The rise of the gig economy: fact or fiction? AEA Papers and Proceedings 109, 357–361.

Agerton, M., Hartley, P.R., Medlock III, K.B., Temzelides, T., 2017. Employment impacts of upstream oil and gas investment in the United States. Energy Econ. 62, 171–180.

Agerton, M., Upton, G.B., 2019. Decomposing crude price differentials: domestic shipping constraint or the crude oil export ban. Energy Journal 40, 155–172.

Alexeev, M., Conrad, R., 2009. The elusive curse of oil. Review of Economics and Statistics 91, 586–598.

Allcott, H., Keniston, D., 2018. Dutch disease or agglomeration? the local economic effects of natural resource booms in modern america. Review of Economic Studies 85, 695–731.

Anderson, S.T., Kellogg, R., Salant, S.W., 2018. Hotelling under presure. Journal of Political Economy 126, 984–1026.

Angrist, J.D., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.

Asiedu, E., Freeman, J.A., Nti-Addae, A., 2012. Access to credit by small businesses: how relevant are race, ethnicity, and gender? American Economic Review 102 (3), 532–537.

Autor, D.H., Dorn, D., 2013. The growth of low skill service jobs and the polarization of the u.s. labor market. American Economic Review 103, 1553–1597.

Bartik, A.W., Currie, J., Greenstone, M., Knittel, C.R., 2019. The local economic and welfare consequences of hydraulic fracturing. American Economic Journal: Applied Economics 11 (4), 105–155.

Beland, L.-P., Unel, B., 2019. Politics and entrepreneurship in the US. Canadian Journal of Economics 52 (1), 33–57.

Bellon, A., Cookson, J.A., Gilje, E.P., Heimer, R.Z., 2021. Personal wealth and self-employment. Review of Financial Studies.

Black, D., McKinnish, T., Sanders, S., 2005. The economic impact of the coal boom and bust. Economic Journal 115 (503), 449–476.

Blanchflower, D.G., Oswald, A.J., 1998. What makes and entrepreneur? Journal of Labour Economics 16, 26–60.

Borjas, G.J., Bronars, S.G., 1989. Consumer discrimination and self-employment. Journal of Political Economy 97 (3), 581–605.

Brown, J., Fitzgerald, T., Weber, J.G., 2019. Does resource ownership matter? oil and gas royalties and the income effect of extraction. Journal of the Association of Environmental and Resource Economists 6 (6), 1039–1064.

Brown, J.P., 2015. The response of employment to changes in oil and gas exploration and drilling. Economic Review-Federal Reserve Bank of Kansas City 57–81.

Cagetti, M., De Nardi, M., 2006. Entrepreneurship, frictions, and wealth. Journal of Political Economy 114 (5), 835–870.

Christiano, L.J., Eichenbaum, M.S., Trabandt, M., 2015. Understanding the great recession. American Economic Journal: Macroeconomics 7, 110–167.

Correia, S., Guimaraes, P., Zylkin, T., 2020. Fast poisson estimation with high-dimensional fixed effects. STATA Journal 20, 95–115.

Cullen, J.B., Gordon, R.H., 2007. Taxes and entrepreneurial risk-taking: theory and evidence for the u.s. J Public Econ 91 (7), 1479–1505.

De Nardi, M., Doctor, P., Krane, S.D., 2007. Evidence on entrepreneurs in the United States: Data from the 1989–2004 Survey of Consumer Finances. Technical Report. Federal Reserve Bank of Chicago,.

Decker, R.A., McCollum, M., Upton Jr., G.B., 2021. Firm dynamics and local economic shocks: evidence from the shale oil and gas boom. Journal of Human Resources - Forthcoming.

Dinopoulos, E., Unel, B., 2015. Entrepreneurs, jobs, and trade. Eur Econ Rev 79, 93–112.

Dismukes, D.E., Upton, G.B., 2022. 2023 Gulf Coast Energy Outlook. Technical Report. Louisiana State University.

Duygan-Bump, B., Levkov, A., Montoriol-Garriga, J., 2014. Financial constraints and unemployment: evidence from the great recession. J Monet Econ 75, 89–105.

EIA, 2011. Review of Emerging Resources: U.S. Shale Gas and Shale Oil Plays. Technical Report. U.S. Energy Information Administration.

Eren, O., Onda, M., Unel, B., 2019. Effects of FDI on entrepreneurship: evidence from right-to-work and non-right-to-work states. Labour Econ 58, 98–109.
Fairlie, R.W., 2014. Kauffman Index of Entrepreneurial Activity 1996–2013. Technical Report. Ewing Marion Kauffman Foundation.
Fairlie, R.W., Meyer, B.D., 2003. The effect of immigration on native self-employment. J Labor Econ 21 (3), 619–650.
Fairlie, R.W., Miranda, J., Zolas, N., 2019. Measuring job creation, growth, and survival among the universe of start-ups in the united states using a combined start-up panel data set. ILR Review 72, 1262–1277.
Feyrer, J., Mansur, E.T., Sacerdote, B., 2017. Geographic dispersion of economic shocks: evidence from the fracking revolution. American Economic Review 107 (4), 1313–1334.
Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, J.R., 2020. Integrated Public Use Microdata Series, Current Population Survey: Version 7.0. Technical Report. Minneapolis, MN: IPUMS.
Fossen, F.M., 2020. Self-employment over the business cycle in the USA: a decomposition. Small Business Economics doi:10.1007/s11187-020-00375-3.
Gourieroux, C., Monfort, A., Trognon, A., 1984. Pseudo maximum likelihood methods: theory. Econometrica 52, 681–700.
Green, D.A., Morissette, R., Sand, B.M., 2019. Economy-wide spillovers from booms: long-distance commuting and the spread of wage effects. J Labor Econ 37 (S2), S643–S687.
Haltiwanger, J., Jarmin, R.S., Miranda, J., 2013. Who creates jobs? small versus large versus young. Review of Economics and Statistics 95, 347–361.
Hurst, E., Lusardi, A., 2004. Liquidity constraints, household wealth, and entrepreneurship. Journal of Political Economy 112 (2), 319–347.
Jacobsen, G.D., 2019. Who wins in an energy boom? evidence from wage rates and housing. Econ Inq 57, 9–32.
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.
Levine, R., Rubinstein, Y., 2017. Smart and illicit: who becomes an entrepreneur and do they earn more? Quarterly Journal of Economics 132 (2), 963–1018.
Levine, R., Rubinstein, Y., 2018. Selection into Entrepreneurship and Self-Employment. Working Paper: 25350. NBER.
Maniloff, P., Mastromonaco, R., 2017. The local employment impacts of fracking: a national study. Resource and Energy Economics 49, 62–85.
Marchand, J., 2012. Local labor market impacts of energy boom-bust-boom in western canada. J Urban Econ 71 (1), 165–174.
Marchand, J., Weber, J., 2018. Local labor markets and natural resources: asynthesis of the literature. J Econ Surv 32 (2), 469–490.
Martin, P.H., Yeates, J.L., 1992. Louisiana and texas oil & gas law: an overview of the differences. LA Law Rev 52 (4), 769–860.
McCollum, M., Upton, G.B., 2018. Local labor market shocks and residential mortgage payments: evidence from shale oil and gas booms. Resource and Energy Economics 53, 162–197.
Michaels, G., 2010. The long-term consequences of resource-based specialisation. Economic Journal 121, 31–57.
Newell, R.G., Prest, B.C., 2019. The unconventional oil supply boom: aggregate price response from microdata. The Energy Journal 40.
Newell, R.G., Prest, B.C., Vissing, A., 2016. Trophy Hunting vs. Manufacturing Energy: The Price- Responsiveness of Shale Gas. Technical Report. Resources for the Future.
Oliver, M., Upton, G.B., 2022. Are energy endowed countries responsible for conditional convergence? Energy Journal 43 (3), 201–224.
Pekkala Kerr, S., Kerr, W., 2020. Immigrant entrepreneurship in america: evidence from the survey of business owners 2007 and 2012. Res Policy 49 (3), 103918.
van der Ploeg, F., 2011. Natural resources: curse or blessing? J Econ Lit 49 (2), 366–420.
Scott, L.C., Upton, G.B., 2019. Louisiana Economic Outlook: 2020 and 2021. Technical Report. Louisiana State University Economics & Policy Research Group.
Silva, J.M.C.S., Tenreyro, S., 2006. The log of gravity. Rev Econ Stat 88.
Silva, J.M.C.S., Tenreyro, S., 2011. Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. Econ Lett 112.
Tolbert, C.M., Sizer, M., 1996. U.S. Commuting Zones and Labor Market Areas. A 1990 Update. Technical Report 9614. Rural Economy Division, Economic Research Service, U.S. Department of Agriculture.
Tsvetkova, A., 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.
Upton, G.B., Yu, H., 2021. Local labor market shocks and employment and earnings differentials: evidence from shale oil and gas booms. Energy Econ. 102 (105462).
Venables, A.J., 2016. Using natural resources for development: why has it proven so difficult? Journal of Economic Perspectives 30, 168–184.
Weber, J.G., 2012. The effects of a natural gas boom on employment and income in colorado, texas, and wyoming. Energy Econ. 34 (5), 1580–1588.