Blue Collar Booms and American Mortality: Evidence from the Fracking Revolution*

Vivek S. Moorthy† and Paul Shaloka‡

February 2024

Abstract

We exploit the positive labor demand shocks driven by the fracking boom to investigate whether a large and persistent increase in economic opportunity reduces mortality. We use variation in geological characteristics amenable to fracking within a difference-in-differences design and confirm that the boom leads to sizeable increases (2-3%) in earnings and employment for both men and women that do not abate for up to six years after widespread fracking begins. While we find that overall mortality decreases by 2%, these effects are not driven by significant reductions in external causes of mortality like suicides. We instead show that treatable, internal causes (primarily cardiovascular deaths) drive the overall fall in mortality. We further find evidence that health insurance coverage increases in the wake of the boom, suggesting that the non-pecuniary benefits of employment may be a mechanism through which improved labor market outcomes reduce mortality.

JEL: I12, I15, J23, Q40, R12, R58

Keywords: Mortality, Suicides, Labor Demand, Fracking, Regional Development

*We thank Taryn Dinkelman, Bill Evans, Dan Hungerman, Ethan Lieber, and participants in the Notre Dame microeconomics seminar for their thoughtful comments and feedback. We gratefully acknowledge the financial support provided by the Institute for Scholarship in the Liberal Arts at the University of Notre Dame. All errors are our own.

†College of the Holy Cross, email: vmoorthy@holycross.edu

‡US EEOC, email: paul.shaloka@eeoc.gov

I Introduction

A growing body of research shows that the negative impacts of job loss permeate well beyond the labor market. The most pernicious of these effects center around health. Job loss increases BMI, alcohol consumption (Deb et al., 2011), depression (Schaller and Stevens, 2015), and can also be fatal, increasing overall mortality (Eliason and Storrie, 2009). However, much of this literature focuses on short term job losses, and a smaller literature looks the mortality consequences of larger long term negative shocks. Further, there is even less evidence on how sizeable and persistent *increases* in labor demand impact mortality. Whether these types of shocks necessarily lower mortality is not ex ante obvious, as short term income receipt has been found to increase certain causes of mortality (Ruhm, 2000;Moore and Evans, 2012). The lack of evidence is partially due to the relative difficulty of finding quasi-exogenous variation that drives large changes in labor demand.

This paper overcomes that challenge and considers the effect of large, sustained, localized labor demand shocks on mortality by exploiting variation in the intensity and location of the hydraulic fracturing (fracking) boom. Feyrer et al. (2017) found that fracking led to the creation of over half a million jobs, with positive spillovers beyond just the mining industry, suggesting that the boom was transformative for local communities. To measure the mortality effects of the fracking boom, we use proprietary data from the National Vital Statistics System (NVSS) to construct mortality rates at the county-level from 1990 to 2018. The setting and these data give us the rare opportunity to consider the effects of large-scale improvements in economic opportunity on an important health outcome.

To estimate the relationship between labor demand shocks and mortality we have to overcome two main challenges. First, local regulations on drilling operations can limit or even outright ban fracking, and these decisions may be directly related to factors which could influence mortality such as the strength of local labor markets and investments in public health. Second, places which benefited from the boom may differ from areas of the country with no fracking potential; for example, increasing opioid mortality was initially more of a rural phenomenon (Rigg et al., 2018). To address these issues, we use a county-level measure of the potential profitability of fracking operations provided by Rystad, a private energy company.

Crucially, this profitability measure is based on detailed geographic surveys, rather than the potentially endogenous realized level of extraction.[4]

We employ a difference-in-differences (DD) strategy that compares counties with higher geological potential for fracking to similar geographically adjacent counties that were less likely to benefit from the fracking boom. We do this in practice by comparing counties in the top-quartile of our profitability measure to others with lower potential within the same shale formation, referred to as shale plays, which are the geological formations that are amenable to fracking.[5] We also use the differential timing of the adoption of modern fracking technologies across shale plays, which enabled producers to construct wells over under-surveyed and previously inaccessible fossil fuel deposits. We find that while counties had similar levels of production and economic activity prior to fracking adoption, there is a sizeable separation in economic activity between our treatment and control counties after the boom begins. Overall, employment and earnings increase by 2-3% over the 6 years following the start of fracking, and the effects increase over time. We show that although the direct beneficiaries of the boom are men who are more likely to be employed in the mining and transportation sectors, women also experience earnings and employment gains, likely through local equilibrium effects such as agglomeration (Allcott and Keniston, 2018).

We then show that overall mortality declines in boom counties. This result is driven by working-age individuals (15-64) who are the direct beneficiaries of the earnings and employment increases. This group experiences a 2% reduction in all-cause mortality. Sommers (2017) finds that state-level Medicaid (a public means-tested insurance program for the poor) expansions reduced overall mortality for working-age adults by 6%, suggesting that while our estimates are sizeable, they are substantially below the mortality responses induced solely by health insurance expansions. We further show that our mortality results are not driven by differential trends in mortality prior to the fracking boom, and find little evidence for migratory responses driving our results.

To better understand the mechanisms underlying the reduction in mortality, we explore changes by more specific causes of death. We show that the fall in mortality attributable to

[4]Bartik et al. (2019) pioneer this identification strategy.

[5]There 16 Shale Plays in our sample, which constitute contiguous counties across different regions of the U.S.

the fracking boom is driven by reductions among treatable, internal causes of death, with the largest declines concentrated in the latest treatment years. This is consistent with Browning and Heinesen (2012), who find that job loss *increases* the risk of internal mortality using administrative data on workers and plant closures from Denmark. Similar to that study, we find that circulatory/cardiovascular mortality drives the changes in internal causes of death. However, unlike previous studies on plant closures (Browning and Heinesen, 2012, Venkataramani et al., 2020), macroeconomic downturns (Hollingsworth et al., 2017), or large persistent negative shocks (Pierce and Schott, 2020), we do not find significant declines in external causes of death like suicide or drug-overdoses, although our point estimates are negative.

There are many potential mechanisms through which improved labor market opportunities could reduce internal causes of death. Additional income is associated with better health (Chetty et al., 2016), and there are non-pecuniary benefits of employment such as increased self-worth (Noordt et al., 2014). However, we do find suggestive evidence that health insurance coverage increases in boom counties by matching our fracking data to county-level coverage estimates constructed from the American Community Survey (ACS).[6] Increases in health insurance coverage have been shown to lead to sizeable mortality declines (Goldin et al., 2021), and Schaller and Stevens (2015) find that workers who lose a job which was their primary source of insurance reduce doctor's visits and prescription drug usage. While Moore and Evans (2012) find that increased income receipt leads to short-run mortality spikes over the following several days, the alternative mechanisms discussed here suggest that our results are driven by very different mechanisms. Additionally, our results are over a longer time period and are based around a sharp, discontinuous unexpected change in employment and earnings rather than receipt of expected payments. These suggest the benefits from longer term employment gains may improve health through expanding access to medical care.

Our paper contributes to work on the effects of labor market outcomes on health and mortality outcomes. While the existing literature has exploited plant closures to generate

[6] Anecdotally, fracking jobs provided fairly robust health insurance. Surveys from Rigzone, a large online oil and gas industry job posting site and career network platform, show that "Oil and gas professionals have become quite accustomed to rich health benefits offerings". An industry health consultant even bemoans the fact that generous health packages have become expected and simply providing good coverage does not grant a competitive advantage in attracting employees. https://www.rigzone.com/news/survey_shows_oil_gas_workers_want_rich_health_benefits-19-sep-2019-159825-article/

quasi-experimental variation in labor market opportunities, we consider the effects of plant (fracking well) *openings* on labor demand and mortality. It is not obvious ex-ante whether the size of the effects we observe would be of similar magnitude to these studies. The shock and stress of job loss are likely to have consequential, immediate health impacts, which may lead to important non-linearities in the effect of employment changes on health outcomes. Iizuka and Shigeoka (2021) finds that demand responses to price increases for child healthcare are twice that of the change induced by price decreases, suggesting increases in income and coverage may not induce as dramatic changes in behavior as decreases along those dimensions.

We can compare our results to the closest papers in the literature to our own study. Sullivan and Von Wachter (2009) exploit plant closings in Pennsylvania and find that sustained employment and earnings losses of around 10% after a decade leads to a 17% increase in mortality, with the effects being larger for displaced workers under 55. Using Danish administrative data, Browning and Heinesen (2012) finds that job displacement leads to slightly smaller earnings declines over a 20 year window following the initial job loss, and that overall mortality increases by almost half the amount found by Sullivan and Von Wachter (2009). Similar to us, Browning and Heinesen (2012) find that changes in mortality from circulatory disease are an important dimension for explaining the overall mortality results. The reductions in mortality we observe relative to the change in earnings and employment are similar in magnitude, suggesting a symmetric response.

We also contribute to the literature on "deaths of despair" by providing some of the first evidence of the effects of a large positive shock to local labor markets. The Case and Deaton (2017) hypothesis that labor market conditions, especially over the long-run and at the time of entry into the labor market, suggests that the fracking boom may lead to reductions in "deaths of despair", and implies that this overall decline may be driven by reductions in external causes of death. However, while our point estimates suggest reductions in suicides and drug overdoses for all working age adults, these results are imprecise. A number of papers find that increased opioid mortality is largely driven by supply-side changes in opioid availability (Currie and Schwandt, 2020, Alpert et al., 2022), suggesting that there is less of a role for increased economic opportunity to play in reducing deaths of despair.

Our paper also contributes to the literature outside of economics on the health effects of hydraulic fracturing.[7] Jemielita et al. (2015) and Denham et al. (2019) suggest that increased fracking correlates with higher hospitalization rates, but these results do not necessarily have clear implications on mortality. The overall health and mortality effects of the fracking boom remain understudied. Despite this, the state of New York banned fracking in 2014, citing health concerns.[8] This paper provides some of the first causal evidence on an economically significant and relevant policy question.

II Background on the Fracking Boom

Oil and natural gas firms drill traditional wells vertically above large concentrated fossil fuel reservoirs. By contrast, unconventional fracking wells exploit far more dispersed fossil fuel reserves that remain trapped within sedimentary, organic-rich rock formations called shale plays. Companies began limited drilling of these shale plays began as early as the 1960s, but the low permeability of the shale prevents oil and gas from pooling into the reservoirs conventional wells are typically drilled over, rendering traditional production techniques unprofitable.

New advancements in horizontal drilling and hydraulic fracturing enabled the fracking boom. Horizontally drilled wellbores can access large areas of shale at once, obviating the need to drill many vertical wells. Fracking also involves injecting a highly pressurized slurry into the wellbore, which fractures the surrounding shale and allows the encased oil and natural gas to flow freely. While the presence of a shale play is a necessary condition for fracking, actual production is sensitive to several geological factors, including the permeability of the rock, as well as the size and density of the hydrocarbon deposits.

Oil and gas firms did not immediately adopt the new technologies that enabled widespread, profitable fracking, partially because private and academic researchers were initially unaware

[7] Further, we contribute to work that exploits the fracking boom to test how economic opportunity impacts other behavior such as fertility (Kearney and Wilson, 2018) and crime (Street, 2018).

[8] Speaking about the ban Department of Health Commissioner Howard Zucker said: "Would I live in a community with HVHF [high-volume hydraulic fracturing] based on the facts I have now? Would I let my child play in the school field nearby, or my family drink the water from the tap or grow their vegetables in the soil? After looking at a plethora of reports ... my answer is no. The potential risks are too great; as a matter of fact they are not even fully known. Until the public health red flags are answered, I cannot support high volume hydraulic fracturing in the great state of New York."

of the true magnitude of the hydrocarbon reserves. For example, the US Geological Survey estimated in 2002 that the Marcellus Shale (covering WV, PA and NY) held two trillion cubic feet of recoverable natural gas. By 2011, these estimates had risen to 84 trillion cubic feet, based on new surveys. This large correction highlights how little understood the shale deposits were before they became exploitable.[9] Figure 1 Panel B plots the dramatic increase in fracking production over time from 2000, where it accounted for barely any of total US oil and natural gas production, to 2014, when it overtook the output of more traditional methods..

Both academic researchers and the popular press have linked the "fracking revolution" to labor market opportunities. Maniloff and Mastromonaco (2017) review various studies of both the local and national earnings gains attributable to fracking, and document estimates of wage growth which range from 2.6% to 16.75%. While the initial job growth is concentrated in the mining industry, the operation of even a single fracking well involves over 6,000 one-way trucking trips (Xu and Xu, 2020) to haul the water and sand needed for the hydraulic fracturing process. Finally, Allcott and Keniston (2018) find that the manufacturing sector actually grows overall in the wake of natural resource booms in the US (driven by upstream and locally-traded sub-sectors), and so there is little evidence of negative spillovers caused by a "Natural Resource Curse".

III Data

We aggregate all our data to the county-year level. We use county definitions as of the 2000 decennial census,[10] and our main sample includes data from 1990 to 2018. As we discuss below, our empirical strategy only compares counties under the same shale play, thus we omit counties that do not intersect with a shale play from our main sample. We further omit two Texas counties with several years of missing mortality data, including Loving Texas, which has fewer than 100 residents as of the 2020 Census. This leaves us with 519 counties (112 of

[9]Source: https://www.usgs.gov/news/usgs-estimates-214-trillion-cubic-feet-natural-gas-appalachian-basin-formations.

[10]If county boundaries change over time, we aggregate to the 2000 boundary definitions using initial population weights. For example, in 2001, Broomfield, Colorado is created from parts of Adams, Boulder, Jefferson, and Weld counties, and the Census Bureau reports the resulting population loss for each of the original counties. Source: https://www.ddorn.net/data.htm

which are in the top-quartile of the within-play RPI) and 29 years of data.

III.A Fracking Data

The U.S. Energy Information Administration (EIA) provides shape files defining every known shale play, which we use to identify counties that have any fracking potential. We also take data on the monthly prices of oil and natural gas from the EIA.[11] We obtained well-level production data from Enverus, a private oil and gas software company, through their academic outreach initiative. These data include information on both production and the orientation of the wellbore, which we use to identify fracking wells.[12]

To capture variation in fracking suitability *within* shale plays, we purchased the NASMaps product from Rystad Energy, a private energy research company. The company produces a Rystad "prospectivity index" (hereafter referred to as RPI), a continuous, non-linear measure of how amenable a specific location within a shale play is to fracking production. Importantly, this measure is not based on realized/actual fracking production, but only on the underlying geological potential of an area. The index ranges from zero to five, with larger numbers representing increased potential fracking yields. We aggregate this measure to the county level, and we show which counties have any fracking potential (RPI greater than zero) in Figure 1 Panel A. Since the methodology used to calculate the RPI is unique to each play, the measure is not directly comparable across broad geographic areas. We therefore follow Bartik et al. (2019) and identify counties which are in the top-quartile of the prospectivity index within each shale play, and these counties (which are more likely to be the most productive: our treatment counties) are shaded darker in Figure 1 Panel A.

While Bartik et al. (2019) has shown that counties within the same shale play are more comparable along many economic dimensions, our analysis requires that these counties be comparable along dimensions that are relevant to our mortality outcomes. We confirm whether our control counties provide a good counterfactual to our top-quartile counties by comparing them along a variety of county-level characteristics from 1990 Census (well before the tech-

[11] For oil prices we use the Cushing, Oklahoma spot price for West Texas Intermediate crude oil, and for natural gas we use the city-gate price.

[12] We identify fracking wells as any well with a non-vertical wellbore orientation. DrillingInfo, the production database provided by Enverus, is also used by the EIA for their official releases concerning US production.

nology that enabled fracking was first applied), including median household income, share of the population that are veterans, and other demographic information. Different demographic groups have varying propensities for succumbing to deaths of despair, and may also be differentially located across shale play counties. Controlling for these baseline characteristics can improve precision and ameliorate worries that our estimates are being driven by factors besides the fracking boom. Table 1 shows baseline 1990 summary statistics for top-quartile and other shale play counties, and shows that there are no statistically or economically significant differences (in terms of 1990 characteristics) between treatment and control counties prior to the boom.[13]

In addition to the cross-sectional variation in fracking potential, the timing of fracking adoption varied across shale plays. The gray bars in Figure 1 Panel B indicate the number of shale plays for which fracking potential became public knowledge in that year, which we take from Bartik et al. (2019). While firms begin exploratory adoption of new fracking technologies in the Barnett shale play in Texas as early as 2001, more well-known fracking hot spots like the Barnett shale play in North Dakota and the Marcellus Shale plays in the Mid-Atlantic do not begin widespread fracking production until 2007 and 2008, respectively. Despite an initial lag, top-quartile RPI counties produce substantially more than other three quarters of shale play counties combined.

III.B Employment and Earnings Data

We use county-level data on earnings and employment from the Quarterly Workforce Indicators (QWI) database, which is an aggregation of micro-level records from the Longitudinal Employer-Household Dynamics (LHED). These data are primarily based on unemployment insurance earnings data from participating states[14] available for a limited number of two-way group tabulations, including sex-age and sex-education. We focus on aggregate changes to employment and earnings instead of restricting attention to the natural resource extraction

[13] We show these same baseline differences across all shale play counties and the rest of the lower 48 states in Appendix Table 1. Shale play counties are poorer and more white than the rest of the country, although residents are more likely to be married. Shale play counties also have a lower age-adjusted death rate per 100,000 residents in 1990.

[14] In the earlier years of our sample, The QWI has limited coverage, which leaves very few observations prior to 2002, after which we have full coverage of every county in our main sample. The main earnings and employment results are robust to limiting the sample to years where we have data on all shale play counties, as shown in Appendix Table B.3 and Appendix Table B.4.

industry. Previous work on agglomeration such as Greenstone et al. (2010) suggests that the opening of large work sites may create positive spillovers for other industries, and Feyrer et al. (2017) finds evidence for such spillovers in response to the fracking boom. We aggregate our main variables of interest, average quarterly earnings and total quarterly employment to the yearly level. Specifically, we take the simple average of employment, and the employment-weighted average of earnings across all 4 quarters in a year.

III.C Mortality Data

We use a restricted-access version of the National Vital Statistics System (NVSS) mortality files from 1990-2018, which represents a census of all deaths in the United States. These data identify basic demographic information, primary/additional causes of death, and contain identifiers for the county of residence and occurrence. We follow Stevens et al. (2015) by separating all causes of death into mutually exclusive categories,[15] further separated into whether the causes of death are internal (cancer, cardiovascular, etc.) or external (homicides, motor vehicle accidents, etc.). For external causes of death, we also include "deaths of despair": suicides, drug-related deaths, and alcohol-related deaths, using the definitions provided by the US Congress' Joint Economic Committee. Since our data span across the use of ICD-9 and ICD-10 codes for reporting causes of death, this use of consistent, broad categories ensures comparability across time.

Our primary outcome is the inverse hyperbolic sine (IHS) of the number of deaths, where the population of the relevant demographic group is included as a control. We take these population data from estimates constructed by the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. The IHS transformation allows us to retain county-year observations with zero deaths, which occur for some of the more uncommon causes of death. We also consider age-adjusted mortality rates per 100,000 residents. While the crude death rate is just the total number of deaths for a specific demographic group divided by the relevant population, the age-adjusted death rate is a weighted average of crude death rates across standard age categories, where the national population shares in those age categories

[15] We do not consider two of the broad death categories used in Stevens et al. (2015): other/unspecified/ill-defined and miscellaneous, in part because this latter category includes drug abuses which we look at separately.

in 2000 are the weights.[16]

IV Empirical Strategy

Advancements in horizontal drilling and slickwater fracturing enabled the extraction of previously inaccessible reserves of oil and natural gas from shale plays. However, since the level of production is endogenous with respect to local labor market characteristics and the regulatory environment, simple comparisons using this measure may introduce bias. For instance, high-productivity areas may have had upward trending economic growth which enabled more widespread and earlier adoption of fracking technologies. In addition, different environmental restrictions and/or zoning regulations may be correlated with factors that influence mortality such as underlying economic trends or the level of public investment. Appendix Table 1 reports baseline comparisons across all shale play counties and the rest of the lower 48 states. Shale play counties are poorer and more white than the rest of the country, and have a lower age-adjusted death rate per 100,000 residents in 1990. These highlight the necessity of carefully constructing the appropriate counterfactual.

Following the approach pioneered by Bartik et al. (2019), we use variation in the RPI and the fracking adoption to account for these issues. The RPI is a time-invariant function of pre-existing geological features that determine both the intensive and extensive margins of potential fracking production. This provides a straightforward approximation of the exogenous variation within a play that determines the extraction potential of fracking wells, and therefore the intensity of the positive labor demand shock. Combining this with temporal variation in the initiation of fracking in each play leads to the following DD specification:

$$y_{cpt} = \beta(\textit{Top-Quartile}_{cp} \times Post_{pt}) + \sum_{t} \Psi_t(\mathbf{I}_{year=t} \times \mathbf{X}_{c,1990}) + \lambda_c + \gamma_{pt} + \epsilon_{cpt} \tag{1}$$

where y_{cpt} is the outcome of interest. $Post_{pt}$ is an indicator for whether shale play p had adopted fracking by year t. $\textit{Top-Quartile}_{cp}$ indicates whether county c is in the top-quartile of

[16]This adjustment is standard in the literature, and accounts for changing demographic patterns over time.

the RPI for shale play p. Ψ_t captures the potentially time-varying effects of $\mathbf{X}_{c,1990}$, a vector of initial county-level characteristics.[17] We also control for time-invariant county characteristics with county fixed effects, λ_c. Regressions are weighted by the 2000 population [18] and all standard errors are clustered at the county level.

Including play by year fixed effects, γ_{pt}, captures play-year shocks and ensures our results are based on variation between counties *within* shale plays. These effectively aggregate estimates from each shale play. Since the timing of fracking adoption varies at the play level, adding these fixed effects safeguard our results from issues associated with staggered treatment timing (e.g. Goodman-Bacon, 2018). Due to this design, we drop any county that is not over a shale play. We further drop two Texas counties with missing mortality data (this includes Loving County, Texas, which has a population of 64 in the 2020 census). This leaves us with 407 shale play (control) counties and 112 top-quartile RPI (treatment) counties.

Since the timing of fracking adoption varies across shale plays (Figure 1 Panel B), the simple difference-in-differences coefficient is subject to composition bias as the number of years post-treatment varies across the sample. We restrict the data to a balanced sample, where the balanced sample is defined by whether every shale play experiences the same number of lags and leads in event years. In the main mortality sample, we have data for each play 11 years prior to treatment and 7 treated years (including the year of initiation of fracking), or 18 event-years of data for each observation.[19]

The identifying assumption of our DD model is that the control counties within plays provide an estimate of the counterfactual time-path of mortality and labor market outcomes had fracking intensity been lower in boom counties. While this assumption can never be directly tested we examine whether our treatment and control counties have the same pre-treatment trends by running the following event study specifications where we replace the $Post_{pt}$ indicator with a vector of event year indicators, omitting the event year prior to fracking's introduction:

[17] All of the variables shown in Table 1 are included as controls aside from the initial age-adjusted death rate.
[18] We show robustness to these weights in the Appendix.
[19] We show that our results are robust to using an unrestricted unbalanced sample in Appendix Figure A.4.

$$y_{cpt} = \sum_{n \neq -1} \beta_n(\textit{Top-Quartile}_{cp} * \mathbf{I}_{year-\tau_p=n}) + \sum_t \Psi_t(\mathbf{I}_{year=t} * \mathbf{X}_{c,1990}) + \lambda_c + \gamma_{pt} + \epsilon_{cpt} \quad (2)$$

where τ_p represents the time of initiation of fracking in a given play. The coefficients β_n trace out the difference in outcomes between top quartile and other counties within a play, in a given event year n relative to omitted year. A key feature of our identification strategy is that the RPI accurately predicts the highest intensity boom counties in terms of actual production. Figure A.1 shows a flat, almost nonexistent pre-trend in production followed by an immediate increase after the boom begins, although production does not really begin to take off until the second and third year after the adoption of fracking technologies.[20] We also show that production increases in a similar, albeit attenuated, manner whether we define treatment using our standard top-quartile definition or whether we expand treated counties to include counties above the median play-level RPI measure or simply look at a standard deviation shift in the actual underlying RPI values. We can see that fracking production is nearly $400 million greater in top-quartile counties six years after the boom begins, indicating a meaningful separation between treatment and control counties.

V Results

V.A Earnings and Employment

To identify who benefited the most from fracking, Figure 2 shows the gender specific results of the fracking boom for earnings and employment using Equation (2). Panel A and Panel B report estimates for the log of average earnings for all employees, while Panel C and Panel D show results for the log of the average employment to population ratio. The dark and lighter shaded bars represent the associated 90% and 95% confidence intervals, respectively. Results from estimating Equation (1) are shown above each event study.

[20] Since treatment timing is determined by when fracking became public knowledge within a play, and because hydrocarbon deposits within shale plays were relatively under-surveyed, an initial lag in production is not surprising.

Overall, Figure 2 shows that earnings and employment increased for both men and women in the wake of the fracking boom, and continued to do so for up to six years after the adoption of fracking technologies. While the average effects show a roughly 3% increase in earnings and employment for men, the coefficients for later event-years are larger and seem to finally begin settling closer to 3-4%. Since our specification only uses within-play variation, and because fracking production is also increasing in our control counties (Figure 1 Panel B), our results do not represent fracking's *overall* impact, but instead leverage variation in plausibly exogenous production ability. Thus these labor market effects are likely smaller than the overall impact of fracking.

Despite the anecdotal evidence that fracking is an almost entirely male-dominated field, we find wage and employment growth for women.[21] For women, both earnings and employment increase by around 2%. However, both Bartik et al. (2019) and Feyrer et al. (2017) show that the boom led to substantial positive spillovers to other industries, with Feyrer et al. (2017) finding that in 2012, half of the overall employment increases attributable to the boom were actually sectors not directly related to extraction, while 30% were concentrated in the transportation sector and only 20% of the overall increase in employment came from the mining sector.

Kearney and Wilson (2018) also find differential sizes of the male and female labor demand shocks in response to the boom, and they find slightly larger effects for male earnings (4%) and the employment-to-population ratio (5%) than our main results suggest. Firstly, we show in Appendix Table B.3 and Appendix Table B.4 that when we do not include county-level population weights our results are roughly similar to Kearney and Wilson (2018), suggesting that some more sparsely populated counties experience the largest relative production booms which men were differentially able to benefit from.[22]

[21] According to the US Bureau of Labor Statistics, men made up over 84% of the workforce in mining, quarrying, and oil and gas extraction industries as of 2019.

[22] Additionally, the Kearney and Wilson (2018) measure boom intensity using a simulated instrument for actual production, comparing fracking counties to others within the same state. The RPI is not constructed from contemporaneous production, and compared counties within the same shale formation.

V.B Overall Mortality Results

Having confirmed that fracking adoption led to a sustained increase in labor demand, we now consider the reduced-form effects of fracking amenability on mortality. Primarily this is because we are unable to disentangle whether the increase in earnings or the increase in employment and the associated non-pecuniary benefits (increased optimism, access to health insurance, etc.) are leading to changes in mortality.[23]

Table 2 looks at the effects of the fracking boom on overall mortality. The dependent variable is the inverse hyperbolic sine of overall mortality. In additional to all controls mentioned earlier, these specifications also include the log of the population of the relevant group as an independent variable. Columns (1) shows that overall mortality fell by 2.1% in top-quartile counties relative to their shale play counterparts. Columns (3)-(6) examine the mortality effects by gender. These reveal a roughly 2% decline in overall mortality for men, and an even larger in magnitude 3% decline in mortality for women.

We also check whether our results are driven by migration. For example, Boslett and Hill (2022) find that migration is a potential driver of their mortality results, and Arthi et al. (2022) highlights the importance of migration in impacting measured mortality. Wilson (2020) found a sizeable migration response to the fracking boom, particular in the Bakken Shale play, intersecting North Dakota and Montana. Thus, we first examine whether migration is influencing our results by following Kearney and Wilson (2018) and omit both North Dakota and Montana from our sample. Columns (2), (4), and (6) of Table 2 show that our mortality results remain of similar magnitude without the inclusion of these high in-migration regions. We further test for migratory responses more directly by estimating (1) with the log number of men and women as the dependent variables. Table B.2 shows these results of these regressions. We find statistically insignificant and modestly sized coefficients.[24]

Figure 3 Panel A plots the estimates from equation (2) for overall Mortality.[25] Panels B and C separately examine overall mortality for men and women, respectively. The shaded

[23]The reduced-form estimates are also more relevant for local municipalities who are deciding on whether or not to allow fracking, since the RPI is measurable before any drilling begins.

[24]Recall for our identification strategy, in or out migration would have to systematically vary between top quartile and other counties within the same shale play to be driving our results.

[25]These figures continue to drop North Dakota and Montana given the unique nature of the Bakken Shale play, however, as in Table 2 the estimates are largely the same with these states included.

bars represent the associated 95% and 90% confidence intervals. Overall, there is reassuringly an absence of differential trends in mortality between treatment and control counties prior to the initiation of fracking. After fracking, there begins a decline in overall mortality. While imprecisely estimated, there is around a 1% decline in mortality starting 3 years following the initiation of fracking. This effect grows to a statistically significant 4% reduction in mortality six years after fracking.

Panel B Looks at Men. Here, while the point estimates do suggest up to 3% decline in overall mortality six years after fracking, the individual point estimates are not statistically significant. For women, Panel C shows that starting 4 years after the initiating of fracking, there is a statistically significant decline in overall mortality, with a 5% decline after six years.

Although we observe overall declines in mortality, we may expect these declines to be driven by the working-age population, given that some combination of income and the non-pecuniary effects of work are the most plausible explanations of the effects we observe. We explore heterogeneity in effects by age in Appendix Table B.6 and Appendix Figure A.3, where the outcome is now just the IHS of the number of deaths for different age groups, and the contemporaneous population counts for that demographic group are included as controls. Here, we can see that the significant declines in mortality are all concentrated among working-age adults 15-64.

Further, we examine whether the results are robust to alternative measure of mortality, specifically the age-adjusted mortality rate, discussed in Section *III.C*. Column (1) of Appendix Table B.5 shows that the overall age-adjusted mortality rate fell in top-quartile counties relative to their shale play counterparts by 8.9 deaths per 100,000 people, which is a 1% decline terms of the sample period mean. This effect remains largely the same when excluding Montana and North Dakota from the sample in Column (2). Columns (3) and (4) show that the coefficients for men are negative and of similar magnitude to the combined death rate, and the remaining coefficients show larger in magnitude declines for women, in line with the IHS specification.[26]

While we do not explicitly consider instrumenting the change in overall earnings or employ-

[26] Appendix Figure A.2 shows the estimates from (2) using the age-adjusted mortality rate.

ment, we can consider what our estimates imply about the elasticity of mortality with respect to the observed change in either variable. However, we caution that these comparisons may be misleading, because changes in both income and employment are occurring at the same time, so simply scaling our mortality results by the magnitude of one of these changes does not consider all of the pecuniary and non-pecuniary changes as a result of a labor demand shock at once. If we take the 2.2% decline in overall working-age mortality from Column (2) of Table 2 and the 2.4% increase in overall wages from Column (4) of Table B.3, our estimates suggest that a 1% increase in wages leads to a 0.92% decline in overall mortality.

V.C Heterogeneity by Cause of Death

While our results so far only consider overall mortality, there may be important differences in effects by cause of death. For example, even though overall mortality declines, increased drilling may reduce air and water quality leading to additional respiratory related deaths. We first consider whether the declines in mortality are driven by internal causes of mortality, such as cardiovascular mortality, or in external causes, such as suicides and homicides. Panels A and B of Figure 4 present estimates of equation (2) for the IHS of internal and external causes of death, respectively, with the estimates from equation (1) presented above. The decline in mortality is driven by declines in internal causes of mortality. Top quartile counties experience a 3% decline in internal causes of death relative to other counties.

Figure 5 explores this heterogeneity further, looking at specific types of death within each category for men and women separately. Panel A looks at internal causes. Each coefficient is estimated from separate regressions, with the bands showing both 90 and 95% confidence intervals.

While most coefficients are negatively signed, there are statistically significant declines in cardiovascular related deaths, infection related deaths (tuberculosis, whooping cough, etc.) and kidney/urethra related deaths (renal failure, kidney infections, etc) for men. For women, there are statistically significant declines in cardiovascular deaths, however there are large in magnitude but imprecisely estimated declines in other internal causes of death, such as respiratory and nutrition-based deaths. These results match Browning and Heinesen (2012), who use plant closures linked with administrative data from Denmark to show that job loss leads to increased risks of mortality from circulatory (cardiovascular, e.g. myocardial infarctions and strokes) causes of death.

Why do we observe reductions in these causes of death? Wherry and Miller (2016) finds substantial increases in high cholesterol diagnosis following Medicaid expansion, and cardiovascular drugs are known to reduce mortality within months of treatment (Aronow et al., 2001; Cannon et al., 2004). Increased income, in addition to expansions in health insurance through increased employment (discussed in Section *V.D*) could all lead to increased access to these treatments. Likewise, Medicaid expansion has been linked to increased access to vaccinations and antibiotics which can reduce death from infectious diseases (Lu et al., 2015), and lower indices of kidney failure among non-elderly adults (Thorsness et al., 2021).

Panel B examines external causes of death. Overall, we find no statistically significant change in external causes of death, however the point estimates provide some suggestive evidence of movements. For example, the positively signed coefficients for motor vehicle accidents are consistent with Moore and Evans (2012), who find that traffic accidents are pro-cyclical. Additionally, transportation jobs are an important driver of the employment growth in response to the fracking boom (Bartik et al., 2019; Feyrer et al., 2017); the operations of just a single well can involve hundreds of commercial truck trips (Goodman et al., 2016) to haul the water and particulate matter needed for hydraulic fracturing.[27]

We find negatively signed and relatively large point estimates for drug overdoses, which is consistent with Pierce and Schott (2020), however we cannot reject a null effect at the 5% significance level. Similarly, we also find negatively signed but imprecisely estimated impacts on suicides which are larger for women. This is consistent with Dow et al. (2020), who find that labor market policies reduce suicides, with stronger effects for women. Overall, while Pierce and Schott (2020) find that reductions in labor demand lead to increased death of despair, we ultimately do not find robust evidence to conclude that even relatively sizeable and sustained *increases* in earnings and employment reduce these causes of mortality. While longer-run sustained increases in economic opportunity could lead to reductions in deaths

[27] Several law firms in Texas (https://www.daxgarzalaw.com/blog/fracking-and-oilfield-trucking-dangers/) and Pennsylvania (https://www.rosenbauminjuryfirm.com/practice-areas/fracking-accidents-damages/fracking-related-truck-and-transportation-accidents/) even specialize in fracking related vehicle accidents.

of despair, our time frame may be too short to observe these changes. Further, increased medical care as a result of expanding insurance and income may actually lead to increases in opioid prescriptions, which could offset any reductions in drug-related mortality.

Appendix Figure A.6 repeats this exercise using age-adjusted death rates. The results for men and women follow the same general pattern, The major exception to this pattern is traffic accidents, which only significantly increase for men. According to the US department of transportation, roughly 80% of the workforce in the transportation sector is male, suggesting that the increased trucking demand and associated traffic risks likely disproportionately affect men.

V.D Mechanisms: Health Insurance Results

While greater income has been closely linked to life expectancy in the US (Chetty et al., 2016), fracking boom counties experienced increases in employment in addition to changes in income. While it is challenging to measure the non-pecuniary benefits of employment such as reduced stress that have been linked to employment opportunities (Marcus, 2013), we can look at suggestive evidence of whether health insurance expanded using data from the Small Area Health Insurance Estimates (SAHIE) Program. The SAHIE is the only source for single-year estimates of health insurance coverage status for all counties in the US, and we use coverage data from 2008-2020. The SAHIE uses data from the American Community Survey (ACS) on whether a person is currently covered by health insurance or health coverage plans to form model-based estimates of coverage.[28] Coverage is estimated based on the proportion of a demographic-group within a specific income category and the proportion insured within that income category. While imperfect, these estimates provide some evidence of changing insurance coverage.

We regress the share of individuals ages 18 to 64 in a county with health insurance on our measure of fracking potential in Table 3. We find evidence that health insurance coverage increased by roughly 1.2 percentage points, or a 1.5% increase off the baseline mean, in the

[28] Although estimates are available for 2005-2007 as well, these prior years use Current Population Survey data with different insurance definitions, and so the results are not comparable across time periods.

wake of the fracking boom. Goldin et al. (2021) show that inducing middle-age adults to enroll in health insurance by informing them of tax penalties led to moderate to large declines in subsequent mortality. Specifically, they find that 1.9% relative increase in coverage led to a per-month effect of coverage on mortality of anywhere from -0.31 to -.0.04 (IV results), a confidence interval which encompasses both very large and moderate reductions. Although we are unable to conduct an IV analysis, the magnitude of our mortality results are much smaller given the insurance coverage increase we observe, although the effects of additional income receipt on some categories of pro-cyclical mortality may be a mitigating factor in our context.

In terms of the specific causes of death that we find are effected by the fracking boom, Thorsness et al. (2021) finds that among non-elderly adults, Medicaid expansion reduces renal failure, and Khatana et al. (2019) find that Medicaid expansion reduces cardiovascular mortality, suggesting that the insurance mechanism can help explain the heterogeneity in causes of death observed in Section *V.C*. Additionally, Schaller and Stevens (2015) find that job loss results in both decreased self-reported health, as well as reductions in doctor's visits and prescription drug usage if employment was the primary source of health insurance. While we are unable to test these latter two mechanisms directly, our findings that health insurance coverage increased suggest that health care utilization may also have increased. Interestingly, Jemielita et al. (2015) finds that increased unconventional drilling is associated with increased hospitalization rates. Although this study is correlational, it does suggest that health care utilization may be increasing with fracking production.[29]

VI Conclusion

While a growing body of evidence finds negative mental and physical health consequences of unemployment, we know less about the role that increased earnings and employment play in terms of mortality. This question has become even more policy salient in recent years, as

[29]Other mechanisms may be at play as well. Bartik et al. (2019) find, using the same source of variation as we do, that local government's increase welfare and hospital expenditures by approximately 24% in the wake of the boom. Although this result was not statistically significant, it suggests that changes in public health investments may also be a contributing factor to the observed mortality declines.

Case and Deaton (2017) have linked declining labor market opportunities to rising suicide, drug-related and alcohol mortality, and the subsequent decline in life expectancy in the US. We show that the positive labor demand shocks driven by the fracking boom led to decreased mortality. While we do not find robust evidence that "deaths of despair" decline in response to these positive labor demand shifts, we do find that treatable, internal causes of death decline. Along with suggestive evidence that health insurance increased, our findings suggest a potential channel behind the positive income and life expectancy gradient (Chetty et al., 2016).

References

Allcott, Hunt, and Daniel Keniston. 2018. "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America." *The Review of Economic Studies* 85 (2): 695–731.

Alpert, Abby, William N Evans, Ethan MJ Lieber, and David Powell. 2022. "Origins of the Opioid Crisis and its Enduring Impacts." *The Quarterly Journal of Economics* 137 (2): 1139–1179.

Aronow, Herbert D, Eric J Topol, Matthew T Roe, Penny L Houghtaling, Katherine E Wolski, A Michael Lincoff, Robert A Harrington, Robert M Califf, E Magnus Ohman, Neal S Kleiman, et al. 2001. "Effect of Lipid-Lowering Therapy on Early Mortality After Acute Coronary Syndromes: An Observational Study." *The Lancet* 357 (9262): 1063–1068.

Arthi, Vellore, Brian Beach, and W Walker Hanlon. 2022. "Recessions, mortality, and migration bias: Evidence from the Lancashire cotton famine." *American Economic Journal: Applied Economics* 14 (2): 228–255.

Bartik, Alexander W., Janet Currie, Michael Greenstone, and Christopher R. Knittel. 2019. "The Local Economic and Welfare Consequences of Hydraulic Fracturing." *American Economic Journal: Applied Economics* 11 (4): 105–55.

Boslett, Andrew, and Elaine Hill. 2022. "Mortality during resource booms and busts." *Journal of environmental economics and management* 115:102696.

Browning, Martin, and Eskil Heinesen. 2012. "Effect of Job Loss Due to Plant Closure on Mortality and Hospitalization." *Journal of health economics* 31 (4): 599–616.

Cannon, Christopher P, Eugene Braunwald, Carolyn H McCabe, Daniel J Rader, Jean L Rouleau, Rene Belder, Steven V Joyal, Karen A Hill, Marc A Pfeffer, and Allan M Skene. 2004. "Intensive versus Moderate Lipid Lowering with Statins After Acute Coronary Syndromes." *New England journal of medicine* 350 (15): 1495–1504.

Case, Anne, and Angus Deaton. 2017. "Mortality and Morbidity in the 21st Century." *Brookings Papers on Economic Activity* 2017 (1): 397–476.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. "The Association Between Income and Life Expectancy in the United States, 2001-2014." *Jama* 315 (16): 1750–1766.

Currie, Janet, and Hannes Schwandt. 2020. *The Opioid Epidemic Was Not Caused by Economic Distress But by Factors that Could be More Rapidly Addressed.* Technical report. National Bureau of Economic Research (NBER) Working Paper 27544.

Deb, Partha, William T Gallo, Padmaja Ayyagari, Jason M Fletcher, and Jody L Sindelar. 2011. "The Effect of Job Loss on Overweight and Drinking." *Journal of health economics* 30 (2): 317–327.

Denham, Alina, Mary Willis, Alexis Zavez, and Elaine Hill. 2019. "Unconventional Natural Gas Development and Hospitalizations: Evidence from Pennsylvania, United States, 2003–2014." *Public health* 168:17–25.

Dow, William H, Anna Godoy, Christopher Lowenstein, and Michael Reich. 2020. "Can Labor Market Policies Reduce Deaths of Despair?" *Journal of Health Economics* 74.

Eliason, Marcus, and Donald Storrie. 2009. "Does Job Loss Shorten Life?" *Journal of Human Resources* 44 (2): 277–302.

Feyrer, James, Erin T. Mansur, and Bruce Sacerdote. 2017. "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution." *American Economic Review* 107 (4): 1313–34.

Goldin, Jacob, Ithai Z Lurie, and Janet McCubbin. 2021. "Health Insurance and Mortality: Experimental Evidence from Taxpayer Outreach." *The Quarterly Journal of Economics* 136 (1): 1–49.

Goodman, Paul S, Fabio Galatioto, Neil Thorpe, Anil K Namdeo, Richard J Davies, and Roger N Bird. 2016. "Investigating the Traffic-Related Environmental Impacts of Hydraulic-Fracturing (fracking) Operations." *Environment international* 89:248–260.

Goodman-Bacon, Andrew. 2018. "Difference-in-differences with Variation in Treatment Timing." National Bureau of Economic Research (NBER) Working Paper 25018.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy* 118 (3): 536–598.

Hollingsworth, Alex, Christopher J Ruhm, and Kosali Simon. 2017. "Macroeconomic Conditions and Opioid Abuse." *Journal of Health Economics* 56:222–233.

Iizuka, Toshiaki, and Hitoshi Shigeoka. 2021. "Asymmetric Demand Response when Prices Increase and Decrease: The Case of Child Healthcare." *The Review of Economics and Statistics,* 1–30.

Jemielita, Thomas, George L Gerton, Matthew Neidell, Steven Chillrud, Beizhan Yan, Martin Stute, Marilyn Howarth, Pouné Saberi, Nicholas Fausti, Trevor M Penning, et al. 2015. "Unconventional Gas and Oil Drilling is Associated with Increased Hospital Utilization Rates." *PloS one* 10 (7): e0131093.

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

Khatana, Sameed Ahmed M, Anjali Bhatla, Ashwin S Nathan, Jay Giri, Changyu Shen, Dhruv S Kazi, Robert W Yeh, and Peter W Groeneveld. 2019. "Association of Medicaid expansion with cardiovascular mortality." *JAMA cardiology* 4 (7): 671–679.

Lu, Peng-jun, Alissa O'Halloran, and Walter W Williams. 2015. "Impact of Health Insurance Status on Vaccination Coverage Among Adult Populations." *American journal of preventive medicine* 48 (6): 647–661.

Maniloff, Peter, and Ralph Mastromonaco. 2017. "The Local Employment Impacts of Fracking: A National Study." *Resource and Energy Economics* 49:62–85.

Marcus, Jan. 2013. "The Effect of Unemployment on the Mental Health of Spouses–Evidence from Plant Closures in Germany." *Journal of health economics* 32 (3): 546–558.

Moore, Timothy, and William Evans. 2012. "Liquidity, Economic Activity, and Mortality." *Review of Economics and Statistics* 94 (2): 400–418.

Noordt, Maaike van der, Helma IJzelenberg, Mariël Droomers, and Karin I Proper. 2014. "Health Effects of Employment: A Systematic Review of Prospective Studies." *Occupational and environmental medicine* 71 (10): 730–736.

Pierce, Justin R., and Peter K. Schott. 2020. "Trade Liberalization and Mortality: Evidence from US Counties." *American Economic Review: Insights* 2 (1): 47–64.

Rigg, Khary K, Shannon M Monnat, and Melody N Chavez. 2018. "Opioid-related mortality in rural America: Geographic heterogeneity and intervention strategies." *International Journal of Drug Policy* 57:119–129.

Ruhm, Christopher J. 2000. "Are Recessions Good for your Health?" *The Quarterly Journal of Economics* 115 (2): 617–650.

Schaller, Jessamyn, and Ann Huff Stevens. 2015. "Short-Run Effects of Job Loss on Health Conditions, Health Insurance, and Health Care Utilization." *Journal of health economics* 43:190–203.

Sommers, Benjamin D. 2017. "State Medicaid Expansions and Mortality, Revisited: A Cost-Benefit Analysis." *American Journal of Health Economics* 3 (3): 392–421.

Stevens, Ann H, Douglas L Miller, Marianne E Page, and Mateusz Filipski. 2015. "The Best of Times, the Worst of Times: Understanding Pro-cyclical Mortality." *American Economic Journal: Economic Policy* 7 (4): 279–311.

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

Sullivan, Daniel, and Till Von Wachter. 2009. "Job Displacement and Mortality: An Analysis Using Administrative Data." *The Quarterly Journal of Economics* 124 (3): 1265–1306.

Thorsness, Rebecca, Shailender Swaminathan, Yoojin Lee, Benjamin D Sommers, Rajnish Mehrotra, Kevin H Nguyen, Daeho Kim, Maricruz Rivera-Hernandez, and Amal N Trivedi. 2021. "Medicaid Expansion and Incidence of Kidney Failure Among Nonelderly Adults." *Journal of the American Society of Nephrology* 32 (6): 1425–1435.

Venkataramani, Atheendar S, Elizabeth F Bair, Rourke L O'Brien, and Alexander C Tsai. 2020. "Association Between Automotive Assembly Plant Closures and Opioid Overdose Mortality in the United States: A Difference-in-Differences Analysis." *JAMA Internal Medicine* 180 (2): 254–262.

Wherry, Laura R, and Sarah Miller. 2016. "Early Coverage, Access, Utilization, and Health Effects Associated with the Affordable Care Act Medicaid Expansions: a Quasi-Experimental Study." *Annals of internal medicine* 164 (12): 795–803.

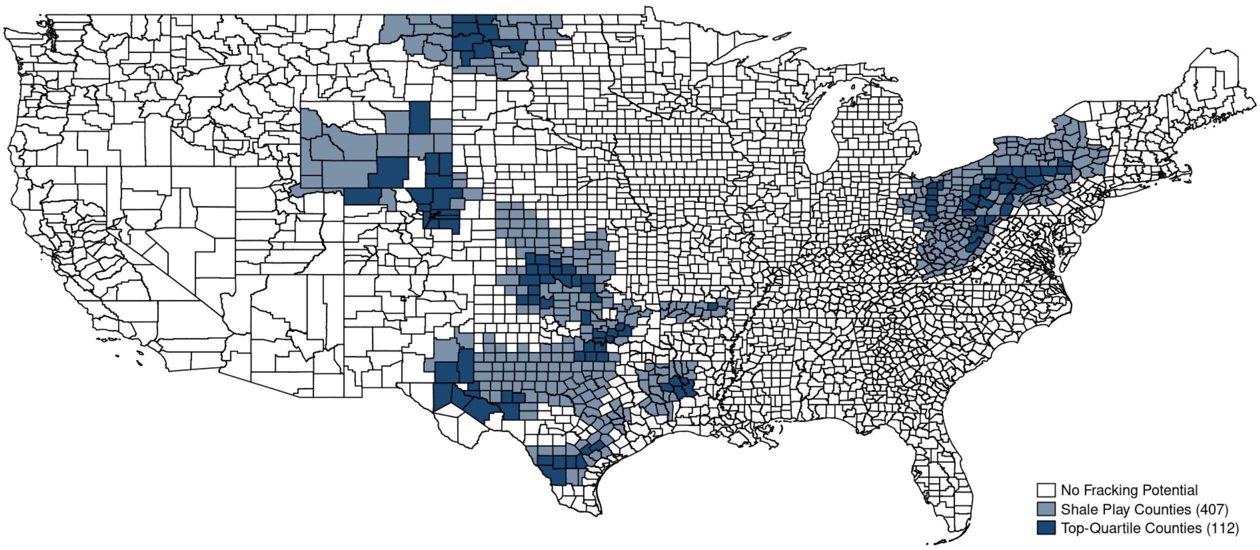
Wilson, Riley. 2020. "Moving to Economic Opportunity: The Migration Response to the Fracking Boom." *Journal of Human Resources,* 0817–8989R2.

Xu, Minhong, and Yilan Xu. 2020. “Fraccidents: The Impact of Fracking on Road Traffic Deaths.” *Journal of Environmental Economics and Management* 101:102303.

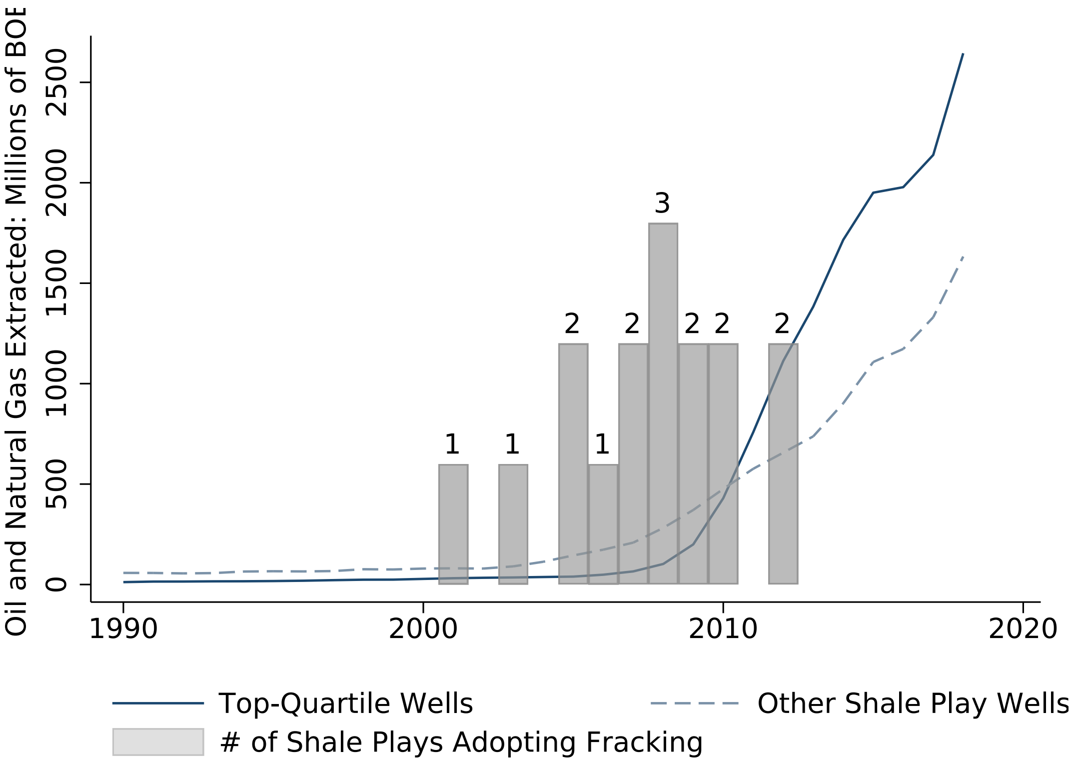
VII Figures

Figure 1: Hydraulic Fracturing Potential and Production - Rystad Prospectivity Index (RPI)

Panel A: Mapping RPI by US Counties



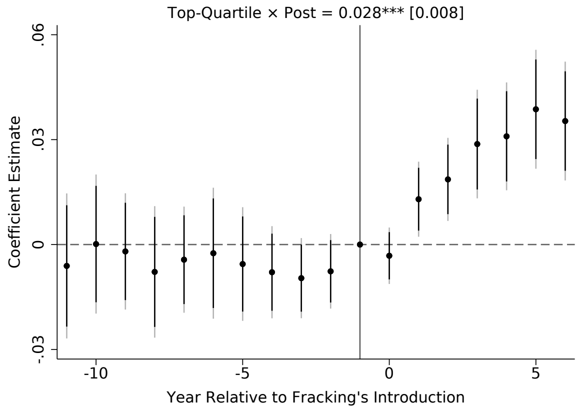
Panel B: County-Level Fracking Production by RPI



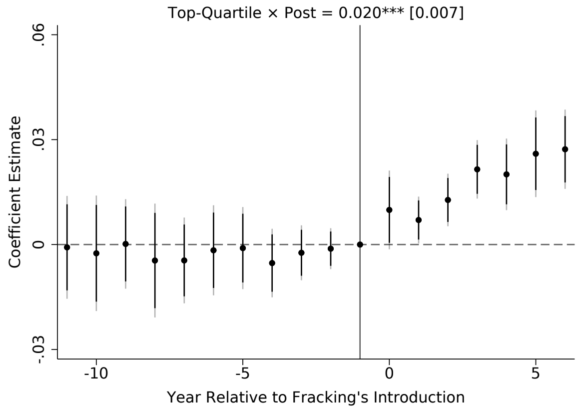
Notes: Panel A plots all US counties from the lower 48 states using 2000 census boundaries. White counties do not intersect with a shale play and are unable to benefit from the fracking boom. Lightly shaded counties (control) intersect with a shale play and are in the bottom three quartiles of the RPI, our measure of fracking potential discussed in Section III.A. Darkly shaded counties (treated) intersect with a shale play and are in the top-quartile of the RPI within a specific shale play. Shale play borders are not shown here for visual clarity. Panel B plots oil and natural gas production measured in millions of barrels of oil equivalent units (BOE) produced by horizontally-drilled wells. These aggregate amounts are calculated from monthly, well-level production data from Enverus. The number of shale plays adopting fracking technology in a specific year (as identified by Bartik et al. (2019), are shown using the shaded gray bars.

Figure 2: Earnings and Employment Effects by Gender

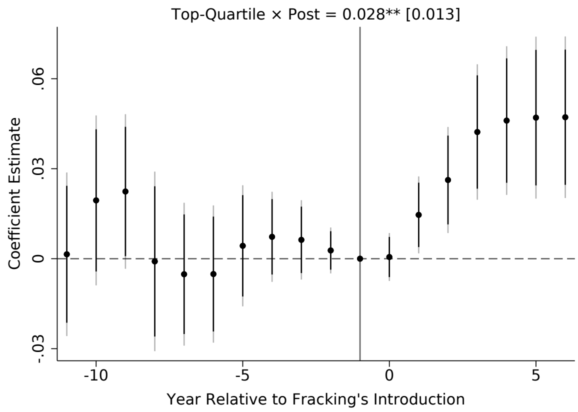
Panel A: Log of Male Earnings



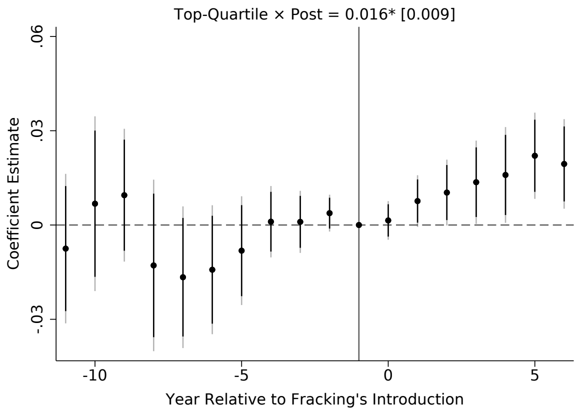
Panel B: Log of Female Earnings



Panel C: Log of Male Employment to Pop. Ratio



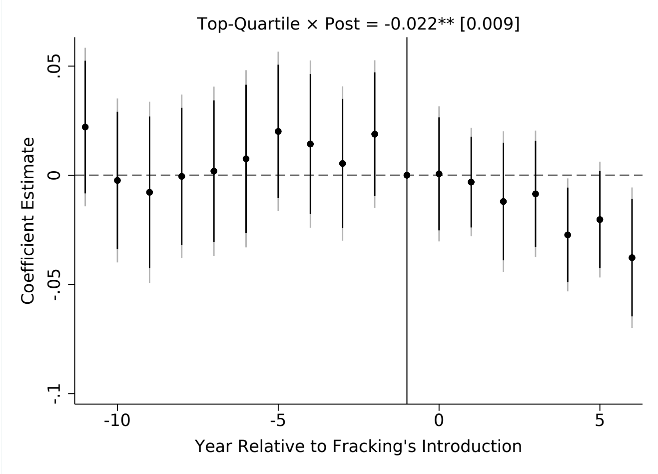
Panel D: Log of Female Employment to Pop. Ratio



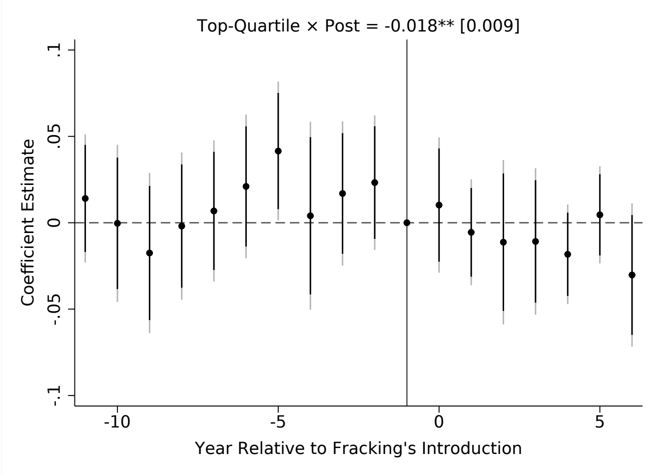
*Notes: Each panel reports the point estimates with their associated 95% and 90% confidence intervals from Equation (2) for the balanced set of event-years. We take earnings measures (adjusted to real 2010 $ amounts) and employment counts from the QWI database. We take population counts from SEER. All values are calculated for 14-99 year old individuals in each county. Each regression include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Each panel omits North Dakota and Montana from the sample. Standard errors are adjusted for clustering at the county level. The difference-in-differences coefficient for each outcome is included above each event-study, as well as the relevant standard error in brackets. *** Significance 1%, ** 5%, * 10%.*

Figure 3: Overall Mortality by Gender

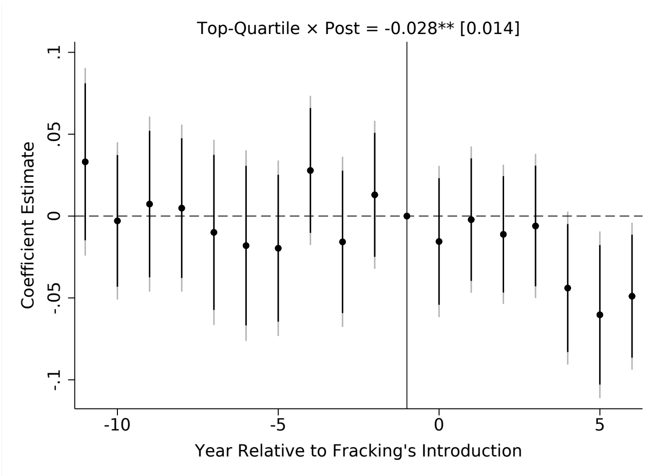
Panel A: Men and Women



Panel B: Men



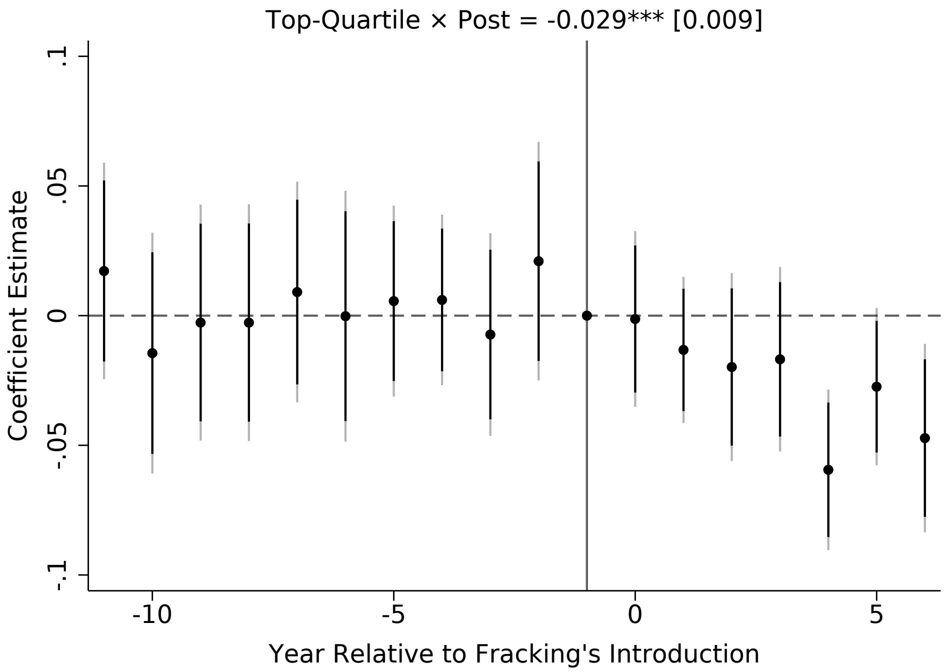
Panel C: Women



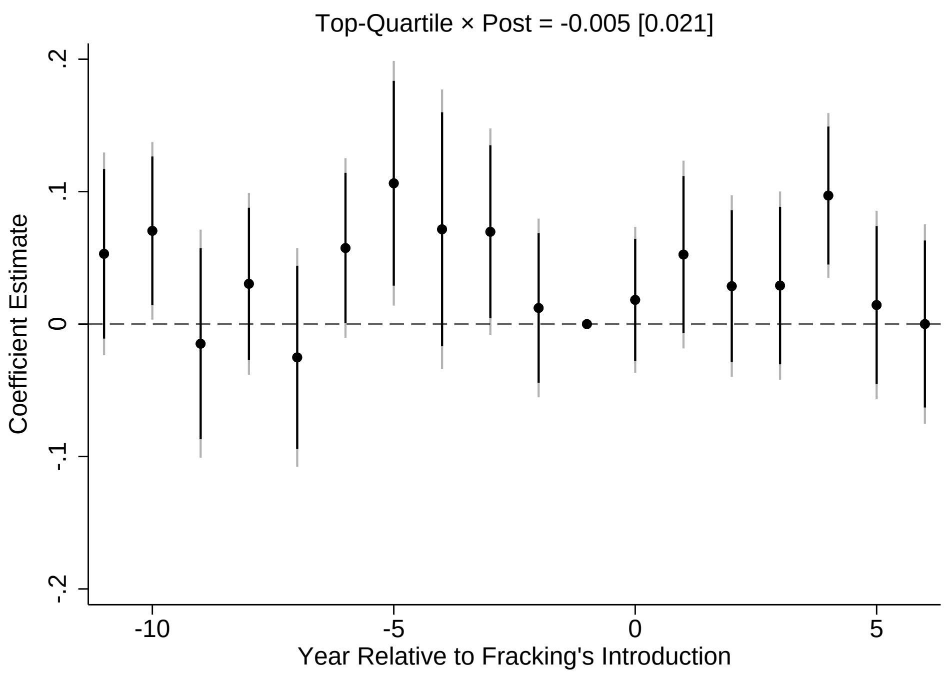
*Notes: Each panel reports the point estimates with their associated 95% and 90% confidence intervals from Equation (2) for the balanced set of event-years. The dependent variable is the inverse hyperbolic sine of overall mortality. We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. Each regression include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level. The difference-in-differences coefficient for each outcome is included above each event-study, as well as the relevant standard error in brackets. *** Significance 1%, ** 5%, * 10%.*

Figure 4: Internal vs. External Causes of Death (Ages 15-64)

Panel A: IHS of Internal Causes of Death



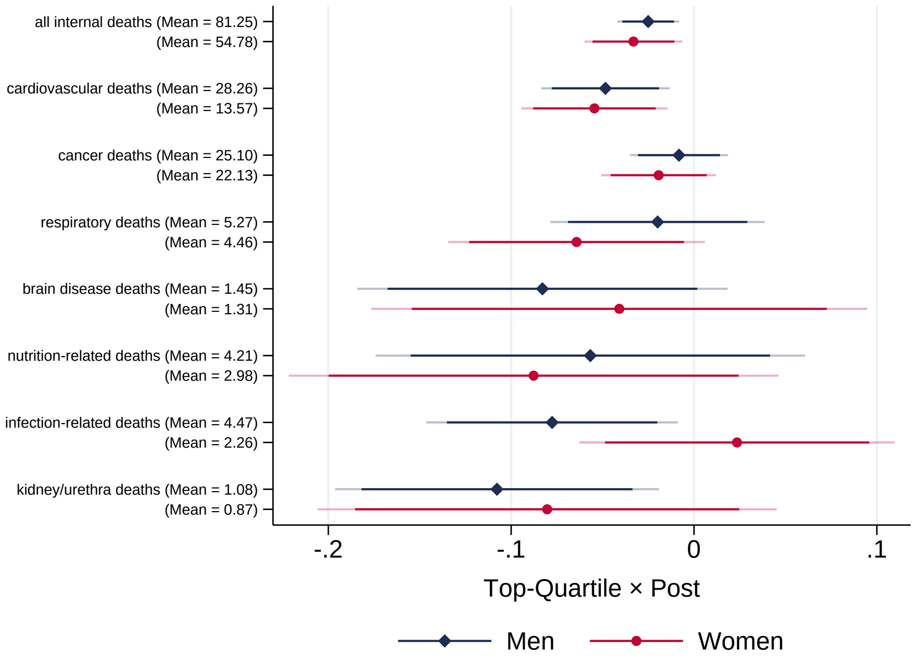
Panel B: IHS of External Causes of Death



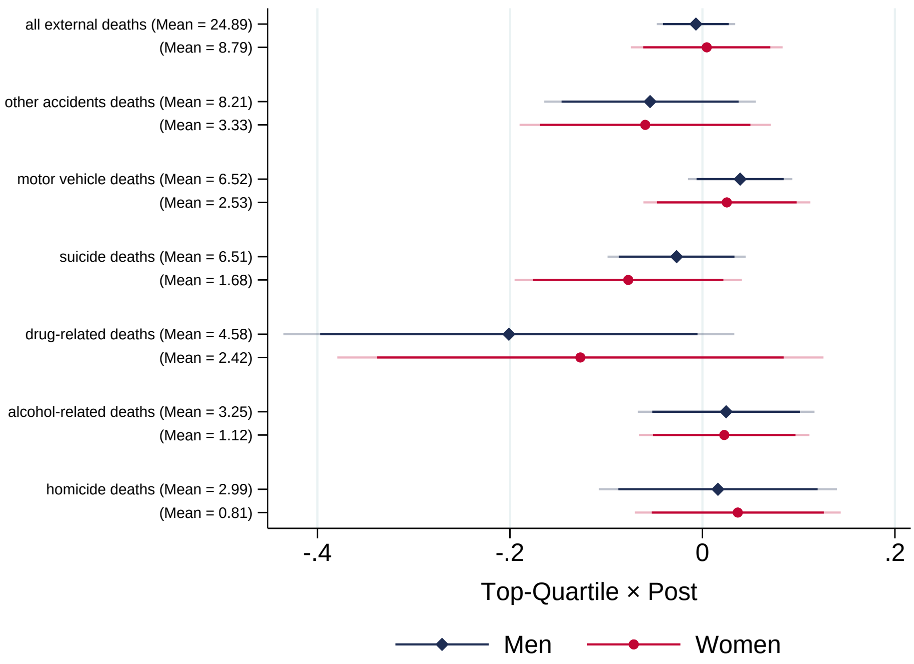
Notes: Each panel reports the point estimates with their associated 95% and 90% confidence intervals from Equation (2) for the balanced set of event-years. We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. Death categories are taken from Stevens et al. (2015), and represent consistent definitions across ICD-9 and ICD-10 cause of death codes. All regressions were estimated using interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.

Figure 5: Internal vs. External Causes of Death: Differences by Gender

Panel A: IHS of Internal Deaths (Ages 15-64)



Panel B: IHS of External Deaths (Ages 15-64)



Notes: Each point represents the outcome from a separate regression (Equation (1)), and the dark and lighter shaded bars represent the associated 90% and 95% confidence intervals, respectively. We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. Death categories are taken from Stevens et al. (2015), and represent consistent definitions across ICD-9 and ICD-10 cause of death codes. All regression were estimated using interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.

VIII Tables

Table 1: Summary Statistics - Treatment and Control Comparisons (1990 Variables)

	Top-Quartile County	Other Shale Play County	Within Play Difference
Age-Adjusted Death Rate	906.23	916.07	-2.60
	(146.75)	(124.18)	[15.17]
Median Household Income	30532.81	29815.35	-111.46
	(7878.33)	(6442.70)	[597.67]
% High School Graduates	34.90	34.83	-0.87
	(7.94)	(6.25)	[0.56]
% in Manufacturing	5.22	5.89	0.42
	(3.88)	(4.54)	[0.39]
% Married	60.76	60.16	-0.38
	(5.66)	(5.40)	[0.50]
% Rural	63.74	61.71	-4.48
	(32.11)	(28.68)	[3.01]
% Veterans	14.59	14.66	-0.03
	(2.38)	(2.14)	[0.23]
% White	91.04	90.88	-0.56
	(10.16)	(10.17)	[0.72]
% Foreign Born	2.80	2.33	-0.32
	(3.82)	(2.97)	[0.32]
% w/ a Bachelors Degree	9.57	8.77	-0.18
	(4.55)	(3.54)	[0.34]
Observations	112	407	519

Notes: All variables are measured at the county-level in 1990. Aside from the age-adjusted death rate, all variables are taken from the 1990 Decennial Census. The age-adjusted death rate is calculated using mortality data from the CDC's National Center for Health Statistics, and all the population data come from SEER. Column (3) reports the regression-adjusted difference between top-quartile counties and other counties in the same shale play (only play-specific fixed effects are controlled for), with standard errors in parenthesis.

Table 2: Working-Age Overall Mortality Rates by Gender

	Men and Women		Men		Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Top-Quartile × Post	-0.021**	-0.022**	-0.017*	-0.018**	-0.028**	-0.028**
	[0.009]	[0.009]	[0.009]	[0.009]	[0.014]	[0.014]
Omits ND & MT?	No	Yes	No	Yes	No	Yes
Outcome Mean	159.55	169.71	99.80	106.14	59.75	63.57
Observations	9,342	8,712	9,342	8,712	9,342	8,712

*Notes: *** Significance 1%, ** Significance 5%, * Significance 10%. All regressions include a time-varying control for the log of the relevant population group. All columns include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census. Some columns also omits all observations from North Dakota and Montana, and include 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.*

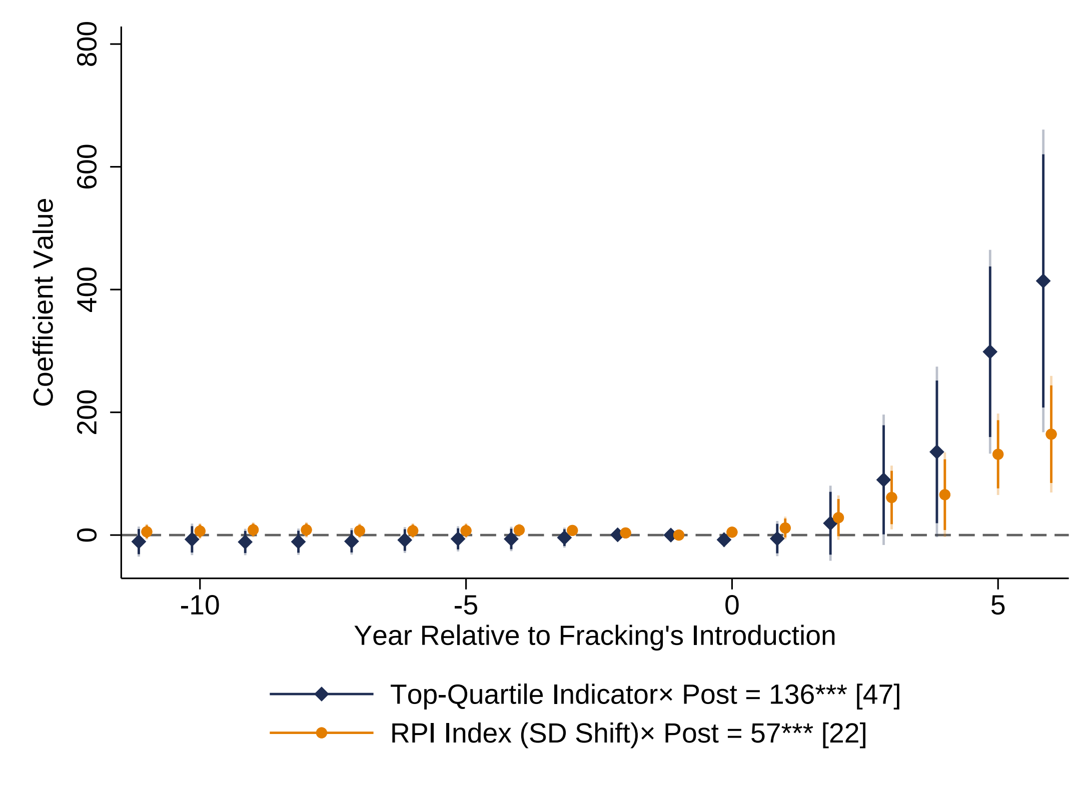
Table 3: Health Insurance Coverage by Gender: Ages 18-64

	men/women	men	women
	(1)	(2)	(3)
Top-Quartile × Post	0.012***	0.009*	0.014***
	[0.004]	[0.005]	[0.004]
1990 Controls?	Yes	Yes	Yes
Outcome Mean	0.81	0.79	0.82
Observations	5,731	5,731	5,731

*Notes: *** Significance 1%, ** Significance 5%, * Significance 10%. The sample is restricted to years after 2008 due to comparibility of the insurance estimates. We take all insurance estimates from the Small Area Health Insurance Estimates (SAHIE) Program, which are calculated using data from the 2008-2018 ACS. All columns include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census. Standard errors are adjusted for clustering at the county level.*

A Appendix Figures

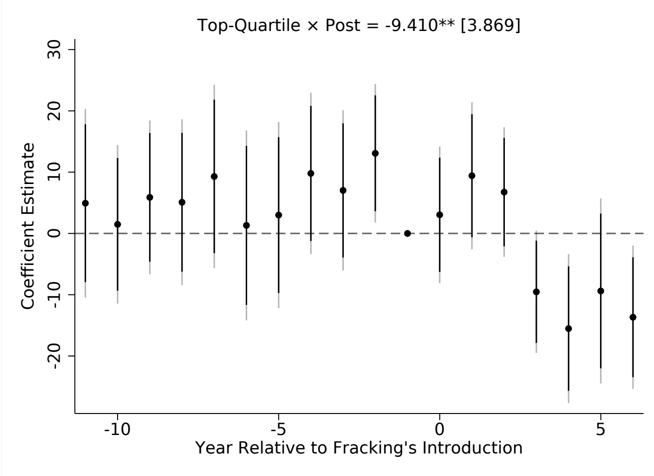
Figure A.1: Horizontal Well Production: Millions of $ of BOE



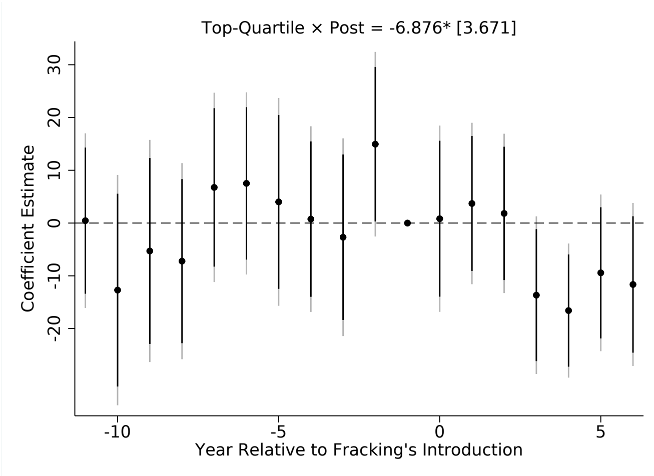
*Notes: Each panel reports the point estimates with their associated 95% and 90% confidence intervals from Equation (2) for the balanced set of event-years. Here, we show coefficients from 2 separate regressions where the coefficient of interest is a different transformation of the RPI. Monthly, well-level production of oil and natural gas data from Enverus, and we aggregate these amounts to the county-level using the latitude and longitude of each well. We use yearly price data from the EIA to calculate the value of fracking production in millions of dollars, transformed into real, 2010 $ using the PCEPI. Each regression include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level. The difference-in-differences coefficient for each outcome is included above each event-study, as well as the relevant standard error in brackets. *** Significance 1%, ** 5%, * 10%.*

Figure A.2: Age-Adjusted Overall Mortality per 100K

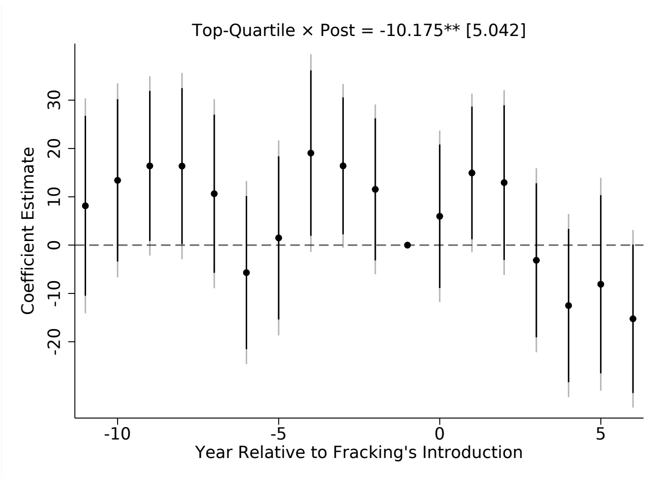
Panel A: Men and Women



Panel B: Men



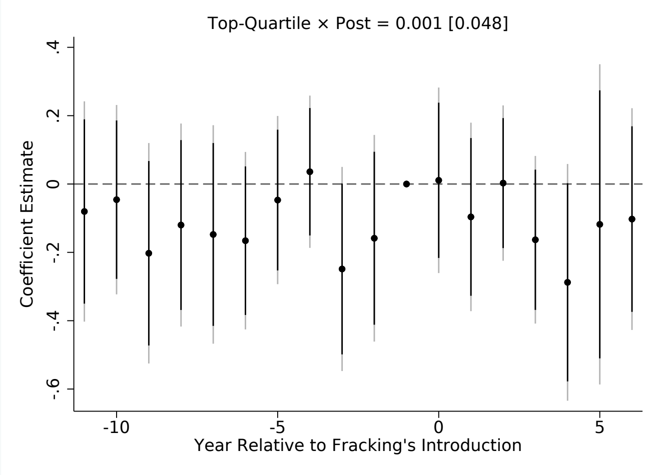
Panel C: Women



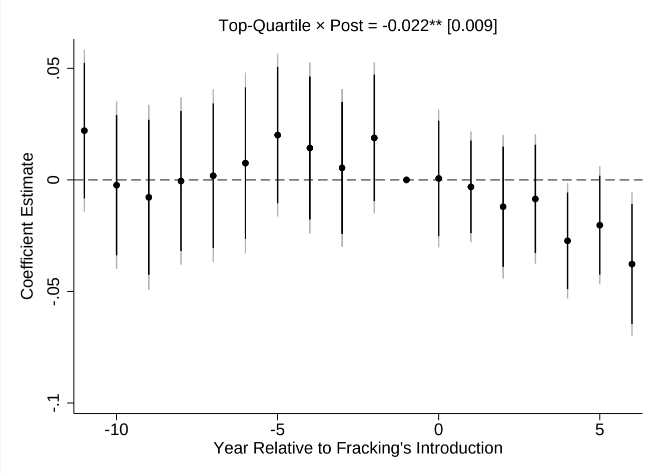
*Notes: Each panel reports the point estimates with their associated 95% and 90% confidence intervals from Equation (2) for the balanced set of event-years. We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. We use the standard method for age-adjustment by taking a weighted average of the crude death rates for different age categories within a county, where the national population shares in those age categories in 2000 are the weights. Each regression include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level. The difference-in-differences coefficient for each outcome is included above each event-study, as well as the relevant standard error in brackets. *** Significance 1%, ** 5%, * 10%.*

Figure A.3: Mortality Effects by Age

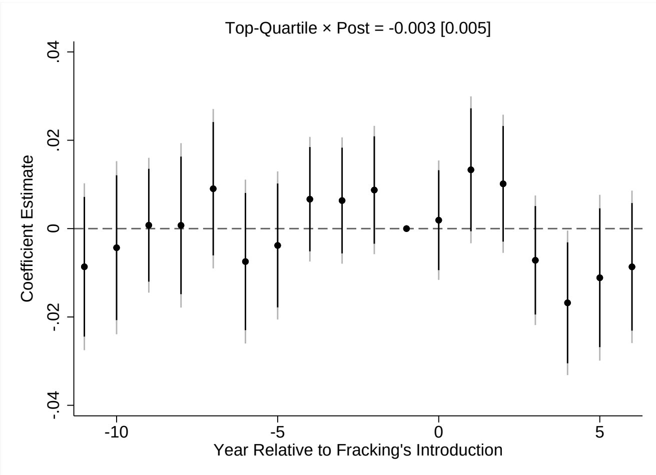
Panel A: IHS of Overall Mortality (Ages 5-14)



Panel B: IHS of Overall Mortality (Ages 15-64)

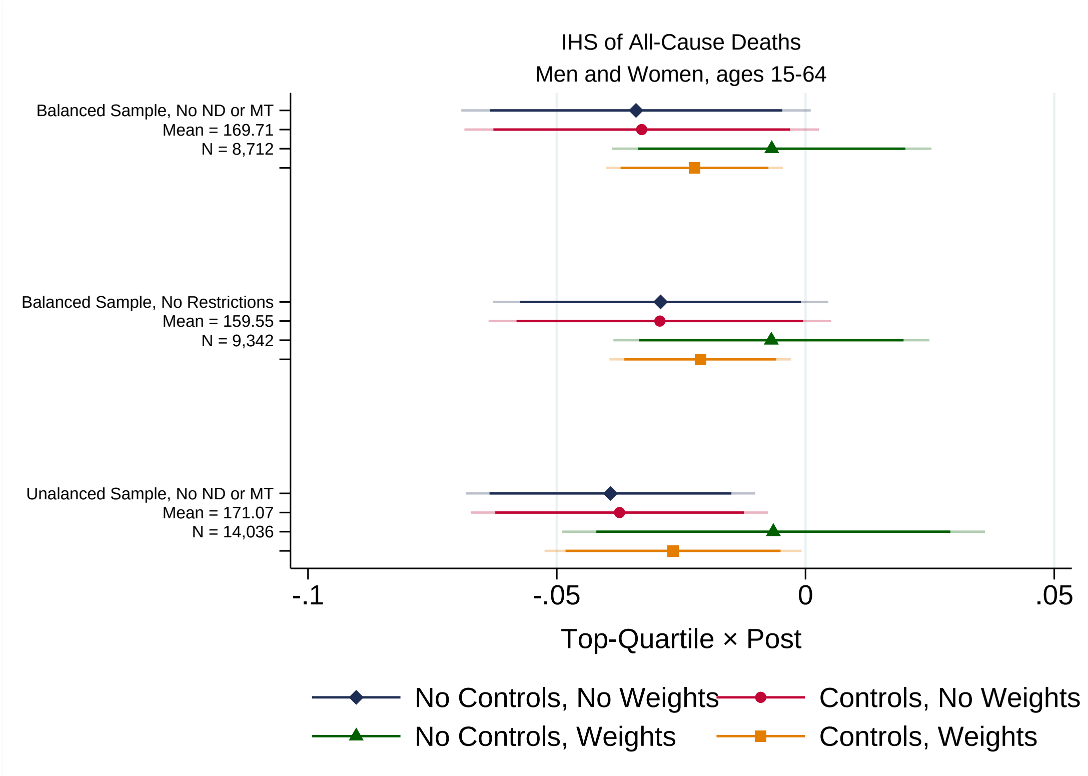


Panel C: IHS of Overall Mortality (Ages 65-99)



*Notes: Each panel reports the point estimates with their associated 95% and 90% confidence intervals from Equation (2) for the balanced set of event-years. We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. Each regression include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level. The difference-in-differences coefficient for each outcome is included above each event-study, as well as the relevant standard error in brackets. *** Significance 1%, ** 5%, * 10%.*

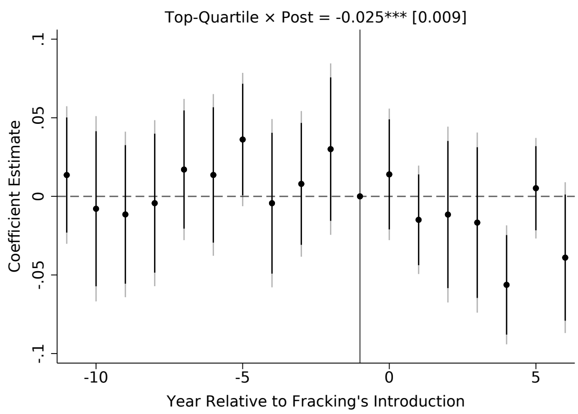
Figure A.4: Men/Women Working-Age Mortality Robustness



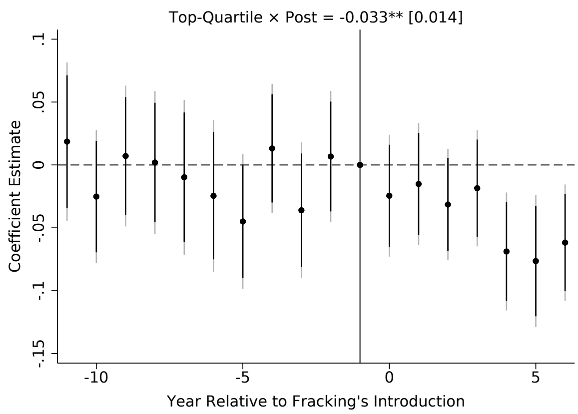
Notes: We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. Death categories are taken from Stevens et al. (2015), and represent consistent definitions across ICD-9 and ICD-10 cause of death codes. Each point represents the outcome from a separate regression (Equation (1)), and the dark and lighter shaded bars represent the associated 90% and 95% confidence intervals, respectively. All regressions include 13,746 observations, (except the specification which omits North Dakota and Montana, which has 12,371 observations).

Figure A.5: Internal and External Causes of Death by Gender

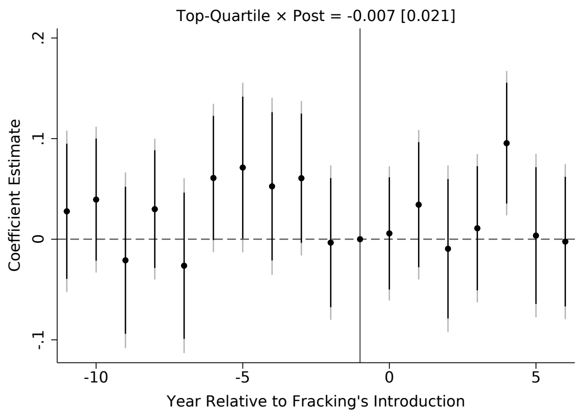
Panel A: IHS of Internal Deaths (Male)



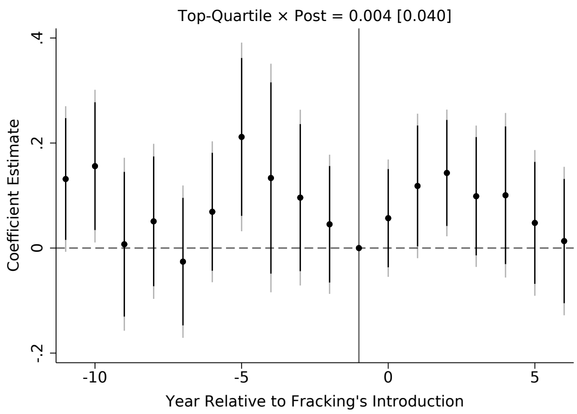
Panel B: IHS of Internal Deaths (Female)



Panel C: IHS of External Deaths (Male)



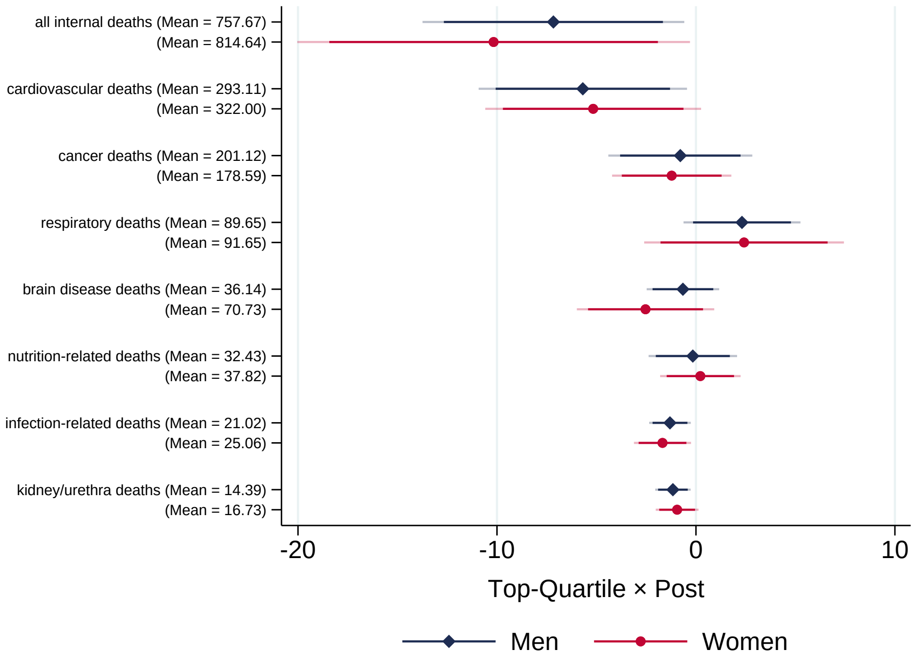
Panel D: IHS of External Deaths (Female)



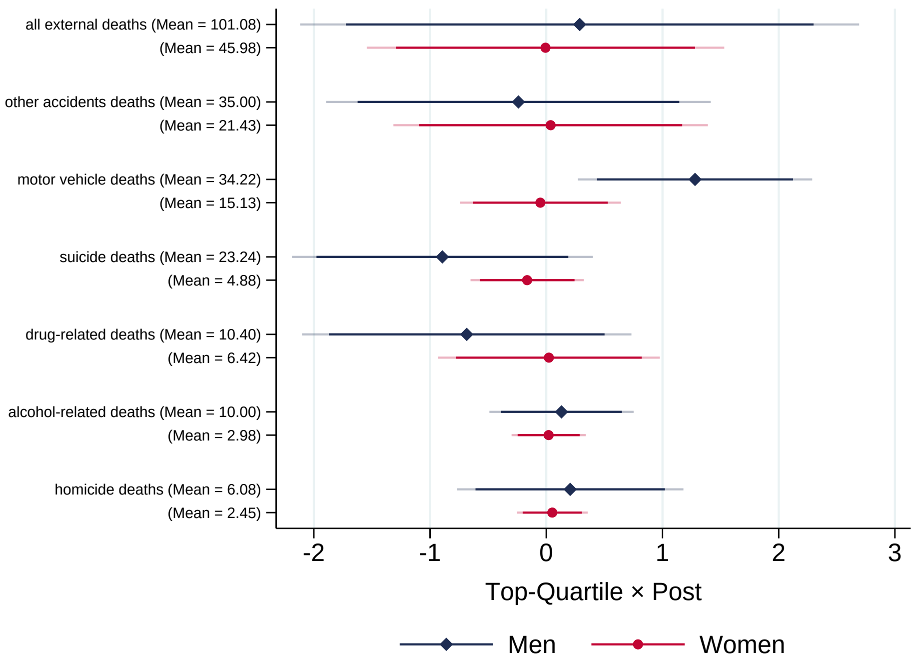
Notes: We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. Death categories are taken from Stevens et al. (2015), and represent consistent definitions across ICD-9 and ICD-10 cause of death codes. Each point represents the outcome from a separate regression (Equation (1)), and the dark and lighter shaded bars represent the associated 90% and 95% confidence intervals, respectively. All regressions were estimated using interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.

Figure A.6: Internal and External Causes of Death: Differences by Gender

Panel A: Internal Causes (Age-Adjusted Death Rate per 100K)



Panel B: External Causes (Age-Adjusted Death Rate per 100K)



Notes: We take all death count data from the CDC's National Center for Health Statistics from 1990-2018. We take population counts for the same time period from SEER. Death categories are taken from Stevens et al. (2015), and represent consistent definitions across ICD-9 and ICD-10 cause of death codes. The definitions of suicides, drug-related and alcohol-related deaths are taken from the Joint Economic Committee of the United States Congress. Each point represents the outcome from a separate regression (Equation (1)), and the dark and lighter shaded bars represent the associated 90% and 95% confidence intervals, respectively. All regressions were estimated using interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census and 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.

B Appendix Tables

Table B.1: Lower 48 States and Fracking Counties Comparison (1990 Variables)

	Any Shale Play	No Shale Play	Diff.
Age-Adjusted Death Rate	913.94	930.40	-16.46**
	(129.30)	(141.21)	[6.32]
Median Household Income	29970.18	31353.13	-1382.95***
	(6776.52)	(8703.30)	[343.13]
% High School Graduates	34.85	34.19	0.66*
	(6.64)	(6.04)	[0.31]
% in Manufacturing	5.75	8.61	-2.87***
	(4.41)	(6.48)	[0.23]
% Married	60.29	58.84	1.45***
	(5.46)	(6.61)	[0.27]
% Rural	62.15	63.67	-1.52
	(29.44)	(30.05)	[1.42]
% Veterans	14.65	14.80	-0.15
	(2.19)	(2.86)	[0.11]
% White	90.91	86.77	4.14***
	(10.15)	(16.14)	[0.55]
% Foreign Born	2.43	2.17	0.26
	(3.17)	(3.67)	[0.16]
% w/ a Bachelors Degree	8.94	9.02	-0.08
	(3.79)	(4.30)	[0.19]
Observations	519	2,589	3,108

Notes: All variables are measured at the county-level in 1990. Aside from the age-adjusted death rate, all variables are taken from the 1990 Decennial Census. The age-adjusted death rate is calculated using mortality data from the CDC's National Center for Health Statistics, and all the population data come from SEER.

Table B.2: Changes in Population Across Samples/Specifications

VARIABLES	(1) Log Male Pop	(2) Log Male Pop	(3) Log Female Pop	(4) Log Female Pop
Top-Quartile × Post	0.00973	0.00848	0.00790	0.00715
	(0.00881)	(0.00860)	(0.00944)	(0.00924)
Observations	9342	8,712	9342	8,712
Omits ND & MT?	NO	YES	NO	YES
Outcome Mean	36,666	38,955	37,798	40,169

*Notes: *** Significance 1%, ** Significance 5%, * Significance 10%. We take population counts from SEER. All values are calculated for individuals of every age in each county. All columns include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census. Some columns also omits all observations from North Dakota and Montana, and include 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.*

Table B.3: Earnings by Gender - Robustness

	Men and Women				Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Top-Quartile × Post	0.046***	0.041***	0.027***	0.024***	0.045***	0.041***	0.028***	0.023**	0.020***	0.017***	0.020***	0.021**
	[0.011]	[0.010]	[0.008]	[0.009]	[0.011]	[0.010]	[0.008]	[0.009]	[0.006]	[0.006]	[0.006]	[0.008]
No Missing Counties?	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
2000 Pop. Weights?	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Omits ND & MT?	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Outcome Mean	34,453	35,475	35,475	35,177	42,660	43,878	43,878	43,516	25,831	26,649	26,649	26,253
Observations	8,513	6,422	6,422	9,778	8,513	6,422	6,422	9,778	8,513	6,422	6,422	9,778

*Notes: *** Significance 1%, ** Significance 5%, * Significance 10%. We take earnings measures (adjusted to real 2010 dollar amounts) and employment counts from the Quarterly Workforce Indicators database. We take population counts from SEER. All values are calculated for 14-99 year old individuals in each county. All columns include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census. Some columns also omits all observations from North Dakota and Montana, and include 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.*

Table B.4: Employment to Population Ratio by Gender - Robustness

	Men and Women				Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Top-Quartile × Post	0.032***	0.028***	0.010*	0.013**	0.048***	0.042***	0.013**	0.016**	0.012**	0.010**	0.005	0.010*
	[0.008]	[0.008]	[0.005]	[0.006]	[0.011]	[0.011]	[0.007]	[0.007]	[0.005]	[0.005]	[0.005]	[0.005]
No Missing Counties?	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
2000 Pop. Weights?	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Omits ND & MT?	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Outcome Mean	0.50	0.50	0.50	0.49	0.52	0.52	0.52	0.51	0.48	0.49	0.49	0.47
Observations	8,513	6,422	6,422	9,778	8,513	6,422	6,422	9,778	8,513	6,422	6,422	9,778

*Notes: *** Significance 1%, ** Significance 5%, * Significance 10%. We take earnings measures (adjusted to real 2010 dollar amounts) and employment counts from the Quarterly Workforce Indicators database. We take population counts from SEER. All values are calculated for 14-99 year old individuals in each county. All columns include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census. Some columns also omits all observations from North Dakota and Montana, and include 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.*

Table B.5: Age-Adjusted Overall Mortality Rates by Gender

	Men and Women		Men		Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Top-Quartile × Post	-8.870**	-9.428**	-6.063	-6.893*	-9.938**	-10.19**
	(3.885)	(3.867)	(3.702)	(3.670)	(5.013)	(5.039)
Omits ND & MT?	NO	YES	NO	YES	NO	YES
Outcome Mean	853.90	861.83	851.71	858.85	851.09	860.72
Observations	9,341	8,711	9,341	8,711	9,341	8,711

*Notes: *** Significance 1%, ** Significance 5%, * Significance 10%. All death rates are age-adjusted using the national age distribution across standard age categories in 2000 to eliminate bias caused by changing demographics over time. All columns include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census. Some columns also omits all observations from North Dakota and Montana, and include 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.*

Table B.6: Overall Mortality - Heterogeneity by Age

	Less than 1		Ages 1-4		Ages 5-14		Ages 15-64		65 and Older	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Top-Quartile × Post	-0.012	-0.020	0.002	-0.085	0.005	-0.010	-0.029*	-0.022**	-0.010	-0.003
	[0.028]	[0.027]	[0.028]	[0.061]	[0.031]	[0.047]	[0.018]	[0.009]	[0.009]	[0.005]
2000 Pop. Weights?	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Omits ND & MT?	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Outcome Mean	6.78	7.19	1.27	1.35	1.71	1.81	159.55	169.71	502.82	533.76
Observations	9,342	8,712	9,342	8,712	9,342	8,712	9,342	8,712	9,342	8,712

*Notes: *** Significance 1%, ** Significance 5%, * Significance 10%. All regressions include a time-varying control for the relevant population group. All columns include interactions of a full set of year dummies (excluding 1990) with time-invariant county characteristics from the 1990 census. Some columns also omits all observations from North Dakota and Montana, and include 2000 county-level population weights. Standard errors are adjusted for clustering at the county level.*