# Local Labor Demand Shocks and Earnings Differentials: Evidence from Shale Oil and Gas Booms.*

Gregory B. Upton Jr.† Han Yu‡

JEL Codes: J16, J23, J24, J31, Q33

October 25, 2019

**Abstract**

In this paper, we investigate the impact of the shale boom on local labor market outcomes. Using a difference-in-difference approach, we show that the boom significantly affects employment and earnings of various groups of workers. We find a significant impact of the boom on employment and earnings differentials between high- and low-skilled workers, as well as between male and female workers. The results show that the impact spillovers to sectors that are not directly impacted by the productivity shock. The results highlight the importance of considering differential effects of technology shocks by education and gender in studying earnings inequality.

*We thank Martin Stuermer for providing valuable comments at 2018 ASSA, Erin Mansur for providing data and feedback, and LSU economics lunch seminar participants. We also thank Naci Mocan for his continued and invaluable feedback.

†Gregory B. Upton Jr.: Center for Energy Studies, Louisiana State University. Email: gupton3@lsu.edu

‡Han Yu: Private Enterprise Research Center, Texas A&M University. Email: yu.han@tamu.edu

# 1 Introduction

Technological developments have made the extraction of previously economically inaccessible energy resources feasible at prevailing market prices. Specifically, the advent of horizontal drilling and hydraulic fracturing techniques have created historic increases in production of oil and natural gas. This created an economic boom in specific geographic areas with oil and gas-rich "shale" geological formations thousands of feet below the surface.

This research focuses on how these natural resource booms impacted employment and earnings differentials between groups of workers including (1) college/high school education levels and (2) males/females.[1] We examine seven geographic areas that were plausibly exogenously located above shale geological formations, namely: *Appalachia*, *Anadarko*, *Bakken*, *Eagle Ford*, *Haynesville*, *Niobrara*, and *Permian*.[2] Appendix Figure A.1 shows the location of these "shale plays."

For decades, earnings inequality has been a focus of the labor economics literature, both focusing on differentials across the income spectrum (Mincer, 1970; Maddison, 1987; Levy and Murnane, 1992; Katz, 1999) and male / female differentials (Blau and Kahn, 2017; Goldin, 2014; O'Neill, 2003; Gunderson, 1989). These dimensions of income inequality have experienced different trends and have been impacted by different factors over the past century in the United States.

Levy and Murnane (1992) describe distinct time periods of changes in income inequality throughout modern history. Since the 1980s, the U.S. economy has experienced increases in income inequality that has persisted to the present day (Attanasio et al., 2012). In contrast, men and women have experienced convergence in earnings, and this has been called the "Grand Gender Convergence" (Goldin, 2014). While differences in pay for men and women still persist, typically at least two-thirds of this differential can be explained by factors such as occupation differences (Blau and Kahn, 2017), career interruptions and hours worked per week (Bertrand et al., 2010), inter-firm mobility (Bono and Vuri, 2011), among others.[3]

Using a specific labor demand shock to a male and blue-collar dominated industry, we provide estimates of the impact of a technology induced labor demand shock on employ-

[1]Hereafter we refer to college-educated workers as all workers with a college degree or more. We refer to high school educated workers as workers with a high school degree or less.

[2]According to EIA, more than 90% of oil production growth as well as all natural gas production growth in the U.S. during 2011-2014 are attributed to these seven regions.

[3]Residuals in earnings differentials that cannot be explained by these factors are typically then attributed to psychological attributes, unobservable non-cognitive skills and/or discrimination. (Blau and Kahn, 2017; Manning and Swaffield, 2008; Wood et al., 1993).

ment and earnings differentials between workers with college/high school educations and males/females.[4] We argue that the oil and gas shale boom of the past decade creates a unique opportunity to study this for three reasons.

First, the shale boom originated from a technology induced labor demand shock. Second, the labor demand shock overwhelmingly directly impacted one demographic of workers, male workers with a high school education or less.[5] Third, the shock is conveniently concentrated in very specific geographic areas that happened to have specific geological formations thousands of feet below the earth's surface. And the timing of these shocks all coincided with technological advancements alongside high oil and natural gas prices that allowed for extraction from these formations. This allows us to identify areas that received the treatment and still have access to plausible control areas with similar pre-treatment characteristics.

We index empirical estimates to the value of oil and gas produced and direct employment shocks utilizing an instrumental variable strategy employed in Feyrer et al. (2017). We produce parameter estimates that might be used more generally in non-oil and gas contexts to understand the sensitivity of earnings differentials to labor demand shocks to sub-sets of workers. We discuss the extent to which empirical point estimates have external validity in contexts outside of localized oil and gas booms.

Finally, we decompose the observed changes in earnings differentials into three channels. First, we observe earnings differentials change within the mining sector. Second, we observe earnings differentials change within non-mining sectors. Third, earnings differentials can be effected by labor migration (into these geographic areas) and substitution of workers across sectors in response to the increased earnings opportunities. Interestingly, we show that the majority of the change in earnings differentials comes from the second channel; changes in earnings differentials within non-mining sectors.

**Economic Impact of Natural Resource Booms** Over the last two decades, the oil and natural gas landscape has changed both suddenly and dramatically. By the mid to late 2000s, after decades of declining production, technological breakthroughs alongside high oil and natural gas prices allowed oil and gas to be extracted from shale geological formations;

[4]Throughout this research we refer to the college/high school employment and earnings differentials. Specifically, we define college as all workers with a college degree or more. We define high school as all workers with a high school diploma or less.

[5]Of course, many female workers and workers with college degrees are employed by the upstream oil and gas sector, but these jobs are primarily office positions in larger cities such as Houston or Oklahoma City, where these companies' headquarters are located. The areas of interest in this study are where the hydrocarbons themselves are actually extracted.

the shale boom was underway.[6] Through a combination of horizontal drilling and hydraulic fracturing (informally referred to as "fracking") the U.S. is now experiencing production at levels not seen since "peak oil" of the 1970s. There has been a growing body of work that quantifies the effects of localized natural resource-based booms. While this literature began before the specific shale boom of this past decade (Black et al., 2005), this new *Era of Shale* has created a significant resurgence in this literature.

Feyrer et al. (2017) find that the shale boom created significant economic shocks to local labor markets. Every million dollars of oil and gas extracted is estimated to generate $243,000 in wages, $117,000 in royalty payments, and 2.49 jobs within a 100-mile radius. In total, the authors estimate that the shale boom was associated with 725,000 jobs in aggregate and a 0.4 percent decrease in the unemployment rate during the Great Recession. Marchand (2012) similarly finds both direct and indirect impacts of production from shale on employment; for every 10 jobs created in the energy sector, 3 construction, 4.5 retail, and 2 service jobs are created. Agerton et al. (2017) find that one additional rig results in the creation of 31 jobs immediately and 315 jobs in the long-run. Other studies corroborated the positive impact of the shale boom on local labor markets (Weber, 2012; Cosgrove et al., 2015; Marchand and Weber, 2018; Komarek, 2016; Bartik et al., 2019; McCollum and Upton, 2018; Decker et al., 2018).

Several recent analyses have exploited natural resource booms as productivity shocks to male and/or high school educated workers in order to study earnings and educational inequality. Cascio and Narayan (2017) exploit the labor demand shock associated with the shale boom to less-educated male labor and finds that this narrowed the male-female gap in teen high school dropout rates by nearly 40%. Related, Aragon et al. (2018) exploit the closing of coal mines in the UK to study employment substitutions across sectors. They provide evidence that men and women are imperfect substitutes for labor in non-primary sectors and therefore a shock to the mining sector can impact employment and earnings in non-mining industries. Similarly, Kotsadam and Tolonen (2016) utilize data from Africa to test the effect of mining on local economies, and find that female employment decreases in response to mining booms that increase male employment.

[6] The official start date of the shale boom can be debated. In the early 2000s, extraction of natural gas in the Barnett Shale in Texas began. But it was not until between 2007 and 2009 that oil and natural gas production led to significant increases in aggregate U.S. production. For our baseline specifications we consider the official start date of the shale boom as 2007, consistent with the beginning of EIA's drilling productivity reports.

# 2 Data

Data on employment and earnings are from the United States Census' Quarterly Workforce Indicators (QWI). QWI contains information on county-level average employment and earnings. We utilize a yearly panel of counties from 2004 until 2013 consistent with Feyrer et al. (2017).[7]

In order to mitigate potential concerns of spatial spillover effects (James and Smith, 2019), we do not include counties that are in geographic proximity to shale counties as potential control counties. Specifically, we remove all counties in shale states, but that themselves are not included in EIA's definitions of shale counties. In addition, states that directly border counties with shale activity are removed from the potential control group[8] and therefore control groups is chosen from non-shale counties in states that did not experience shale activity. We utilize propensity score matching to identify a control group of counties from across the U.S. that are not in proximity to shale counties whose demographic characteristics are similar to the boom counties in the pre-boom time.

County-level data on new wells drilled and the value of new production from new wells come from data published alongside Feyrer et al. (2017), that was initially sourced from *DrillingInfo*. The price of crude oil and natural gas used for calculating the value of total production are WTI crude oil spot price and Henry Hub natural gas spot price also taken directly from Feyrer et al. (2017), initially sourced from EIA. For all difference-in-differences estimates, we consider 2007 the beginning of the treatment time period consistent with when EIA's Drilling Productivity reports began tracking production and rig counts in shale regions.[9] We show that results are robust across shale plays, different control groups, and accounting for differential intensity of activity across both shale plays and time.

Table 1 show the labor market characteristics in counties with shale oil and/or gas activity

[7] Feyrer et al. (2017) are able to extend their analysis to 2014, but due to limitations in QWI data, the fourth quarter of 2013 is the most recent consistently available quarter at the time of analysis. While many states and sub-sectors within states do have more recent data, we used this cutoff in order to preserve a balanced panel across states and relevant sectors.

[8] After applying these decision rules, control counties are pulled from the following 28 states: AL, AZ, CA, CT, DE, FL, GA, HI, ID, IL, IN, IA, ME, MI, MN, MS, MO, NV, NH, NJ, NC, OR, RI, SC, TN, VT, WA, and WI. It should be noted that this conservative approach of choosing control counties comes with both costs and benefits. It allows us to rule out, for all intents and purposes, the possibility of spillovers into non-shale counties that are in close proximity. But, on the other hand, precludes us from considering spatial spillovers as has been done in other research. Given the research question at hand and the prior work of spatial spillovers, we choose the conservative approach and save any analysis of spatial spillovers of earnings differentials for future research.

[9] Other papers utilizing a DD specification choose different starting dates. For instance, Cascio and Narayan (2017) use 2006, while McCollum and Upton (2018) use 2007.

compared to the propensity score matched control group. More specifically, Table 1 shows the change in employment and earnings for the treated (shale) counties and the propensity score matched control counties. We present average employment and earnings in the pre-shale (2004 to 2006) and post shale (2007 to 2013) time periods. In the control counties, overall employment decreased by an average of 4.42 percent between these time periods. This is unsurprising as the Great Recession of 2009 coincided with the time period of the shale boom. In the shale counties, employment remained relatively flat, slightly decreasing by around 0.07 percent. Earnings, on the other hand, increased in the control areas by about 13 percent in nominal dollars between the pre-and-post 2007 time periods, while earnings in shale areas increased by an even larger 22 percent. Thus, earnings growth in shale areas outpaced non-shale area earnings growth by about 9.14 percentage points.

Table 1 also breaks down these relative changes by demographic of workers. We point out three notable items. First, the relative employment and earnings growth are largest for workers with a high school diploma or lower and male workers. More specifically, male workers and workers with a high school diploma or less experienced a 6.1 and 4.8-percentage point faster increase in employment and a 9.6 and 11.2-percentage point faster increase in earnings relative to control groups. Second, we find relatively large increases in earnings across all demographics of workers in shale areas. In fact, even female workers with a college degree, the group least directly affected by the shale boom, experienced a 4.6-percentage point increase in earnings relative to control areas. Third, is the relative changes in employment and earnings in mining compared to non-mining sectors. Specifically, we estimate that mining sector employment increased by around 41% in shale counties relative to non-shale control counties, while non-mining sector employment increased by less than 4% relative to non-shale sectors. Earnings, similarly, increased by more than 15.5% and 8.6% respectively in the mining and non-mining sectors.

All three of these observations lead to a simple conclusion. While the largest percent increases in employment were experienced by male workers with a high school diploma in the mining sector, earnings increased across all subsets of workers. These results are corroborated in the regression framework presented below.

# 3 Empirical Specification

## 3.1 Difference-in-Differences

As a first empirical test, we utilize a difference-in-differences (DD) approach to test for the impact of shale oil and gas booms on local labor markets. Specifically, we consider the following specification:

$$Y_{c,t} = \beta_0 + \beta_1(Shale_C * Shale_T) + \lambda_c + \alpha_t + \epsilon_{i,t} \tag{1}$$

where $Y_{c,t}$ is the outcome of interest, including employment, earnings, employment differentials and earnings differentials, in county $c$ and year $t$. We consider employment and earnings of workers across sectors, educational level, gender, and seven shale plays geographically dispersed across the U.S. $Shale_C$ is an indicator equal to 1 if county $i$ is a county located within one of the seven key shale regions; otherwise, $Shale_C$ equals to 0. Similarly, $Shale_T$ is a dummy variable indicating the beginning of the shale boom. For purposes of this first specification, and for simplicity of interpreting the coefficient estimates, the shale boom begins in 2007 and continues until the end of the sample period (2013). $\lambda_c$ and $\alpha_t$ stand for county fixed effects and year fixed effects, respectively. $\beta_1$ is the parameter of interest which shows the estimated average treatment effect.

When using the DD approach, it is important to find an appropriate counterfactual county for each boom county. We utilize propensity score matching to identify a control group of counties from across the U.S. that are not in proximity to shale counties whose demographic characteristics are similar to the boom counties in the pre-boom time. We select the counties that are similar to the treated counties in employment counts, aggregate earnings, the ratio of workers with a college degree, and rate of changes in employment and earnings differentials between different groups of workers. In this way, the control group contains counties which are similar to the shale counties in size, average educational level of workers, average earnings, and have similar trends employment and earnings differentials. As a robustness check, we show that results are robust to alternative choices of control groups.

## 3.2 Instrumental Variables

While the DD approach is convenient to estimate and coefficient estimates have simple interpretation, it is also is subject to some inherent limitations. First, it requires the establishment of a treatment date. In this context, we choose 2007 as the treatment date consistent with when EIA begins its Drilling Productivity Reports that track shale production. But in reality, the timing and intensity of activity varied significantly across areas. To illustrate the importance in this context, consider that natural gas wellhead prices fell from a peak of more than $10 per thousand cubic feet in July of 2008 to less than $3 in September of 2009.[10] This price drop occurred about mid-way through our sample and impacted plays in different ways. A "dry gas" play like the Haynesville experienced a quick subsequent drop in drilling activity, while producers substituted towards oil plays such as the Bakken and Permian. The evolution of the value of production is illustrated in Appendix Table A.2. A subsequent price drop for oil did occur starting in mid-2014, after our sample concludes.

In order to capture variation both across time and across plays, we implement the instrumental variable specification utilized in Feyrer et al. (2017). We provide a brief overview of the instrument. The goal of the instrument is to model the intensity of activity by examining the value of new production per capita in a county such that intensity of activity varies across time and across the seven shale plays. In order to accomplish this, data on oil and natural gas extracted from *new* wells are aggregated to the county level. New wells are defined as having been producing for a year or less. Production is then multiplied by prices of oil and natural gas, respectively.[11] The key independent variable is the total estimated value of oil and natural gas extracted from wells that started producing in the current year measured in millions of dollars per worker, $NewValue_{it}$.

Because new production depends on a firms decision to drill in a county, which is endogenous, a two-step process is utilized. First:

$$\ln(NewValue_{i,t} + 1) = \alpha_i + \lambda_{jt} + \varepsilon_{it} \tag{2}$$

where $\alpha_i$ is a dummy for each county and $\lambda_{jt}$ represents a set of dummy variables for each play-year combination. 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. Next, these predictions are transformed into new production per employee for each county-year pair using (3).

$$New\hat{V}alue_{it} = (e^{\hat{\alpha_i} + \hat{\lambda_{jt}}} - 1)/emp_{i,2004} \tag{3}$$

This approach uses the time series of the value of production within a shale play to predict

[10] Based on U.S. Monthly Natural Gas Wellhead Price from EIA.

[11] Agerton and Upton (2019) show that prices vary significantly across shale plays. We utilize WTI prices for oil and Henry Hub natural gas prices for gas for consistency.

county-level value of production such that the predicted values of new production per worker are based on the timing of new production for all counties within the play. The data used to construct the instrument is borrowed directly from Feyrer et al. (2017), but is re-constructed using the shale play geographic definitions from EIA's Drilling Productivity Reports shown in Figure A.1.[12]

We utilize this instrument in two sets of results. First, we estimate the effect of a million dollar change in the value of new production per worker on earnings differentials. This provides a robustness check for the baseline DD approach that better captures variation in intensity across both plays and time, but also creates a parameter estimate that can be applied in other resource booms.

Second, we utilize the IV strategy to estimate the impact of a direct labor demand shock to both males and workers with a high school education on aggregate labor market earnings differentials. In other words, we ask the following questions. *If a labor demand shock increases high school (male) employment by 10 percentage points, what is the economy-wide impact on high school / college (male/female) earnings differentials?* This provides an elasticity estimate of the sensitivity of economy-wide earnings differentials based on a percentage point increase in employment induced by the productivity shock.

Specifically, Equations (4) and (5) describe the elasticity of earnings differentials with respect to labor demand shocks.

$$\frac{\%\Delta \frac{e_c}{e_{hs.}}}{\%\Delta L_{hs.}} = \epsilon_{c,hs} \tag{4}$$

$$\frac{\%\Delta \frac{e_m}{e_f}}{\%\Delta L_m} = \epsilon_{m,f} \tag{5}$$

We hypothesize that $\epsilon_{c,hs} < 0$ and $\epsilon_{m,f} > 0$. Equation (6) and (7) show the first and second stage equations that is estimated. H.S. $Emp_{c,t}$ is employment of high school workers in year $t$ and county $c$ and $\overline{PreShaleEmp_c}$ is the average total employment in the pre-shale time period.

$$\textbf{First Stage:} \frac{H.S.Emp_{c,t}}{\overline{PreShaleEmp_c}} = \alpha + \beta_3 New\widehat{Value}_{it} + \lambda_c + \alpha_t + \theta_{c,t} \tag{6}$$

$$\textbf{Second Stage:} \ln(\frac{e_c}{e_{h,s.}}) = \gamma\Big(\frac{H.S.\widehat{Emp}_{c,t}}{\overline{PreShaleEmp_c}}\Big) + \lambda_c + \alpha_t + \nu_{c,t} \tag{7}$$

[12]We differ slightly from Feyrer et al. (2017) who use a one-year lag of employment. Based on feedback, we instead use the employment in the earliest sample year to avoid the problem of endogenous employment growth after the shale boom. This choice does not materially impact results.

Corollary estimates, are presented for the impact of male employment shocks on male/female earnings differentials, where $\gamma$ is the coefficient of interest. For all IV estimates, we utilize the same propensity score matched control group. All standard errors are clustered at the shale-by-year level.[13]

# 4 Results

## 4.1 Difference-in-Differences

Before discussing the estimated effects on all the outcomes of interests, we implement a standard event study to test the parallel trend assumption between the treated and control group. The results are depicted in Figure 1. It is clear that, in all four event studies (college / high school — male / female — earnings differentials / employment differentials), there is no statistically significant difference in the shale counties and non-shale control counties in the pretreatment years. The results indicate that the parallel trend assumption is fulfilled.[14]

### 4.1.1 Employment and Earnings

Table 2 shows estimated increases in employment and earnings by education level and gender across sectors; all sectors, mining, and non-mining. For total earnings and employment, we consider two samples. First, we utilize the largest balanced panel of counties available in QWI. These results presented in columns (1) and (5) are labeled "Full Sample." Next we restrict the sample to counties where information regarding employment and earnings in the mining sector is available. Due to data censoring, if there are not sufficient number of employees in a given sector, the data is not available for that county/year combination. We note that the difference in estimated treatment effects between the larger and smaller sample can be meaningful, especially for employment results.

For the large sample, we estimate total employment increased by an average of 1,381 workers in shale counties relative to non-shale counties. Comparing this to the pre-shale average employment in Table 1 (29,230), this yields an estimated treatment effect of about

[13]Statistical significance of results do not seem to be sensitive to standard errors chosen.

[14]When the outcome is the employment differential between high- and low-skilled workers, there is a seemlingly slightly and insignificant unparallel trend before the shale boom. This could be caused by some of the shale counties which started gas production before 2007. When we drop such shale plays, such as Niobrara and Permian, the insignificant unparallel trend is elimited. And dropping these shaleplays out of the sample does not affect our main findings.

4.7% percent. The estimated treatment effect for earnings is around $240 in the full sample, or about a 9.3% increase relative to pre-shale levels.

Mining employment increased by 356 workers relative to controls. While this is only about 1% of total employment, this is about a 41% increase above pre-shale mining employment. We estimate a non-mining employment treatment effect that is larger in absolute value, 5,617 workers, but about 14% increase above pre-shale non-mining employment.

For workers with a college degree or higher, employment increased by 866, 48, and 822 workers in the all-sector sample, the mining sector, and the non-mining sector, respectively relative to controls. All the coefficients are statistically significant at 1 percent level. Similar patterns are found for workers with a high school diploma or less, male and female workers.

The impact of the shale-boom on earnings is reported in column (5)-(8). The effect is the largest for workers in the mining sector, and this is true for all subsets of workers examined. For example, and shown in Panel A, the shale boom is associated with a $673 increase in monthly average earnings for workers in the mining sector, while the increase in earnings for all workers on average and workers in the non-mining sectors are $241 and $174, respectively.

Effects on earnings are largest for workers with a high school diploma or less ($279) and male workers ($323), and are even larger for high school educated and male workers in the mining sector ($720 and $713, respectively). Overall, results in Table 2 corroborate the summary statistics in Table 1 and concurrent research finding that the shale boom has significantly increased labor demand for less-educated male workers (Cascio and Narayan, 2017; Bartik, 2017; Kearney and Wilson, 2018).

### 4.1.2 Employment and Earnings Differentials

Table 3 presents our main result, namely the estimated effect on employment and earnings differentials. All dependent variables are presented in log differences of employment and earnings for each respective group and therefore can be interpreted as percent change in the differential of employment and earnings associated with the shale boom.

We find that college / high school employment differentials decreased by about 4.4% in the full sample and 3.2% in the small sample. Employment differential decreases are also observed in the mining (3.8%) and non-mining (2.9%) industries. For earnings differentials, we estimate a 4.4% decrease between workers with college/high school educations, with a larger effect in the non-mining sector of 4.1% than that in the mining sector, 2.4%.

For male / female workers, we find that employment differentials increased by around 3.7% in the full sample and about 5.8% in the small sample. We find no change in employment differentials within the mining sector, but do find a 4.3% increase in the non-mining sectors. Earnings differentials increased by around 2.7% in the full sample and 3.5% in the small sample. Similar to employment differentials, point estimates for the mining sector are not statistically different from zero, but earnings differentials increased by an estimated 2.9% in the non-mining sector.

We note a few broad observations to put these results into context.

First, we observe a *decrease* in earnings differentials between workers with college and high school educations, while we observe an *increase* in earnings differentials between male and female workers. Second, we observe a change in earnings differentials even within the non-mining sectors. In particular, college/high school earnings differentials decreased by 4.1% in the non-mining sectors. Similarly, the male/female earnings differentials increased by 2.9% in the non-mining sectors. Thus, not only did earnings differentials overall change because of labor migration and/or substitution into the higher paying mining sector, but also because of changes within sectors not directly related to mining. Third, we observe a *decrease* in employment differentials between workers with college and high school educations and an *increase* in employment differentials between male and female workers within non-mining sectors.

One potential reason for the observed change in employment differentials within non-mining sectors might be due to the composition of the industries indirectly effected by the shale boom. As shown in Table 1, the construction and transportation sectors experienced relatively large employment growth relative to controls. This is consistent with Feyrer et al. (2017) that tests for effects across sectors and finds that the construction and transportation sectors were the most impacted, less of course the oil and gas industry. And, like the oil and gas sector, the construction and transportation sectors are heavily employed by male workers with high school educations.[15]

Thus, potentially, the estimated change in employment differentials within the non-oil and gas industries is due to the simple fact that the two most indirectly impacted industries also have a similar composition (predominantly male high school educated workers). To provide insight into this hypothesis, Table 4 presents results for the construction, transportation, and all other non-construction, non-transportation, and non-oil and gas sectors separately.

[15]More specifically, 83% and 71% of employment in the construction and transportation sectors nationally are male (as compared to 52% of labor force is male). Similarly, 47% and 45% of these workers have a high school degree or less, compared to 37% of the U.S. labor force. Source: Quarterly Workforce Indicators, U.S. Census Bureau. Beginning of Quarter Employment counts. Q1 2017 to Q4 2017. Transportation sector includes transportation and warehousing.

We observe a decrease in the college/high school employment differentials in all three categories. Specifically, we estimate a 2.9% and 4.8% decrease in employment differentials in the construction and transportation sectors, respectively. Within the other non-mining sectors, we still observe a 1.9% decrease in the college/high school employment differential. For male/female employment differentials, we find no statistically significant effect within the construction or transportation sector, but we find a statistically significant 1.4% increase in employment differentials in other non-mining sectors. While point estimates are smaller in magnitude than the overall estimated treatment effects, we still find evidence that employment differentials changed in seemingly unrelated sectors.

Table 4 also shows the effects on earnings differentials. For college/high school earnings differentials, we find no effect in construction, a 4.6% decrease in transportation, and a 2.6% decrease in other non-mining sectors. For male/female earnings differentials, we find a 2.3%, 2.0%, and 2.5% increase across these sectors respectively. Thus, while the oil and gas industry only employed about 2% of employment in the pre-boom period in treated areas, a productivity shock to this one (relatively small) sector had implications for earnings differentials in sectors not only outside of the oil and gas sector itself, but also outside of the two other most effected sectors, namely construction and transportation. This speaks to the importance of labor demand shocks to a small subset of workers on seemingly unrelated sectors of the economy.

#### 4.1.3 Employment and Earnings Differentials by Region

In Table 5 we disaggregate our main result (Table 3) by shale play. We do this for two reasons. First, we want to ensure that results are robust across different plays, to mitigate the concern that one area is driving all results. Second, this provides point estimates that might be useful for policy makers interested in geographic specific regions. We conduct analysis on the *Anadarko, Appalachia, Bakken, Eagle Ford, Haynesville, Niobrara,* and *Permian* regions per the geographic definitions of EIA's Drilling Productivity Reports.

For college/high school earnings differentials, we estimate a significant and negative treatment effect in six out of the seven regions. Of these seven regions, the magnitude of the effect ranges from around 1.5% (Permian) to 9% (Bakken). We estimate a positive and significant treatment effect for male/female earnings differentials in five out of the seven regions. Point estimates range from 10.8% (Bakken) to less than around 2% (Appalachia and Permian).

### 4.1.4 Alternative Control Groups

We next test the sensitivity to alternative control groups. In all prior analysis, we utilize propensity score matching to identify a control group of counties whose demographic characteristics are similar to the boom counties in the pre-boom time. In order to test the sensitivity of results to a different choice of control groups, Table 6 shows the results for employment and earnings differentials using 20 random control groups taken from all counties in the United States not in proximity to shale counties.[16] We simply select a random control county for each treatment county in lieu of the propensity score match. This process of generating a random control group is then performed 20 times, and for each of these iterations we estimate a treatment effect. In total, the 80 treatment effects estimated are presented.

We highlight two observations. First, all of the estimated treatment effects for college/high school employment and earnings differentials are negative, and all estimated treatment effects for male/female employment and earnings differentials are positive, consistent with estimates using the propensity score match control group. Second, notice that for three of the four categories, the baseline point estimates (from Table 3) fall in the range of the random control groups. For earnings differentials, the baseline results are almost identical to the average treatment effect from the random control groups.

## 4.2 Instrumental Variable Results

We next utilize the instrumental variables approach used in Feyrer et al. (2017). This approach has two benefits. First, it takes into account the variation in intensity of the shale boom both across time and between shale play areas. Second, it allows parameter estimates to be scaled such that they can be used more generally in other contexts. For comparison, we show results using both OLS and IV.

IV results in Table 7 show that one million dollars of oil and gas production per person is associated with about a 8.1% decrease in the college/high school employment differential, and about a 6.1% decrease in the earnings differential. Similarly, one million dollars of oil and gas production per person is associated with about a 21.8% increase in the male/female employment differential and approximately a 6.5% increase in the male/female earnings differential. In comparison to the OLS estimates, all four IV estimates are lower in absolute value. Both OLS and IV estimates are in the expected direction. Unsurprisingly, OLS estimates have lower standard errors. IV estimates do not produce statistically significant

[16] Again, we pull from the following states with no shale activity: AL, AZ, CA, CT, DE, FL, GA, HI, ID, IL, IN, IA, ME, MI, MN, MS, MO, NV, NH, NJ, NC, OR, RI, SC, TN, VT, WA, and WI.

results. Standard errors are clustered at the county and year level, consistent with Feyrer et al. (2017).

Table 8 presents estimates of the elasticity of earnings differential shocks with respect to an initial labor demand shock illustrated in equations (4) and (5). In essence, we take advantage of a plausibly exogenous technology induced labor demand shock to trace the change in relative earnings associated with the demand shock. We then index the change in relative earnings to the equilibrium change in the composition of male and high school educated workers.

Results are again presented both with OLS and IV. IV estimates suggest that a 10 percentage point increase in the share of workers with high school education induced by the shale boom is associated with a about 2.4% decrease in college/high school earnings differentials. Similarly, we estimate that a 10 percentage point increase in the share of male employment induced by the shale boom is associated with an about 1.5% increase in male/female earnings differentials. For college/high school differential, IV estimates are very similar in magnitude, while male/female earnings differential estimates IV are about twice the magnitude higher compared to OLS estimates.

## 4.3 Decomposition

In this final results section, we address the plausible channel through which the productivity shock impacted earnings differentials. We consider earnings differentials within a region broken out into two representative firms as follows:

$$\frac{e_H}{e_L} = \frac{N_m}{N}\frac{e_{H,m}}{e_{L,m}} + \frac{N_o}{N}\frac{e_{H,o}}{e_{L,o}} \tag{8}$$

where $N_{m|o}$ is employment counts in the mining and non-mining (other) sectors, and $N$ is total employment. Equation 8 states that the earnings differential in a region is an employment weighted average of earnings differentials within the two representative sectors. Any change in $\frac{e_H}{e_L}$ can therefore be decomposed into three channels.

**Channel 1: Earnings Differentials within Mining Sector** The first channel is a change in $\frac{e_{H,m}}{e_{L,m}}$. In words, this is the change the earnings differential within the mining sector, holding constant the earnings differential in the non-mining sectors ($\frac{e_{H,o}}{e_{L,o}}$), and the relative employment shares in the respective sectors ($\frac{N_m}{N}$, $\frac{N_o}{N}$).

**Channel 2: Earnings Differentials within Non-Mining Sector** The second channel is a change in the earnings differential within the non-mining sector ($\frac{e_{H,o}}{e_{L,o}}$), again, holding constant the earnings differential in the mining sectors ($\frac{e_{H,m}}{e_{L,m}}$), and the relative employment shares in the respective sectors ($\frac{N_m}{N}, \frac{N_o}{N}$).

**Channel 3: Employment Migration and Substitution Between Sectors** The third channel through which a change in $\frac{e_H}{e_L}$ can occur is through the relative share of the mining and non-mining employment, respectively, to total employment. By construction, $\frac{N_m}{N} + \frac{N_o}{N} = 1$, i.e. 100% of the employment comes from these representative two sectors. Therefore, holding constant both $\frac{e_{H,m}}{e_{L,m}}$ and $\frac{e_{H,o}}{e_{L,o}}$, a change in $(1 - \frac{N_m}{N})$ (or similarly $(1 - \frac{N_o}{N})$) can create a change in the earnings differentials if earnings differentials are different between the sectors.

As already presented, we find evidence for all three of these effects. Thus, next we decompose the relative contribution of these three effects to understand their relative importance utilizing point estimates from Table 3 alongside summary statistics from Table 1. The specific algebra and calculations are shown in Online Appendix A.2 and results are summarized in Table 9.

The first rows in Panels A and B list point estimates from the difference-in-differences specification from Table 3. Results are broken out for college / high school and male / female in Panel A and B respectively.

In the second row of each panel, point estimates for the mining and non-mining sectors are scaled by the share of employment in the mining and non-mining sectors in the pre-boom time period. For college / high school earnings differentials, we estimate that of the 4.4% decrease, only 0.05% is associated with the change in earnings differential within the mining industry, while 4.01% is associated with a change in the non-mining industries. This is due to two factors. First, the estimated change in the earnings differential in the non-mining industry was almost twice the magnitude of the change in the mining industry. But second, only about 2% of the employment was in the mining sector before the boom.

The residual, 0.34% is associated with workers substituting and/or migrating into the mining sector. For male / female earnings differentials, of the 3.5% increase, less than 0.01% is associated with a change in the earnings differential within the mining-sector, while 2.9% is associated with a change in the non-mining sectors. The residual, 0.66%, is associated with labor substitution and migration.

The fourth row divides each of these three effects into the relative contribution such that the sum of all three effects is equal to 100%. Interestingly, the vast majority of the change in earnings differentials in these areas (relative to controls) for both college / high school and male /female come from changes within the non-mining industries. Specifically, 91% and 81% of the change in earnings differentials for college / high school and male / female respectively come from *within* the non-mining sectors. Very little of the change comes from within the mining sector itself. For college / high school and male / female, 7.7% and 18.8% respectively, of the change is associated with labor migration and substitution between sectors.

This decomposition highlights how a specific labor demand shock to a specific subset of workers in a specific sector (that is a relatively small share of total employment) can have a meaningful impact on earnings differentials in sectors that are not directly impacted by the shock. Further, while employment substitution and migration can impact earnings differentials overall in the labor market, the majority of the estimated change comes from *within* sectors that were not directly shocked. This highlights the importance of considering productivity shocks to subsets of workers in explaining broadly earnings differentials.

# 5 Discussion

It is important to consider the extent to which parameters estimated in this research can be generally applied to the U.S. labor market broadly. We point out three factors that should be considered when applying results of this research in other contexts.[17]

First, the areas impacted by the shale oil and gas booms are relatively rural, and the response of a rural labor market to a productivity shock might not be representative of the U.S. economy as a whole.

Second, the shale boom occurred around the time of the Great Recession; a time of historic slackness in the labor market. Had this labor demand shock occurred at a time with a tighter labor market, magnitudes might be different. For instance a demand shock during a tight labor market might experience more earnings gains relative to employment gains, while under slack conditions, the opposite is true.

Third, the oil and gas sector has relatively low barriers to entry. A male can plausibly get a job working as a "roustabout" on a rig out of high school, especially during a boom time. Similarly, the construction and transportation sectors have relatively low barriers to entry. A productivity shock in an industry with higher barriers to entry would be expected to have higher earnings effects in the short-run, and less employment response.

[17]While these factors have been pointed out by various colleagues, this is by no means intended to be an exhaustive list.

# 6 Conclusions

In this paper we exploit a plausibly exogenous labor demand shock associated to a specific subset of workers (male workers with high school education) in a specific industry (oil and gas) on earnings and employment differentials within sectors not directly impacted by the shock. We find that earnings differentials between men and women increase while earnings differentials between workers with college and high school education decrease. These effects are also observed within sectors not directly impacted by the shock.

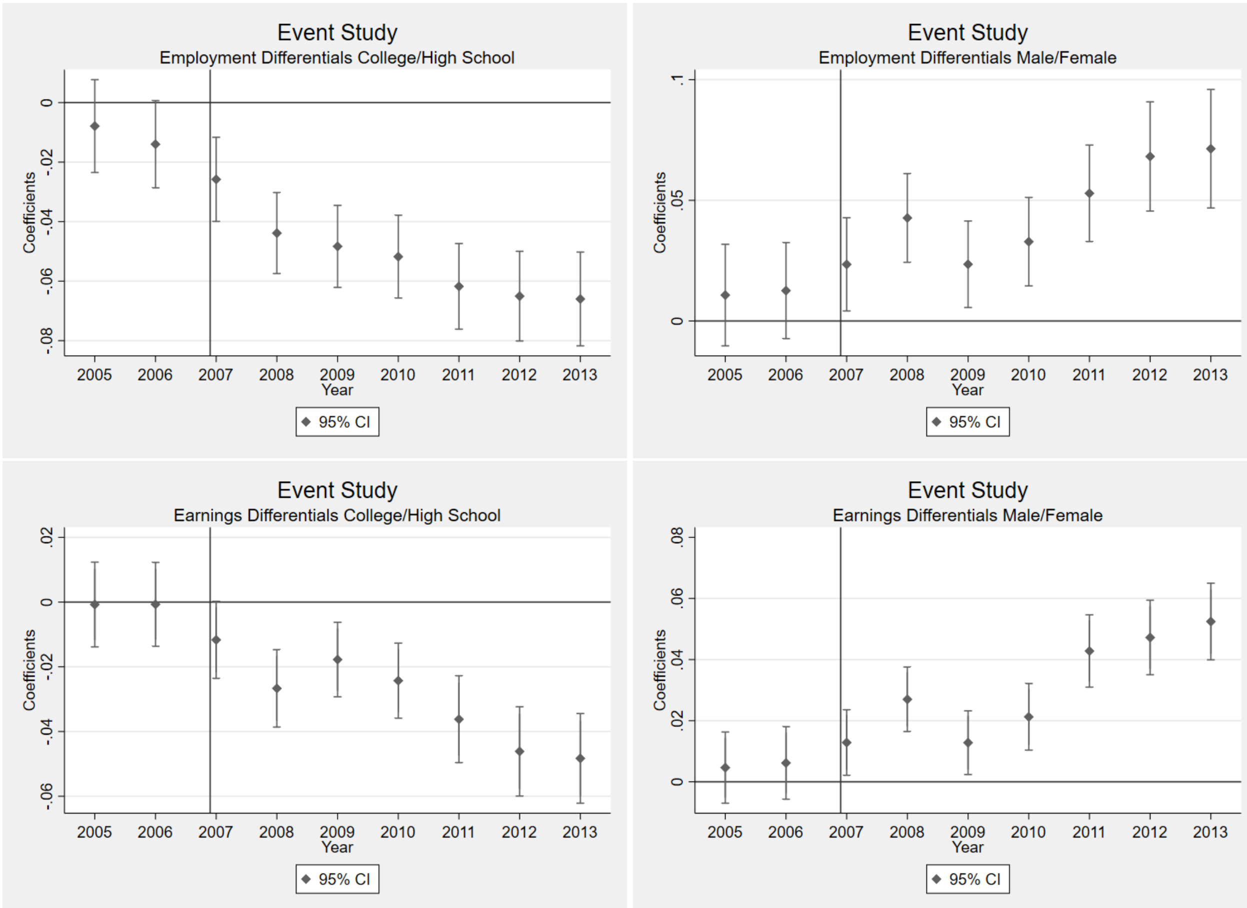
We also observe changes in employment differentials between men and women and workers with college and high school educations within non-oil and gas related sectors. This result is observed in related construction and transportation sectors, but also in plausibly not directly related non-construction and non-transportation sectors.

We show that results for both earnings and employment differentials are robust across seven geographic regions, alternative control groups, and an IV strategy that takes into account varying intensity and timing across time and geographic regions. We also estimate elasticities of the labor demand shocks on earnings differentials. Specifically, we find that a 10 percentage point increase in the high school employment rate is associated with a 2.6% decrease in the college/high school earnings differential, and that a 10 percentage point increase in the male employment rate is associated with a 1.5% increase in male/female earnings differentials.

Finally, we decompose these effects into three channels. The first channel is that the labor demand shock might impact the earnings differentials *within* the affected sector, namely the mining sector in this context. The second channel is that the labor demand shock might impact the earnings differentials in the *non-mining* sectors. The third plausible channel is that male workers with high school education substitute into the mining sector or migrate into the area due to employment opportunities in the mining sector.

For college / high school and male / female earnings differentials 91% and 81% respectively of the observed change in earnings differentials can be explained by changes within the non-mining sectors. Approximately 8% and 19% respectively can be explained through the employment substitution and migration channel. Very little of the change in earnings differentials can be explained by changes within the mining sector.

Results of this research have significant policy implications. We show that labor demand shocks to specific subsets of workers can have significant impacts on earnings differentials within seemingly unrelated sectors. Policies aimed at reducing income inequality across the income spectrum and/or at reducing income inequality between men and women should be aware of the sensitivity of labor demand shocks in seemingly unrelated sectors on earnings differentials within other sectors.




*Note:* The figure contains event study results for the effect of the shale boom on employment differentials and earnings differentials between high- and low-skilled workers, as well as between male and female workers. The X axis shows the years, where 2004 is omitted from the analyses. The Y axis is the scale of the treatment effect.

Figure 1: Event Study

## Table 1: Summary Statistics: Baseline Sample

| | Treatment Group | | | Control Group | | | |
|---|---|---|---|---|---|---|---|
| | Pre 2007 | Post 2007 | Percent Change | Pre 2007 | Post 2007 | Percent Change | %ΔTreatment - %ΔControl |
| **Employment (Thousands of Workers)** | | | | | | | |
| All | 29.23 | 29.21 | -0.07% | 31.71 | 30.31 | -4.42% | 4.35% |
| College + | 5.77 | 5.77 | 0.00% | 7.00 | 6.84 | -2.29% | 2.29% |
| Highschool | 10.35 | 10.71 | 3.48% | 10.69 | 10.55 | -1.31% | 4.79% |
| Male | 14.94 | 14.97 | 0.20% | 16.12 | 15.17 | -5.89% | 6.09% |
| Female | 14.29 | 14.24 | -0.35% | 15.59 | 15.14 | -2.89% | 2.54% |
| Male College(+) | 2.91 | 2.88 | -1.03% | 3.53 | 3.37 | -4.53% | 3.50% |
| Female College(+) | 2.86 | 2.89 | 1.05% | 3.48 | 3.47 | -0.29% | 1.34% |
| Male High School (-) | 5.65 | 5.87 | 3.89% | 5.79 | 5.62 | -2.94% | 6.83% |
| Female High School (-) | 4.70 | 4.83 | 2.77% | 4.90 | 4.93 | 0.61% | 2.15% |
| Mining Sector | 0.87 | 1.23 | 41.38% | 0.37 | 0.37 | 0.00% | 41.38% |
| Construction Sector | 1.97 | 1.91 | -3.05% | 2.03 | 1.54 | -24.14% | 21.09% |
| Transportation Sector | 1.27 | 1.34 | 5.51% | 1.41 | 1.36 | -3.55% | 9.06% |
| Non-Mining Sector (Mining Sample) | 39.82 | 39.53 | -0.73% | 136.58 | 130.67 | -4.33% | 3.60% |
| **Earnings ($)** | | | | | | | |
| All | $2,591 | $3,164 | 22.12% | $2,564 | $2,896 | 12.94% | 9.14% |
| College + | $3,892 | $4,524 | 16.24% | $3,863 | $4,242 | 9.80% | 6.43% |
| Highschool | $2,282 | $2,856 | 25.15% | $2,256 | $2,570 | 13.91% | 11.24% |
| Male | $3,227 | $3,928 | 21.72% | $3,136 | $3,518 | 12.16% | 9.56% |
| Female | $1,922 | $2,311 | 20.24% | $1,987 | $2,267 | 14.10% | 6.15% |
| Male College(+) | $4,886 | $5,706 | 16.78% | $4,807 | $5,292 | 10.08% | 6.69% |
| Female College(+) | $2,936 | $3,366 | 14.65% | $2,969 | $3,267 | 10.04% | 4.60% |
| Male High School (-) | $2,849 | $3,537 | 24.15% | $2,754 | $3,094 | 12.37% | 11.78% |
| Female High School (-) | $1,601 | $1,980 | 23.67% | $1,688 | $1,961 | 16.17% | 7.49% |
| Mining Sector | $4,339 | $5,522 | 27.26% | $4,330 | $4,840 | 11.77% | 15.49% |
| Construction Sector | $2,794 | $3,527 | 26.23% | $2,820 | $3,227 | 14.43% | 11.80% |
| Transportation Sector | $3,145 | $3,731 | 18.63% | $2,967 | $3,219 | 8.51% | 10.12% |
| Non-Mining Sector (Mining Sample) | $2,567 | $3,090 | 20.37% | $2,957 | $3,306 | 11.81% | 8.57% |
| **Other Variables** | | | | | | | |
| Value of Production $/Worker | $11,965 | $33,572 | 181% | $16 | $7 | -54.45% | 235.04% |
| IV: Value of Production $/Worker | $9,889 | $20,141 | 104% | $6 | $7 | 8.70% | 94.97% |
| Share White | 92% | 91% | -0.33% | 86% | 85% | -1.16% | 0.84% |
| College Share | 16% | 16.1% | 0.63% | 16% | 17% | 6.25% | -5.63% |

Averages of annual data for treatment and control groups. Pre-2007 period is 2004-2006. Post-2007 period is 2007-2013. Total employment in counts (a thousands people). Earnings are average monthly earnings of full time stable workers.

Table 2: Impact of Shale Boom on Employment and Earnings by Sector

| | Employment | | | | Earnings | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) Full Sample | (2) Small Sample | (3) Mining Sector | (4) Non-Mining Sector | (5) Full Sample | (6) Small Sample | (7) Mining Sector | (8) Non-Mining Sector |
| *Panel A: All Workers* | | | | | | | | |
| Treated | 1.381*** | 5.962*** | 0.356*** | 5.617*** | 240.4*** | 240.6*** | 673.4*** | 174.0*** |
| | (0.214) | (0.620) | (0.034) | (0.617) | (12.453) | (14.785) | (59.268) | (12.686) |
| *N* | 6,220 | 4,060 | 4,060 | 4,060 | 6,220 | 4,060 | 4,060 | 4,060 |
| *Panel B: Workers with at Least A College Degree* | | | | | | | | |
| Treated | 0.167*** | 0.866*** | 0.0475*** | 0.822*** | 253.1*** | 200.5*** | 828.4*** | 133.5*** |
| | (0.049) | (0.145) | (0.005) | (0.145) | (20.136) | (24.312) | (157.803) | (24.410) |
| *N* | 6,220 | 4,060 | 4,060 | 4,060 | 6,220 | 4,060 | 4,060 | 4,060 |
| *Panel C: Workers with A High School Diploma or Lower* | | | | | | | | |
| Treated | 0.496*** | 0.770*** | 0.173*** | 0.600*** | 260.0*** | 278.8*** | 719.7*** | 205.5*** |
| | (0.081) | (0.197) | (0.016) | (0.194) | (12.159) | (14.855) | (39.936) | (12.429) |
| *N* | 6,220 | 4,060 | 4,060 | 4,060 | 6,220 | 4,060 | 4,060 | 4,060 |
| *Panel D: Male Workers* | | | | | | | | |
| Treated | 0.984*** | 4.297*** | 0.304*** | 3.999*** | 319.5*** | 322.6*** | 712.8*** | 245.7*** |
| | (0.138) | (0.370) | (0.030) | (0.366) | (16.314) | (18.824) | (68.494) | (18.177) |
| *N* | 6,220 | 4,060 | 4,060 | 4,060 | 6,220 | 4,060 | 4,060 | 4,060 |
| *Panel C: Female Workers* | | | | | | | | |
| Treated | 0.397*** | 1.665*** | 0.0514*** | 1.618*** | 109.1*** | 79.41*** | 401.6*** | 65.62*** |
| | (0.087) | (0.267) | (0.006) | (0.266) | (6.187) | (7.023) | (38.472) | (5.777) |
| *N* | 6,220 | 4,060 | 4,060 | 4,060 | 6,220 | 4,060 | 4,060 | 4,060 |

Dependent variables are total employment (1,000 people) and monthly average earnings (USD) in column (1)-(4) and (5)-(8), respectively. Standard errors are clustered at county and year level and are reported in parentheses. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table 3: Impact of Shale Boom on Employment and Earnings Differentials by Sector

| | Employment Differentials | | | | Earnings Differentials | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) Full Sample | (2) Small Sample | (3) Mining Sector | (4) Non-Mining Sector | (5) Full Sample | (6) Small Sample | (7) Mining Sector | (8) Non-Mining Sector |
| *Panel A: College (+) / High School (-)* | | | | | | | | |
| Treated | -0.044*** | -0.032*** | -0.038*** | -0.029*** | -0.030*** | -0.044*** | -0.024*** | -0.041*** |
| | (0.003) | (0.003) | (0.010) | (0.003) | (0.003) | (0.003) | (0.009) | (0.004) |
| *N* | 6,220 | 4,060 | 4,060 | 4,060 | 6,220 | 4,060 | 4,060 | 4,060 |
| *Panel B: Male / Female* | | | | | | | | |
| Treated | 0.037*** | 0.058*** | -0.013 | 0.043*** | 0.027*** | 0.035*** | 0.002 | 0.029*** |
| | (0.005) | (0.005) | (0.015) | (0.005) | (0.003) | (0.003) | (0.010) | (0.003) |
| *N* | 6,220 | 4,060 | 4,060 | 4,060 | 6,220 | 4,060 | 4,060 | 4,060 |

Dependent variables are the natural log differentials in employment and monthly average earnings in columns (1)-(4) and (5)-(8), respectively. Standard errors are clustered at county and year level and are reported in parentheses.***p≤0.01, **p≤0.05, *p≤0.1.

Table 4: Impact of Shale Boom on Employment and Earnings Differentials by Sub-Sector

| | Employment Differentials | | | Earnings Differentials | | |
|---|---|---|---|---|---|---|
| | (1) Construction Sector | (2) Transportation Sector | (3) Other Non-Mining Sectors | (4) Construction Sector | (5) Transportation Sector | (6) Other Non-Mining Sectors |
| *Panel A: College (+) / High School (-)* | | | | | | |
| Treated | -0.029*** | -0.048*** | -0.019*** | 0.004 | -0.046*** | -0.026*** |
| | (0.006) | (0.008) | (0.003) | (0.006) | (0.007) | (0.003) |
| *Panel B: Male / Female* | | | | | | |
| Treated | 0.004 | 0.022 | 0.014*** | 0.023*** | 0.020** | 0.025*** |
| | (0.009) | (0.014) | (0.003) | (0.006) | (0.008) | (0.003) |
| $N$ | 5340 | 4660 | 3600 | 5340 | 4660 | 3600 |

The dependent variables are the natural log differentials in employment and monthly average earnings in columns (1)-(3) and (4)-(6), respectively. Other non-mining sectors are all sectors excluding the mining, construction, and transportation. Standard errors are clustered at county and year level and are reported in parentheses. ***p≤0.01, **p≤0.05, *p≤0.1.

Table 5: Impact of Shale Booms on Earnings Differentials by Region

| | Shale Play | | | | | | |
|---|---|---|---|---|---|---|---|
| | Anadarko | Appalachia | Bakken | Eagle Ford | Haynesville | Niobrara | Permian |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| *Panel A: College (+) / High School (-) Earnings Differential* | | | | | | | |
| Treated | -0.065*** | -0.016*** | -0.089*** | -0.022* | -0.034*** | -0.000 | -0.015* |
| | (0.011) | (0.004) | (0.015) | (0.013) | (0.007) | (0.009) | (0.009) |
| *N* | 580 | 2,520 | 400 | 460 | 480 | 700 | 1,080 |
| *Panel B: Male / Female Earnings Differential* | | | | | | | |
| Treated | 0.038*** | 0.025*** | 0.108*** | 0.062*** | -0.000 | -0.007 | 0.017* |
| | (0.010) | (0.004) | (0.015) | (0.012) | (0.008) | (0.008) | (0.009) |
| *N* | 580 | 2,520 | 400 | 460 | 480 | 700 | 1,080 |

Dependent variable is logged differentials in earnings (USD). Earnings are average monthly earnings of full time stable workers. County and year fixed effects are controlled in all regressions. Standard errors are clustered at county and year level and are reported in parentheses. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table 6: Impact of Shale Boom on Employment and Earnings Differentials - Random Control Groups

| | Employment Differentials | | Earnings Differentials | |
|---|---|---|---|---|
| | College/ High School | Male/ Female | College/ High School | Male/ Female |
| Iteration | (1) | (2) | (3) | (4) |
| 1 | -0.0308*** | 0.0380*** | -0.0254*** | 0.0263*** |
| 2 | -0.0358*** | 0.0420*** | -0.0351*** | 0.0271*** |
| 3 | -0.0342*** | 0.0383*** | -0.0282*** | 0.0314*** |
| 4 | -0.0351*** | 0.0438*** | -0.0357*** | 0.0274*** |
| 5 | -0.0333*** | 0.0442*** | -0.0353*** | 0.0300*** |
| 6 | -0.0291*** | 0.0306** | -0.0249*** | 0.0241*** |
| 7 | -0.0334*** | 0.0448*** | -0.0336*** | 0.0297*** |
| 8 | -0.0330*** | 0.0443*** | -0.0279*** | 0.0278*** |
| 9 | -0.0308*** | 0.0405*** | -0.0317*** | 0.0228*** |
| 10 | -0.0377*** | 0.0409*** | -0.0310*** | 0.0254*** |
| 11 | -0.0362*** | 0.0415*** | -0.0293*** | 0.0238*** |
| 12 | -0.0342*** | 0.0419*** | -0.0308*** | 0.0277*** |
| 13 | -0.0345*** | 0.0446*** | -0.0312*** | 0.0256*** |
| 14 | -0.0353*** | 0.0423*** | -0.0289*** | 0.0277*** |
| 15 | -0.0313*** | 0.0427*** | -0.0298*** | 0.0285*** |
| 16 | -0.0361*** | 0.0386*** | -0.0303*** | 0.0303*** |
| 17 | -0.0314*** | 0.0468*** | -0.0341*** | 0.0268*** |
| 18 | -0.0389*** | 0.0453*** | -0.0360*** | 0.0278*** |
| 19 | -0.0342*** | 0.0363*** | -0.0328*** | 0.0244*** |
| 20 | -0.0346*** | 0.0417*** | -0.0278*** | 0.0296*** |
| High | -0.0389 | 0.0468 | -0.0360 | 0.0303 |
| Average | -0.0340 | 0.0415 | -0.0310 | 0.0272 |
| Low | -0.0291 | 0.0306 | -0.0249 | 0.0228 |
| Baseline Results (Table 3) | -0.044*** | 0.037*** | -0.030*** | 0.027*** |

Dependent variables are natural log of employment and earnings differentials, respectively. Standard errors are clustered at county and year level and are omitted for brevity. ***p≤0.01, **p≤0.05, *p≤0.1.

Table 7: Value of Production from New Wells and Employment and Earnings Differentials

| | (1) Employment Differentials | (2) Earnings Differentials |
|---|---|---|
| **Panel A: OLS** | | |
| *College (+) / High School (-)* | | |
| County Value of Production/Capita | -0.105*** | -0.077** |
| | (0.040) | (0.030) |
| County Value of Production/Capita | 0.293** | 0.071** |
| | (0.115) | (0.029) |
| **Panel B: IV** | | |
| *College (+) / High School (-)* | | |
| County Value of Production/Capita | -0.081 | -0.061 |
| | (0.050) | (0.047) |
| *Male / Female* | | |
| Value of Production/Capita by Shale Play-Year | 0.218 | 0.065 |
| | (0.158) | (0.040) |
| $N$ | 6,220 | 6,220 |

Dependent variables are natural log differentials in employment and monthly average earnings in columns (1) and (2), respectively. First stage F-value is 144 for IV regressions. County fixed effects and year fixed effects are controlled in all regressions. Standard errors are clustered at county and year level and are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Labor Demand Shocks and Economy Wide Earnings Differentials

| | College/H.S. Earnings Differentials (1) | Male/Female Earnings Differentials (2) |
|---|---|---|
| **Panel A: OLS Estimates** | | |
| H.S. Employment Share | -0.237*** | |
| | (0.027) | |
| Male Employment Share | | 0.083*** |
| | | (0.018) |
| **Panel B: IV Estimates** | | |
| H.S. Employment Share | -0.244*** | |
| | (0.076) | |
| Male Employment Share | | 0.150*** |
| | | (0.044) |
| First Stage F-Stat | 591 | 565 |
| N | 6,220 | 6,220 |

Employment shares defined as the total H.S. and Male employment respectively in a given year as a share of average total employment pre-2007. County fixed effects and year fixed effects are controlled for in all regressions. Standard errors are clustered at county and year level and are reported in parentheses. ***p≤0.01, **p≤0.05, *p≤0.1.

Table 9: Decomposing Changes in Earnings Differentials

| | (1) Mining Sector | (2) Non-Mining Sectors | (3) Employment Migration & Substitution | (4) Total |
|---|---|---|---|---|
| *Panel A: College (+) / High School (-)* | | | | |
| Point Estimates (Table 3) | -2.4% | -4.1% | - | -4.4% |
| Share of Pre-Boom Employment (Table 1) | 2.14% | 97.86% | | 100% |
| Percent Change in Earnings Differential | -0.05% | -4.01% | -0.34% | -4.4% |
| Relative Contribution | 1.17% | 91.19% | 7.64% | 100% |
| *Panel B: Male / Female* | | | | |
| Point Estimates (Table 3) | 0.2% | 2.9% | - | 3.5% |
| Share of Pre-Boom Employment (Table 1) | 2.14% | 97.86% | | 100% |
| Percent Change in Earnings Differential | 0.004% | 2.838% | 0.66% | 3.5% |
| Relative Contribution | 0.12% | 81.09% | 18.79% | 100% |

Detailed calculations can be found in Appendix A.2.

# References


**Agerton, Mark and Gregory B. Upton**, "Decomposing Crude Price Differentials: Domestic Shipping Constraint or the Crude Oil Export Ban.," *The Energy Journal*, 2019, *40* (3).

**_ , Peter R. Hartley, Kenneth B. Medlock III, and Ted Temzelides**, "Employment impacts of upstream oil and gas investment in the United States," *Energy Economics*, 2017, *62*, 171–180.

**Aragon, Fernando M., Juan Pablo Rud, and Gerhard Toews**, "Resource shocks, employment, and gender: Evidence from the collapse of the UK coal industry," *Labour Economics*, 2018, *52*, 54–67.

**Attanasio, Orazio, Erik Hurst, and Luigi Pistaferri**, "The Evolution of Income, Consumption, and Leisure Inequality in the US, 1980-2010," *NBER Working Paper No. 17982*, 2012.

**Bartik, Alexander W.**, "Worker Adjustment to Changes in Labor Demand: Evidence from Longitudial Census Data.," *Job Market Paper*, 2017.

**_ , Janet Currie, Michael Greenstone, and Christopher R. Knittel**, "The Local Economic and Welfare Consequences of Hydraulic Fracturing," *American Economic Journal: Applied Economics*, 2019, *Forthcoming.*

**Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz**, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sector," *American Economic Journal: Applied Economics*, 2010, pp. 228–255.

**Black, Dan, Terra McKinnish, and Seth Sanders**, "The economic impact of the coal boom and bust," *The Economic Journal*, 2005, *115* (503), 449–476.

**Blau, Francine D. and Lawrence Kahn**, "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 2017, *55*, 789–865.

**Bono, Emilia Del and Daniela Vuri**, "Job Mobility and the Gender Wage Gap in Italy," *Labour Economics*, 2011, *18*, 130–142.

**Cascio, Elizabeth U. and Ayushi Narayan**, "Who Needs a Fracking Education? The Educational Response to Biased Technological Change," *NBER Working Paper*, 2017, (21359).

**Cosgrove, Brendan M., Daniel R. LaFave, Sahan T. M. Dissanayake, and Michael R. Donihue**, "The Economic Impact of Shale Gas Development: A Natural Experiment along the New York / Pennsylvania Border," *Agricultural and Resource Economics Review*, 2015, *44* (2), 20–39.

**Decker, Ryan A., Meagan McCollum, and Gregory B. Upton**, "Firm Dynamics and Local Economic Shocks: Evidence from the Shale Oil and Gas Boom," *Working Paper*, 2018.

**Feyrer, James, Erin T Mansur, and Bruce Sacerdote**, "Geographic dispersion of economic shocks: Evidence from the fracking revolution," *The American Economic Review*, 2017, *107* (4), 1313–1334.

**Goldin, Claudia**, "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, 2014, *104* (4), 1091–1119.

**Gunderson, Morley**, "Male-Female Wage Differentials and Policy Responses," *Journal of Economic Literature*, 1989, *27*, 46–72.

**James, Alexander and Brock Smith**, "Geographic Dispersion of Economic Shocks: Evidence from teh Fracking Revolution: Comment," *American Economic Review*, 2019.

**Katz, Lawrence F.**, "Changes in the wage structure and earnings inequality," *Handbook of Labor Economics*, 1999, *3*, 1463–1555.

**Kearney, Melissa S. and Riley Wilson**, "Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom," *The Review of Economics and Statistics*, 2018, *100* (4), 678–690.

**Komarek, Timothy M**, "Labor market dynamics and the unconventional natural gas boom: Evidence from the Marcellus region," *Resource and Energy Economics*, 2016, *45*, 1–17.

**Kotsadam, Andreas and Anja Tolonen**, "African Mining, Gender, and Local Employment," *World Development*, 2016, *83*, 325–339.

**Levy, Frank and Richard J. Murnane**, "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," *Journal of Economic Literature*, 1992, *30*, 1333–1381.

**Maddison, Angus**, "Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment," *Journal of Economic Literature*, 1987, *25*, 649–698.

**Manning, Alan and Joanna Swaffield**, "The Gender Gap in Early-Career Wage Growth," *The Economic Journal*, 2008, *118*, 983–1024.

**Marchand, Joseph**, "Local labor market impacts of energy boom-bust-boom in Western Canada," *Journal of Urban Economics*, 2012, *71* (1), 165–174.

_ **and Jeremy Weber**, "Local Labor Markets and Natural Resources: A Synthesis of the Literature," *Journal of Economic Surveys*, 2018, *32* (2), 469–490.

**McCollum, Meagan and Gregory B. Upton**, "Local Labor Market Shocks and Residential Mortgage Payments: Evidence from Shale Oil and Gas Booms," *Resource and Energy Economics*, 2018, *53*, 162–197.

**Mincer, Jacob**, "The Distribution of Labor Incomes: A Survey," *Journal of Economic Literature*, 1970, *8* (1), 1–26.

**O'Neill, June**, "The Gender Gap in Wages, circa 2000," *The American Economic Review Papers and Proceedings*, 2003, *93* (2).

**Weber, Jeremy G**, "The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming," *Energy Economics*, 2012, *34* (5), 1580–1588.

**Wood, Robert G., Mary E. Corcoran, and Paul N. Courant**, "Pay Differences among the Highly Paid: The Male-Female Earnings Gap in Lawyers' Salaries," *Journal of Labor Economics*, 1993, *11*, 417–441.

# A For Online Publication

## A.1 Appendix Figures

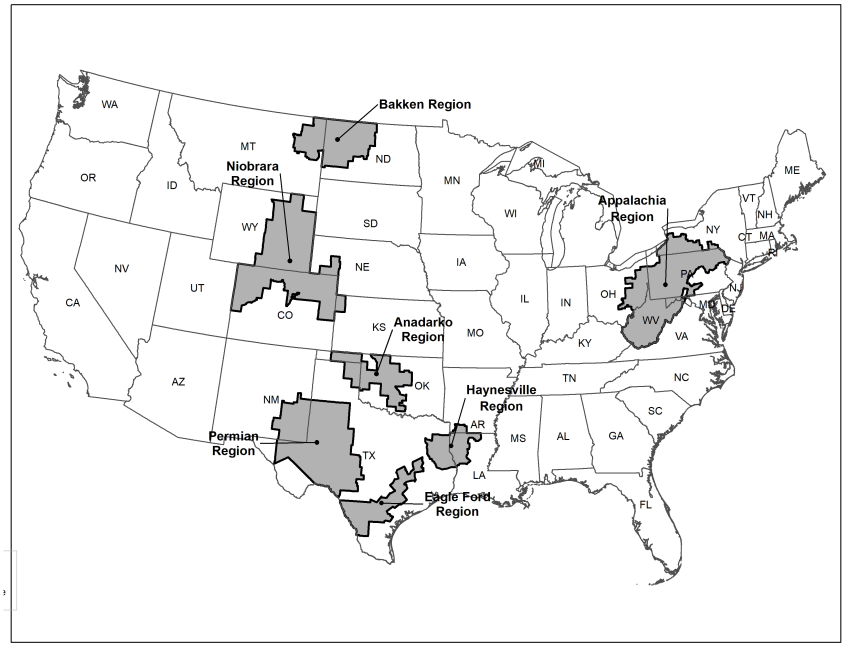



Figure A.1: U.S. Shale Plays
Source: EIA Drilling Productivity Report

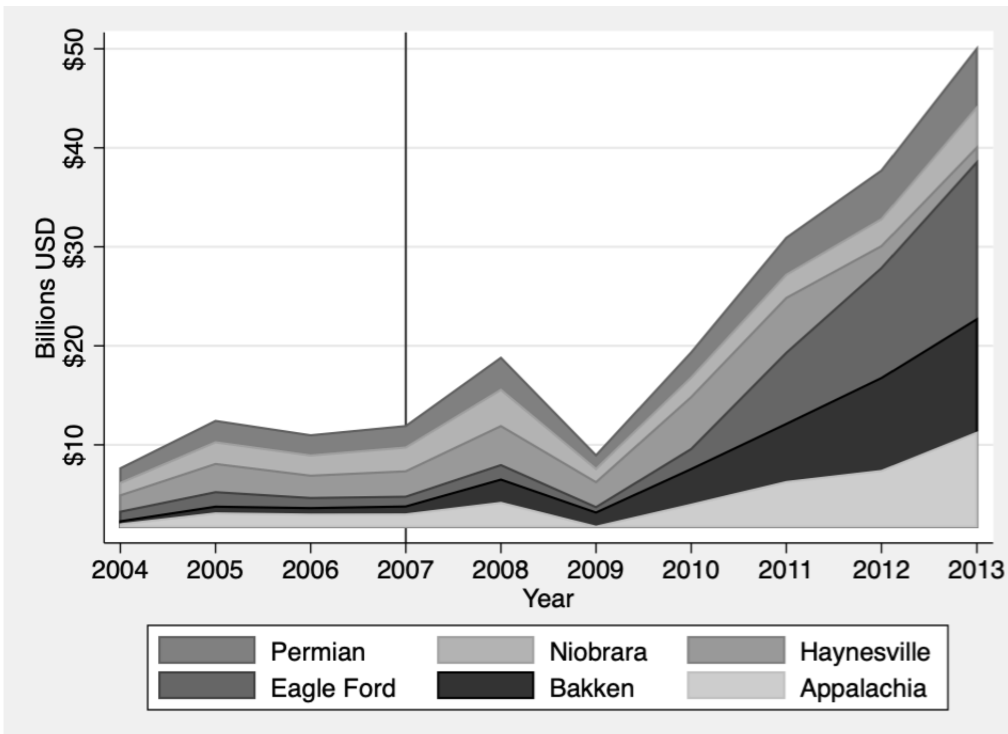



Figure A.2: Value of New Production by Shale Play
Source: Feyrer et al. (2017) and EIA.

## A.2 Decomposition

**Channel 1: Earnings Differentials within Mining Sector** We define $\Gamma_m$ as the percent change in the earnings differential in the local labor market associated with only the change in the earnings differential within the mining industry. In words, it is simply the share of the economy in the mining industry in the pre-boom time period multiplied by the estimated percent change in the earnings differential *within* the mining industry, where $\left(\frac{\overline{N_{m,t0}}}{\overline{N_{t0}}}\right)$ is simply the average share of employment ($N$) in the mining sector in the pre-boom time period. We sum mining employment and non-mining employment within the mining sample (from Table 1) for consistent comparison. $\left(\%\Delta\widehat{\frac{e_{H,m}}{e_{L,m}}}\right)$ is the estimated percent change in the earnings differential between high and low skilled workers in the mining industry associated with the shale boom.

$$\textit{Skilled/Unskilled: } \Gamma_m = \left(\frac{\overline{N_{m,t0}}}{\overline{N_{t0}}}\right) \times \left(\%\Delta\widehat{\frac{e_{H,m}}{e_{L,m}}}\right) = \frac{870}{40,690} * -2.4\% = -0.051\%$$

In words, of the 4.4% decrease in earnings differentials between skilled and unskilled workers, ≈0.05% is associated with a change in the earnings differential within the mining sector.

$$\textit{Male/Female: } \Gamma_m = \left(\frac{\overline{N_{m,t0}}}{\overline{N_{t0}}}\right) \times \left(\%\Delta\widehat{\frac{e_{M,m}}{e_{F,m}}}\right) = \frac{870}{40,690} * 0.2\% \approx 0.004\%$$

In words, of the 3.5% increase in earnings differentials between male and female workers, 0.004% is associated with a change in the earnings differential within the mining sector.

**Channel 2: Earnings Differentials within Non-Mining Sector** We define $\Gamma_o$ as the percent change in the earnings differential associated with only the change in the earnings differential within the non-mining (i.e. other) industries. In words, it is simply the share of the economy in the non-mining industries in the pre-boom time period multiplied by the estimated percent change in the earnings differential *within* the non-mining industries.

$$\textit{Skilled/Unskilled: } \Gamma_o = \left(\frac{\overline{H_o + L_o}}{\overline{H + L_{t0}}} \times \%\Delta\widehat{\frac{e_{H,m}}{e_{L,m}}}\right) = \left(1 - \frac{870}{40,690}\right) \times -4.1\% \approx -4.01\%$$

In words, of the 4.4% decrease in earnings differentials between skilled and unskilled workers, 4.01% is associated with a change in the earnings differential within the non-mining sectors.

$$\textit{Male/Female: } \Gamma_o = \left( \frac{M_o + F_o}{\overline{M + F}_{t0}} \times \%\Delta \widehat{\frac{e_{M,m}}{e_{F,m}}} \right) = \left(1 - \frac{870}{40,690}\right) \times 2.9\% \approx 2.84\%$$

In words, of the 3.5% increase in earnings differentials between male and female workers, 2.84% is associated with a change in the earnings differential within the non-mining sectors.

**Channel 3: Employment Migration and Substitution Between Sectors** We define $\Gamma_s$ as the percent change in the earnings differential associated with workers substituting into the high paying mining industry. It is simply the total change in earnings differentials ($\Gamma_T$) less that share associated with the changes within the mining ($\Gamma_m$) and non-mining ($\Gamma_o$) industries respectively.

$$\textit{Skilled/Unskilled: } \Gamma_s = \Gamma_T - (\Gamma_m + \Gamma_o) = -4.4\% - (-0.05\% + -4.01\%) = -0.34\%$$

In words, of the 4.4% decrease in earnings differentials between skilled and unskilled workers, 0.34% is associated with a disproportionate growth in unskilled employment in the higher paying mining sector.

$$\textit{Male/Female: } \Gamma_s = \Gamma_T - (\Gamma_m + \Gamma_o) = 3.5\% - (0.004\% + 2.84\%) = 0.66\%$$

In words, of the 3.5% increase in earnings differentials between male and female workers, 0.66% is associated with a disproportionate growth in male employment in the higher paying mining sector.