NBER WORKING PAPER SERIES

MALE EARNINGS, MARRIAGEABLE MEN, AND NONMARITAL FERTILITY: EVIDENCE FROM THE FRACKING BOOM

Melissa S. Kearney
Riley Wilson

Working Paper 23408
http://www.nber.org/papers/w23408

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2017

We are grateful to Phil Levine, Kevin Lang, Seth Sanders, Na'ama Shenhav, and Jim Ziliak for helpful comments. We also thank our University of Maryland colleagues for many helpful comments during the applied microeconomics and population center workshops, as well as seminar participants at the University of Virginia, Boston University, and University of Kentucky. We gratefully acknowledge financial support from a University of Maryland Population Research Center seed grant. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom
Melissa S. Kearney and Riley Wilson
NBER Working Paper No. 23408
May 2017
JEL No. I3,J1,J12,J13,J18,R2

ABSTRACT

There has been a well-documented retreat from marriage among less educated individuals in the U.S. and non-marital childbearing has become the norm among young mothers and mothers with low levels of education. One hypothesis is that the declining economic position of men in these populations is at least partially responsible for these trends. That leads to the reverse hypothesis that an increase in potential earnings of less-educated men would correspondingly lead to an increase in marriage and a reduction in non-marital births. To investigate this possibility, we empirically exploit the positive economic shock associated with localized "fracking booms" throughout the U.S. in recent decades. We confirm that these localized fracking booms led to increased wages for non-college-educated men. A reduced form analysis reveals that in response to local-area fracking production, both marital and non-marital births increase and there is no evidence of an increase in marriage rates. The pattern of results is consistent with positive income effects on births, but no associated increase in marriage. We compare our findings to the family formation response to the Appalachian coal boom experience of the 1970s and 1980s, when it appears that marital births and marriage rates increased, but non-marital births did not. This contrast potentially suggests important interactions between economic forces and social context.

Melissa S. Kearney
Department of Economics
University of Maryland
3105 Tydings Hall
College Park, MD 20742
and NBER
kearney@econ.umd.edu

Riley Wilson
University of Maryland, College Park
Department of Economics
3114 Tydings Hall
College Park, MD, 20742
wilson@econ.umd.edu

I. INTRODUCTION

There is widespread interest among social scientists and policy observers in the declining rates of marriage among less-educated individuals and the concomitant rise in non-marital childbearing. In 2014, over 40 percent of all births in the U.S. were to an unmarried mother, with an even higher rate of 62 percent among non-college educated mothers.[1] A leading conjecture as to why so many less-educated women are choosing motherhood without marriage points to the weak economic prospects of their male partners. The idea is that changing labor market structures and economic conditions have adversely affected the economic prospects of less educated men, making them less "marriageable" from the perspective of the women with whom they sexually partner. This concept was first posited in the seminal work of Wilson and Neckerman (1986) and Wilson (1987) who were writing about the rise in non-marital childbearing among African American women in urban settings. But the rise in non-marital childbearing is now much more widespread, reflecting a dramatic, steady shift over the past 50 years, and it is no longer a distinctly urban or minority experience.

Though conceptually compelling and generally consistent with empirical patterns, there is scant empirical evidence about the causal link between male economic status and rates of marriage or non-marital childbearing. It is very difficult to empirically isolate the causal relationship of interest because high earning men might be more likely to marry for other reasons, and places with better male economic opportunities might also be places where more marriage- or family-inclined individuals choose to live for other reasons. A recent working paper by Autor, Dorn, and Hansen (2017) provides a notable exception. Exploiting trade shocks during the period 1990 to 2010, the authors find that a decline in male employment opportunities driven

[1] For a recent review of the relevant literature and a conceptual framework for thinking about non-marital childbearing from the perspective of children's outcomes, see Kearney and Levine (2016).

by import shocks leads to a decline in births, a decline in marriage, a rise in births to teen mothers, and an increase in the number of children being raised in single-mother families.[2]

We attempt to shed additional light on the causal relationship between *improved* labor market opportunities for less educated men and subsequent fertility and marriage outcomes using the "fracking boom" as a source of improved wage prospects for less educated men. We investigate a reverse marriageable men hypothesis, asking if an improvement in the economic position of men associated with local fracking production leads to a positive effect on marriage and a corresponding decrease in non-marital birth rates. The notion underlying this empirical approach is that fracking production occurred in some areas and not others due to pre-existing geological factors that but for the fracking boom, would not otherwise have had a determinant effect on subsequent marriage and birth trends.

Our empirical analysis examines the relationship between annual local-area fracking production between 1997 and 2012 and family formation outcomes among 18 to 34 year olds. Our baseline models define a local area as a Census Public Use Microdata Area (PUMA) and econometrically control for fixed PUMA effects, state by year effects, and a set of time-varying PUMA level characteristics. We use a continuous measure of simulated fracking production from new wells (described in detail below) at the PUMA/year level as our measure of the extent

[2] Earlier work offered some corroborating support for the notion. For example, work by Charles and Luoh (2010) finds that increased rates of male incarceration due to policy shifts led to decreased rates of marriage in affected marriage markets. McLanahan and Watson (2011) document that for a given earnings level, men who are lower ranked among their peers are less likely to be married, perhaps because they view themselves as less "marriageable". Kearney and Levine (2014) find that conditional on becoming pregnant, low-socioeconomic status young women are less likely to marry before having the baby if they live in a place with a greater level of income inequality, consistent with a story of them being less likely to find their (presumably low-SES) male partners to be desirable marriage partners. Cherlin et. al (2016) finds that part of that documented relationship appears to be due to fewer middle-skill male jobs in more unequal places. The ethnographic work of Edin and Kafalas (2011) suggests that among their sample of interviewed single mothers, an important factor in their decision not to marry the child's father relates to a perceived lack of economic security that he would bring to the family.

of the localized boom. In the years we study, 611 PUMA in the United States – out of a total of 2,057–had positive fracking production from new wells.

A key element to the strategy of using this context to advance our understanding about the link between improved male economic opportunity and family formation is that the localized fracking boom had a first order effect on the earning potential of less-educated males. We begin by documenting this relationship. We then conduct a reduced form analysis documenting the relationship between PUMA/year level simulated new fracking production and birth and marriage outcomes. This analysis does not indicate a shift toward marriage in response to an increase in the potential wages of less-educated men associated with localized fracking booms. But, both marital and non-marital births increase significantly. This is consistent with the notion that children are "normal goods," offering further support for Becker's (1960) observation and confirmation of what previous empirical papers have found (e.g., Black et al, 2013; Dettling and Kearney, 2014; Lindo, 2010; Lovenheim and Mumford, 2013). We confirm that the results do not appear to be driven by two potential confounding factors: the sex composition (male/female ratio) and house prices.

For the sake of comparison, we revisit the family formation response to the Appalachian coal boom and bust of the 1970s and 1980s, building on the work of Black, Kolesnikova, Sanders, and Taylor (2013). The data indicate that the increased earnings associated with the coal boom during those earlier decades led to an increase in marriage rates, an increase in the marital birth rate, and a decrease in the non-marital birth rate. This contrast in findings between periods might suggest that as non-marital births have become increasingly common, individuals are more likely to respond to increased income with increased fertility, whether or not they are married, and not necessarily with an increased likelihood of marriage. We conclude by speculating that social norms play an important role in determining the response of family formation outcomes to economic conditions.

II. BACKGROUND ON FRACKING

The exogenous economic shock of fracking production underlying our empirical approach arises from the technological advancements over time in the mining of shale gas and tight oil, combined with predetermined geological differences across place in fracking potential. For thousands of years, shale plays have trapped deposits of natural gas and oil far below the surface of the earth. It wasn't until the mid-2000's that the technology behind hydraulic fracturing and horizontal drilling made extraction of oil and natural gas from these shale plays economically profitable, resulting in an explosion of growth and well drilling starting around between 2004 and 2007. The geographic variation in fracking potential is dependent on the predetermined location of shale plays and fuel deposits that took shape thousands of years ago. As these deposits had no economic value and were relatively unknown for many years, communities and economies grew and developed largely independent of shale play location.

Hydraulic fracturing (also fracing, fracking, hydrofracturing or hydrofracking) is a well-stimulation technique in which the rock is fractured by a pressurized liquid. The process involves the high-pressure injection of 'fracking fluid' (primarily water, containing sand or other thickening agents) into a wellbore to create cracks in the deep-rock formations that will release natural gas, petroleum or brine. This technology has been in use since the 1950s, but experimentation with the fracking fluid formulation in the late 1990s and early 2000s led to cheaper, more cost effective applications that were capable of splitting shale rock and releasing the oil and gas reserves (Gold, 2014). Because shale plays trap oil and gas at a molecular level (rather than in a pool or pocket) conventional vertical drilling to extract resources from shale plays was not economical. Horizontal drilling was first introduced in the 1980s, allowing wells to be drilled at an angle following layers of fuel deposits, rather than vertically pass through the deposits.

Together these two technologies made oil and gas extraction from shale plays both feasible and economical, resulting in a flurry of drilling and production, and the widely publicized (in both good and bad terms) "fracking boom". In the popular press this boom has been touted as creating tens of thousands of jobs and providing starting salaries at $50,000 for recent high school graduates, with average earnings in oil and gas between $70,000 and $80,000.[3]

Recent papers on the economic impacts of fracking

There are a number of recent papers and working papers investigating the impact of fracking on wages and employment. Freyer, Mansur, and Sacerdote (2017) show that wage gains not only accrue in oil and gas extraction industries and the counties where wells are located, but in other industries and neighboring counties as well. Their estimates imply that every million dollars of new oil and gas extracted produces $80,000 in wage income, $132,000 in royalty payments and business income, and 0.85 jobs within the county in the year production occurs. They find that total economic impacts in the region are three times larger than the county-specific estimates. They also document that the impacts of new production on wages are persistent, with 2/3 of the wage income increase persisting two years after the initial shock. Working papers by

Allcot and Keniston (2014), Eliason & Timmins (2014), Fetzer (2014), and Maniloff and Mastromanaco (2014) also reveal large wage gains associated with fracking, on the order of 5 to

[3] See for example, https://www.minneapolisfed.org/publications/fedgazette/desperately-seeking-workers-in-the-oil-patch, http://www.newyorker.com/magazine/2011/04/25/kuwait-on-the-prairie, or http://www.nytimes.com/2012/12/26/us/26montana.html?rref=collection%2Ftimestopic%2FGas%20(Fuel).

24 percent. A working paper by Cascio and Narayan (2015) suggests that by increasing the wages of less educated men, fracking has led to an increased propensity among high school age males to drop out of school.

Wilson (2016) examines the migration response to the higher wages brought about by the fracking boom. His paper finds a substantial wage migration response in North Dakota, but much smaller migration elasticities in other states involved in fracking in the West, South Central, and Rust Belt. Migration increased the population of fracking counties in North Dakota by 12-25 percent, but by less than one percent in fracking counties in other states. His paper explores potential explanations for the difference. Given the large migration response to the Bakken region, we exclude North Dakota and Montana from our analysis and focus on states that where the changes in the population composition was minimal.

There are also a few papers looking at outcomes other than labor market related outcomes. Muehlenbachs, Spiller, & Timmins (2015) estimate hedonic models of property value impacts of shale gas development in Pennsylvania and New York. They find that there are negative impacts on the property valuation of groundwater-dependent homes close to wells, but small positive impacts for piped-water dependent homes, which they interpret as consistent with benefits from lease payments. Using data from a fracking county in Pennsylvania, Gopalakrishnan & Klaiber (2014) also document heterogeneous impacts of shale gas exploration activity on property values, with a modest reduction in property values for houses within one mile of a shale well. A recent working paper by Bartik, Currie, Greenstone, and Knittel (2017) estimates that households' willingness to pay for allowing fracking ranges from $1,300 to $1,900 per household annually, reflecting a positive net valuation of improved economic conditions over amenity losses (e.g., increased crime and noise.) The current paper fits into this new, innovative line of research examining the economic and social consequences of the fracking boom.

III. EMPIRICAL APPROACH

We exploit cross-sectional, time-series variation in PUMA-level fracking production to estimate a causal relationship between local economic shocks and subsequent family formation outcomes among less-educated individuals. As described above, prior to the technological innovations of the early 2000s, the oil and gas deposits extracted from shale plays through fracking were previously unattainable and had no existing economic value. The sudden shock to local economies when fracking came to their area led to increased labor demand, putting upward pressure on wages, including the wages of less educated individuals. We thus use the fracking boom of the late 2000's as an exogenous positive shock to the potential wages of less-educated men living in counties that cover a shale play. This gives us an opportunity to estimate a causal relationship between a local economic shock and family formation decisions, and under certain assumptions, interpret this as a response to an increase in male earnings.

Following Feyrer et al. (2017) we capture the extent of the fracking boom in county *c* by assigning to that county a share of shale play *j*'s production, based on the existence of geographic overlap and information about actual oil and gas production measured in dollar terms. (We subsequently map counties to PUMAs.) The data on well location and production was provided from a private company DrillingInfo through a special use agreement, as described in Section IV below. Because actual production might be correlated with unobservable characteristics related to economic and demographic conditions, we simulate production using only geographic variation in county exposure to a shale play interacted with year effects. The year interaction serves to adjust production amounts for time-varying changes in relevant prices and technology. We estimate equation [1] for all counties over shale plays, and then take the exponential of the predicted value.

$$\ln\left(new\ production_{cy} + 1\right) = \alpha_c + \sum_{\tau=y}\sum_{j=1}^{J} \hat{\theta}_{\tau j} I\{county\ c\ over\ shale\ play\ j\} * I\{year = \tau\} + \nu_{cy} \quad [1]$$

$$sim.\,new\ production_{cy} = \exp\left(\hat{\alpha}_c + \sum_{\tau=y}\sum_{j=1}^{J} \hat{\theta}_{\tau j} I\{county\ c\ in\ shale\ play\ j\} * I\{year = \tau\}\right) - 1$$

The main explanatory variables are a set of interactions between an indicator that equals one if county *c* intersects shale play *j* (measured using ArcGIS software, as described below) and an indicator that equals one in year *y*. This estimates the average impact of being in shale play *j* on new production and allows this relationship to vary over time, as technology and prices change. The variable $new\ production_{cy}$ is the dollar value of oil and gas production from wells drilled in the current year located in county *c*. This simulated measure is then aggregated up from the county/year level to the PUMA/year level. The observed correlation between actual new production and simulated new production is p=.69, suggesting that geological constraints and temporal variation play an important role in determining production intensity. To better capture the intensity of treatment on population outcomes, we divide our measure of simulated new production by the baseline population in 2000, and scale the measure to represent simulated new production per capita, in thousands of dollars.

We conduct our analysis at the PUMA-level rather than the county-level for several reasons. First, public-use Census data on annual level marriage outcomes are not available at areas smaller than the PUMA. Second, PUMA are constructed to encompass geographic areas that will maintain a total population over 100,000 between 2000 and 2010. By conducting our analysis at the PUMA-level, we thus avoid comparing sparsely populated counties to more populous counties. We also avoid the issue of having observation cells with zero or close to zero marital or non-marital births in any given year. Finally, previous work has suggested that the labor market impacts of fracking propagate beyond county borders, and this is especially true among less populated counties. PUMA-level analysis captures these cross-county spillovers.

To begin our analysis, we first estimate "first stage" effects of simulated new production on the wages of men and women, separately by gender and education. Our baseline specification is not a two-stage least squares model (2SLS) because of concerns about the exclusion restriction (as we describe below), so this is not formally a first-stage model in a 2SLS estimation. The model is described by the following equation:

$$wages_{py} = \alpha_0 + \alpha_1\left(simulated\ new\ production_{py}\right) + X'_{py}\xi + \mu_p + \phi_{sy} + \eta_{py} \qquad [2]$$

The subscript p refers to PUMA and y refers to calendar year. The matrix X_{py} is a vector of time-varying PUMA-level controls which includes the ratio of 18 to 34 year old females to males, the log average house price, gender specific shares of 18 to 34 year olds Non-Hispanic black, Hispanic, and Non-Hispanic other, and gender specific shares of 18 to 34 year olds with less than high school, some college, or a four year college degree. Age specific sex ratios and race by gender shares are calculated from the Surveillance, Epidemiology, and End Result Program (SEER) population data, which are derived from the U.S. Census population estimates. The housing price data is obtained from the Federal Housing Finance Agency three digit zip code housing price index, which we then link to counties and PUMA and convert to dollars using the median house value from the 2000 Census. We aggregate up the county level estimates from the 2000 Census and the ACS to calculate PUMA-level age specific gender by education shares.

The model also includes controls for time-invariant PUMA effects μ_p and state-specific year effects,ϕ_{sy}, to account for fixed differences across PUMA and time trends or shocks in wage outcomes experienced at the state level.[4] As described below, the data show a strong positive relationship between simulated fracking production and the wages of men, and especially non-college educated men, in the PUMA. We are not literally estimating this as a first stage relationship because our main analysis is a reduced form analysis. This is an important motivating equation, however, showing that the fracking boom had a first order effect on male wages.

Figure 1 presents a map showing where shale plays are located and plots the total production value from new wells per capita between 2000 and 2012, in bins, by county. Production values were uniformly high in North Dakota. Of the 16 counties in North Dakota with any new production, seven produced over one hundred thousand dollars per capita from new wells between 2000 and 2012, with an average of $244 thousand per capita over the entire period. This massive amount of production is due to the large oil reserves in the Bakken shale play. The other locations with the highest levels of simulated production – over twenty five thousand dollars per capita between 2000 and 2012 – include counties in Arkansas, Colorado, Louisiana, Montana, New Mexico, Oklahoma, Pennsylvania, Texas, Utah, and Wyoming. As noted above, we exclude North Dakota and Montana from our analysis. With that exclusion in place, there are 88 PUMA with total simulated new production above $25 thousand per capita, and an additional 516 with positive total simulated new production below $25 thousand per capita.

[4] We do not include PUMA-specific trends in the model because these will (over)control for the response to the shock, an econometric point discussed by Wolfers (2006). When we do estimate the models with PUMA-specific trends included, there is still a statistically significant increase in marital birth rates, albeit the point estimate is much smaller, and there is no longer a discernible increase in the non-marital birth rate. There is still no discernible change in marriage related outcomes.

We then estimate the reduced-form relationship between simulated fracking production from new wells and subsequent birth and marriage outcomes:

$$Y_{py} = \beta_0 + \beta_1\left(simulated\ newproduction_{py}\right) + X'_{py}\omega + \mu_p + \phi_{sy} + \varepsilon_{lpy} \quad [3]$$

The key outcome variables of interest (Y_{py}) are defined at the level of a PUMA p and year y. We separately examine the non-marital birth share, total births, married births, and non-marital births. The year is the year of conception. We then consider a set of marriage related outcomes. *Simulated new production$_{py}$* is measured in thousands of dollars per capita. The matrix X_{py} is the same vector of time-varying PUMA-level controls included in equation [2]. Again, the model also includes controls for time-invariant PUMA effects μ_p and state-specific year effects ϕ_{sy}, to account for fixed differences across PUMA and time trends or shocks in birth and marriage outcomes experienced at the state level. Estimation of equation [3] yields estimates of the reduced-form impact of simulated fracking production on birth and marriage outcomes.

We are ultimately interested in identifying the effect of potential male earnings – an indicator of the "marriageability" of men, in the conceptual framework of Wilson (1987) – on family formation outcomes. Ordinary Least Squares (OLS) estimation of the relationship between family formation outcomes and male earnings would likely yield a biased estimate of the causal parameter of interest as there are almost certainly unobserved changes across time within a county that affect both male wages and family formation outcomes. That is why we instead attempt to gain insight from the reduced form effect of a localized economic shock that increased the potential wages of men without a college degree.

For this reduced form approach to be informative about the relationship between male economic prospects and family formation outcomes, it must be true that simulated fracking production only, or at least primarily, affects subsequent trends in birth and marriage outcomes through its effect on male economic prospects, as captured by measured wages. This condition might not be satisfied if the fracking induced by the existence of shale play affected other factors that might affect family formation, such as house prices. (Kearney and Dettling (2014) and Lovenheim and Mumford (2013) have documented a relationship between house prices and birth rates.) For this reason, our model includes house prices directly as a control variable. If the estimated reduced form relationship changes with the inclusion of this variable, it would indicate that such variables are potentially driving part of the observed relationship between the local fracking boom and subsequent family formation outcomes. As it turns out, the data do not indicate that to be the case.

Under strong assumptions, we can use simulated fracking production as an instrument for male wages, and then relate predicted wages to birth and marriage outcomes. Two potential confounders that we worry about are an increase in female wages and an increase in unearned income from land leases and royalties from mineral rights. Nonetheless, for the sake of comparison to existing studies, in particular Black et al. (2013), and to gauge what the implied responsiveness would be if all of the effect on family formation came through the channel of male earnings, we report the results of such an IV analysis below.

IV. DATA

This analysis requires detailed information on local level fracking production, wages, and birth and marriage outcomes. In this section we provide an overview of the various data sources we draw upon for our analysis. More detailed information is available in the Data Appendix.

Data on fracking production

To construct PUMA level measures of fracking production, we overlay shale play boundary shapefiles from the Energy Information Administration (EIA) onto U.S. Census Bureau county boundary shapefiles. ArcGIS software is used to if counties and shale play intersect.[5] These geographic measures are then combined with well level quarterly oil and gas production data obtained through a restricted access agreement with DrillingInfo, a private firm that collects lease, permit, and production data on all wells drilled in the United States.[6] Relevant to our data needs, the DrillingInfo data file indicates drill date, quarterly production amount, reservoir name, drilling direction (vertical or non-vertical), and latitude and longitude, which allows us to identify the county location of the well. Oil and gas production is reported in barrels and thousands of cubic feet respectively. We use average annual national prices for oil and gas, recorded by the Energy Information Administration (EIA), to convert production amounts into dollar amounts. We then use the Personal Consumption Expenditure (PCE) price index calculated by the United States Bureau of Economic Analysis to adjust all dollar amounts to year 2010 dollars. Actual production values as well as simulated production described above are then aggregated up to the PUMA-level using county to PUMA mappings from the U.S. Census. In large urban counties that contain multiple PUMA, production values are assigned according to the population share.

Data on wages

We use the Quarterly Workforce Indicators (QWI) as our primary source of wage data. The QWI is aggregated from the Longitudinal Employer-Household Dynamics (LEHD) micro-level data collected from unemployment insurance earnings data from participating states and

---

[5] A special thanks to University of Maryland Geography students Lisa Boland and Michael Bender for their research assistance using ArcGIS software.

[6] These proprietary data are obtained through an academic use agreement with DrillingInfo, available through their academic outreach initiative.

several other sources (U.S. Census, 2014). The QWI is aggregated to the county level, and can be tabulated by firm characteristics (industry, size) or worker characteristics (gender, age, education). When tabulating by worker characteristics, only two levels of tabulation are feasible: gender by age or gender by education. Because the QWI is constructed from firm employment, all measures are constructed for jobs, not individual people. Ideally we would have a measure of *potential* wages to test the hypothesis that an improvement in male economic prospects positively affect marriage rates and negatively affect non-marital birth rates. What we observe in the data – e.g., job-level quarterly wages for jobs held by men or women with various levels of education – is an imperfect proxy for potential wages of a particular group.[7] As with the production data, the wage data is aggregated up to the PUMA-level by summing total wage earnings for a given group across all counties in the PUMA and dividing by the total number of jobs for that same group.

Data on birth outcomes

We use restricted access Vital Statistics data obtained from the National Center of Health Statistics to construct PUMA level measures of birth outcomes. These files provide the universe of births between 1997 and 2013 with county identifiers, which we then aggregate to the PUMA-level. We date births back to the time of conception by subtracting the length of gestation from the fifteenth of the month of birth. We do this to consider the time the pregnancy decision was made, or more accurately said, the time the decision was made to contracept or not. Our analysis sample thus consists of live births that were conceived between 1997 and 2012. For most of our

[7] It is conceptually appealing to use a measure of potential wages to test the hypothesis that male economic prospects are a key driver of marriage and birth decisions. An alternative measure would be actual earnings, which would account for employment, hours, and wages. This data could be obtained from the American Community Survey (ACS), but public-use data from the ACS does not include information on geographic units smaller than a Public Use Micro Area (PUMA), which rules out a county-level analysis. For the sake of exploration, we have experimented with a commuting zone level analysis using ACS data on actual earnings; the results are generally in line with what we find in our county level analyses, but less precise.

analysis we focus on births to women, ages 18 to 34. During our period of analysis over 71 percent of these births were to single women, consistent with the general finding that births to young women are births to less-educated women. To construct age specific birth rates, we divide the number of births by the number of 18-34 year old women in each county from the SEER gender by age population estimates. We then multiply this by 1000 to be interpreted as birth per 1000 women.

Given our interest in the labor market opportunities and family formation decisions of low-income, less-educated men and women, we would ideally examine births to non-college women specifically. Our baseline analysis focuses on total births to 18 to 34 year old women largely for data reasons. The non-reporting of maternal education in some state/year cells is a well-known issue with using the natality files for group-level analyses. The NCHS began requiring states to report maternal education according to a 2003 classification starting in 2009. For the 20 states that did not comply with this requirement, maternal education is not recorded in 2009. Starting in 2010, some of these states' education measures were included again. To limit our sample to states with consistent reporting of maternal education severely limits the sample. We thus do not limit the sample by education level in our main sets of analysis, but in a subsequent analysis we examine birth and marriage outcomes for women separately by education group for the restricted sample.

PUMA level data on marriage outcomes

We use the 2000 Decennial Census and 2005-2011 American Community Survey (ACS) microdata to construct the PUMA-level share of 18 to 34 year old women who are never married, married, divorced, and cohabitating (Ruggles, 2015). Recently, many partners that do not marry still cohabitate, which is imperfectly captured in the Census and ACS. To measure cohabitation we expand the method use previously and count women 18-34 as cohabitating if they are either the head of the house and an unmarried partner is present, or if they are listed as an unmarried partern (Shenav, 2016). The 2005-2011 ACS also includes an indicator that equals one if the individual was married in the last year, which allows us to measure PUMA-level marriage rates

Analysis sample construction

Our analysis is estimated at the level of PUMA by year. There are 2,057 total PUMA in the lower 48 states (as compared to 3,109 counties). Two sample restrictions reduce our sample to 2,044 PUMA. First, given the unique context of fracking in the Bakken region, in particular the migration response documented in Wilson (2016), we exclude the 12 PUMA in North Dakota and Montana. Second, we exclude the PUMA with the highest level of simulated new production, which corresponds to Webb County, TX. Simulated production in this PUMA was over 125 percent larger than the second highest producing PUMA, and it is excluded to limit the influence of outliers.

The final sample consists of a balanced panel of 2,044 PUMA observed over the 16 years from 1997 to 2012.[8] PUMA in our analysis sample with any simulated new production are located in the following 26 states: Alabama, Arkansas, California, Colorado, Georgia, Illinois, Indiana, Kansas, Kentucky, Maryland, Michigan, Mississippi, Missouri, Nebraska, New Jersey, New Mexico, New York, Ohio, Oklahoma, Pennsylvania, Tennessee, Texas, Utah, Virginia, West Virginia, and Wyoming. Eight of these states (Alabama, Georgia, Illinois, Indiana, Maryland, Missouri, NJ, and Tennessee) do not have any actual fracking production, but wind up with predicted fracking production because they have land that overlaps with a shale play. For

[8] When looking at marriage outcomes from the Census and ACS, we use PUMA-level measures created from the 2000 Census and 2005-2011 ACS.

the PUMA with positive simulated fracking that are in states with no actual production (, simulated production is less than $30 per capita. Recall that our simulation method is intended to create a measure that is a function of geological attributes, not choice variables like local ordinances.

Summary Statistics

In Table 1 we present mean values from the year 2000 (before the fracking boom) for our birth and marriage outcomes, labor market characteristics, and several population characteristics for both non-fracking and fracking PUMA, where a fracking PUMA is defined as a PUMA with positive simulated production at any point between 2000 and 2012. The statistics reported in this table indicate that the fracking counties in our sample are not substantively different from other counties in the country along the dimensions on which this paper is focused. In particular, pre-boom birth rates and non-marital birth shares are quite similar across county groups. The non-marital birth share in 2000 for women age 18 to 34 was 34.2 percent in non-fracking PUMA and 33.4 percent in fracking PUMA. Marriage outcomes are also similar across non-fracking and fracking PUMA at baseline, although women age 18 to 34 in fracking PUMA are slightly more likely to have been married at some point in time (1.1 percent less likely to be never married, one percent more likely to be married, 0.3 percent more likely to be divorced). The share of women age 18 to 34 married was 37.6 percent in non-fracking PUMA and 38.6 percent in fracking PUMA. Overall, labor markets and population characteristics were similar between non-fracking and fracking PUMA, though fracking PUMA had a lower share of college educated individuals (18.0 versus 15.5 among men and 21.5 versus 18.4 among women, with both differences being statistically significant.)

The number of PUMA with active wells increased dramatically over the sample period. The count of PUMA in our analysis sample with active fracking wells in a given year rises from 331 in 2004 to 581 in 2012. Summary statistics about fracking production are reported in Appendix Table A1. Fracking production is highly skewed across fracking PUMA. For instance, in the year 2012 (the end of our sample period), annual simulated new production for the median fracking PUMA is three dollars per capita; but the PUMA at the 90th percentile of the production distribution has simulated production of 71 dollars per capita. Annual simulated new production among the top 10 percent of producing PUMA averaged between 500 and 600 dollars per capita.

IV. RESULTS

A. Effect of localized fracking boom on labor market outcomes

We begin our empirical analysis by verifying that local fracking activity led to an increase in potential earnings for non-college educated men. Table 2 reports the results of the estimation of OLS equation [2], specifying log average earnings in a PUMA/year cell (from the QWI files) as a function of PUMA/year simulated fracking production. There are 29,471 PUMA/year cell observations in this regression, coming from 45 states.[9] The coefficient of interest in column (1) indicates that an additional thousand dollars of simulated new production per capita is associated with a statistically significant 3.8 percent increase in average earnings for men. If we separately look at earnings by educational attainment we see that this increase is concentrated among non-college men, where an additional thousand dollars of simulated new production per capita is associated with a statistically significant 4.4 percent increase in average earnings.

[9] Alaska, Hawaii, Montana, North Dakota, South Dakota and Massachusetts are excluded. During this period earnings data from South Dakota and Massachusetts was not available through the QWI.

The increase in earnings is not limited to oil and gas extraction jobs. Column 2 reports that one thousand dollars per capita of fracking production is associated with a 2.6 percent increase in average earnings for men in jobs outside those industries, with impacts slightly larger for non-college men (2.9 percent). This is consistent with positive spillover effects on other earning opportunities. That is, the fracking boom leading to a localized economic boom that extends beyond the new fracking jobs created. This finding is consistent with the findings of Feyrer et. al (2017). In column (3) we see that fracking production is also associated with an increase in the jobs to population ratio. Specifically, one thousand dollars per capita of fracking production is associated with a 5.2 percent increase in the number of jobs (standard error 0.2). This increase in jobs was experienced by both non-college and college educated men. Columns (4), (5), and (6) report analogous results for regressions estimated for women. In earnings specifications, the data indicate similar signed effects, with roughly half the estimated magnitude as for men. However, non-college women do not observe an increase in jobs and college educated women actually observe a small decrease in average earnings. This will be important to keep in mind when we return to interpreting the results below.

B. *Effect of localized fracking boom on birth and marriage outcomes*

Having established that local fracking production has a positive effect on the economic prospects of non-college educated men, as captured by earnings and jobs, we turn to an estimation of the reduced form relationship between simulated fracking production and birth and marriage outcomes. We start by looking at the impact on the non-marital birth share, defined as the share of births that are born to an unmarried mother. A reduction in the non-marital birth share would be consistent with a reverse marriageable male hypothesis, suggesting that a localized economic shock that raises male earnings is associated with a decrease in the non-marital birth share. However, to interpret this reduction as a response along those lines, it is necessary to consider what happened to total births and to marriage rates. If the number of total births remained constant, then a reduction in the non-marital birth share would reflect a shifting from non-married births to married births, driven by an increase on the extensive margin of marriage. As shown in Table 3, this is not what the data reveal.

Table 3 column (1) indicates that an additional thousand dollars of simulated fracking production per capita leads to a statistically insignificant 0.11 percentage point decrease in the non-marital birth share. The result in column (2) indicates that a localized fracking boom leads to an increase in total births. The point estimate implies that one thousand dollar of fracking production per capita is associated with an increase of 5.96 births per 1000 women (standard error of 0.96). In the peak years of the boom, simulated production per capita in the most intensive fracking counties was between $500-$600 per capita, which would suggest that total births increased by 3-3.6 births per 1000 women, or around three percent. This is consistent with a positive income effect of income on fertility, as has been found in previous work (for example, Black et al, 2013; Dettling and Kearney, 2014; Lindo, 2010).

Results reported in Table 3 columns (3) and (4) show that both marital and non-marital births increased. The point estimates imply a greater proportional increase in marital births, but the estimated effects are not statistically different. The estimated coefficients are, respectively, 3.57 (standard error of 1.0) and 2.39 (standard error of 0.6). Though this is a reduced form result and not a direct measure of the elasticity of fertility with respect to income, we know of no previous work directly comparing marital and non-marital birth responses to the same economic shock.

In Appendix Table A2 we explore the robustness of the estimated birth effects. The response of both marital and non-marital birth rates is robust to excluding house prices and the sex ratio in the regression model (column 2). It is also robust to estimating the model unweighted (column 3), including year fixed effects rather than state by year (column 4), including shale play by year fixed effects rather than state by year (column 5), and defining the outcome as the natural log of birth rates (column 6).

We also investigate heterogeneous effects by demographic groups. The results are reported in Appendix Table A3. Column (1) considers older women, and the results do indicate that for women ages 35-44, there is an increase in the marital birth rate and a decrease in the non-marital birth rate. The point estimates in columns (2) through (5) imply that the main effects are being driven by non-Hispanic whites. There is a sizable increase in both the marital and non-marital birth rate for non-Hispanic whites, whereas for other race/ethnic groups the effects are not statistically different from zero. But, for those groups, the birth effects are imprecisely measured and racial/ethnic differences cannot be confirmed. The final two columns in the table show that for both the marital and non-marital birth rate, there is an increase in both first and higher-parity births.

The pattern of results is consistent with a positive income effect on fertility for both married and unmarried couples, but no obviously with a reverse marriageable men hypothesis. To gain more insight into this, we look directly at marriage outcomes. Table 4 reports the results from estimating equation [3] with the dependent variable defined as the percent of women 18 to 34 who are never married, married, divorced, cohabitating, or newly married (married in the last year). The data give no indication that the economic activity associated with fracking production led to a reduction in the percent never married (column 1) or an increase in the percent married (column 2) or newly married (column 3). Nor does the data give any indication that divorce fell (column 4). In fact, although the point estimates are small and not statistically different from zero, they are all of the opposite sign of what the reverse marriageable men hypothesis would predict; the direction of the point estimates suggest that the share married and newly married fell and the share never married and divorced rose. [10] Given the rising secular trend in cohabitation, we also test to see if the share of women cohabitating rises with simulated production. We find no evidence of such an effect (column 5).[11]

We next consider whether the effect of the localized economic shock on fertility and marriage outcomes varies by social context. In particular, we investigate whether the estimated effects vary with the baseline non-marital birth share. If a higher rate of non-marital births is associated with a more accepting social norm than in places where non-married births are less common, we might expect to see a larger increase in non-marital births in response to the

---

10 A positive effect on non-marriage and divorce in this context is inconsistent with a reverse marriageable men hypothesis, and also inconsistent with the predictions of a bargaining model based on relative wages. The standard model of marriage in the economics literature posits that as female wages rise relative to male wages, there will be a reduction in marriage, because the returns to marriage are lower (Becker, 1974). Furthermore, there will be an increase in divorce because the female outside option has increased (Browning et al, 1994). Shenhav (2016) provides empirical support for this prediction for the time period 1980 and 2010. Exploiting demand shifts in industry/occupation employment cells as an exogenous shock to sex-specific wages, she finds that increases in the relative wage of women led to a decline in the likelihood of marriage for those on the margin of a first marriage. But, here we see that a local economic shock that had the effect of decreasing the female/male wage ratio potentially led to an increase in non-marriage and divorce, the opposite of what the bargaining model would predict. A positive point estimate on divorce is consistent with prior evidence of the pro-cyclicality of divorce (e.g., Hellerstein and Morrill, 2011; Amato and Beattie, 2011; and Schaller, 2013).

[11] As another check on specification, we consider whether the results are sensitive to using the simulated production measure versus actual production. The results of using actual production are reported in Appendix Table A4. The table shows that the pattern of results are similar to our baseline findings: there is a significant increase in marital and non-marital birth rates with no significant impacts on marriage. But, the magnitude of the point estimates is greatly attenuated because the simulated measure is much lower than actual production; an additional one thousand dollars of simulated new production per capita is associated with a five thousand dollar per capita increase in actual production.

positive economic shock associated with fracking in places with higher baseline non-marital birth shares. As reported in Table 5, the data is consistent with this prediction.

In Table 5, we report the results of estimating equation [3] separately for counties with a high or low non-marital birth share, where we define those categories as relative to the median non-marital birth share among women age 18 to 34 in the year 2000, which is 33.6 percent of births. The results indicate a pattern that is consistent with the social norms prediction stated above. The impacts on marital and non-marital birth rates are more similar in places where non-marital births were a larger share of all births and the point estimate on non-marital birth rates is larger in "high" non-marital birth share PUMA than in "low" non-marital birth share PUMA. However, some of these relationships are imprecisely estimated and we cannot rule out that births responded similarly across "high" and "low" places. To push on this further, we also ran the analysis at the county level (not reported in the table); in that case, the data indicate that the response of non-marital births in "low" non-marital birth share counties is close to zero and significantly different than the marital birth response as well as the non-marital birth response in "high" non-marital birth share counties.[12] Though not conclusive of social norm effects, these patterns of responses are consistent with the notion that social context partially determines the family formation response to a positive income or earnings shock. We return to this notion when we explicitly compare the findings from the current fracking context to the experience of the Appalachian coal boom and bust of the 1970s and 1980s.

[12] In specifications not reported in the table, we also estimate the model with an interaction term between simulated new production and a continuous variable measuring year 2000 county level non-marital birth share. That specification yielded a statistically insignificant coefficient on the interaction term of interest. In addition, we estimated the model with an interaction of simulated new production and year 2000 measure of percent religious, as measured from the Association of Religion Data Archives county level church membership. That interaction term did not enter the model with statistical significance for any of the birth or marriage outcomes.

C. Alternative specifications

One possible explanation for the lack of a finding of a marriage effect is that any effect of the positive economic shock associated with fracking on marriage rates might occur with a lag. This could be because women need time to update their expectations about male economic status or perhaps to observe whether economic improvements are persistent. (Recall that Freyer et al. 2017 document that the wage increases associated with fracking production show persistence.) To allow for this possibility, we alternatively specify the regression analysis model as a long difference. We define the outcome variables to be the difference in the birth or marriage outcome of interest from 2000 and 2011 and estimate that as a function of total simulated production in the PUMA between 2000 and 2011. This definition of fracking production is meant to proxy for the total size of the economic shock during this extended period.

The results of this long-difference specification are reported in Table 6. As with the annual specification of the regressions, the data indicate a sizable and statistically significant increase in both the marital and non-marital birth rate. But now the marital birth effect is statistically larger than the non-marital birth effect. An additional one thousand dollars of simulated production per capita increased the marital birth rate by 0.8 births per one thousand women age 18 to 34 and the non-marital birth rate by 0.2 births per one thousand women. Counter to the marriageable men hypothesis, the long-difference results show an increase in the percent never married, though there is no discernible change in the percent married, divorced, or cohabiting. We are hesitant to draw too strong a conclusion from this one coefficient on percent never married (column 3) and say that marriages fell. But, given the weight of the evidence in Tables 2 through 6, we are fairly confident in concluding that there is no evidence to suggest that the economic activity associated with fracking led to an increase in marriage, but there is evidence that the positive income effects associated with the fracking boom led to an increase in births among both married and unmarried women.

Another possible explanation for the lack of evidence of an increase in marriage is that the effect is concentrated among less educated women, and since our analysis sample includes all women age 18 to 34, we are unable to detect the group-specific changes. In Table 7 we report the results of estimating our baseline birth and marriage regressions separately for women age 18 to 34 with and without a college degree. To do this requires limiting the sample to the subsample of states that record maternal education in the vital statistics natality files for all years. This results in a much smaller sample of 31 states plus the District of Columbia, yielding observations from 1,393 PUMAs. Panel A reports the results for all women estimated on this restricted sample. Panel B reports the results for non-college women and Panel C reports the results for women with a college degree.

As found with the full sample in the earlier tables, the results in Panel A show statistically significant increases in the marital birth rate and the non-marital birth rate. Very similar point estimates are found when the regressions are estimated for non-college educated women (reported in Panel B). In contrast, for college educated women, the effect on the marital birth rate is much smaller and the effect on the non-marital birth rate is small and negative. The estimated impacts in columns (1) and (2) for marital and non-marital birth rates are statistically different for non-college and college educated women. There are no statistically significant effects on the marriage related outcomes for either non-college or college educated women.

D. Comparison to the effects of the Appalachian coal boom in the 1970s and 1980s

In a paper titled "Are Children Normal?" Black, Kolesnikova, Sanders, and Taylor (2013) analyze how fertility among married couples responded to the increase in male earnings in the Appalachian coal-mining region of the United States in the mid-1970s. The Appalachian coal boom began in the 1970s when energy prices spiked and continued through the 1980s until energy prices plummeted. Black et al. (2013) exploit variation in county level coal reserves and yearly energy prices to estimate the causal effect of changes in earnings on marital birth rates. Specifically, they estimate an IV regression of the marital birth rate in a county as a function of total county earnings, where county earnings are predicted as a function of the value of coal reserves in the county. (They estimate the model in first differences.) They motivate and interpret their analysis as an empirical test of Becker's (1960) contention that children are normal goods, meaning that the demand for children increases in response to rising income. Black et al. estimate that a 10 percent increase in county earnings associated with the coal boom leads to a 7 percent increase in the marital birth rate.

During the time period studied by Black et al., non-marital birth rates were much less common than they are today. It is thus interesting to consider whether the non-marital birth response to the coal boom might have been different than what we are observing in the context of the localized fracking boom. In Table 8, we revisit the fertility response to the Appalachian Coal Boom by looking at both married and unmarried births. This extension builds directly on the reduced form results reported in Table 5 which suggested that places with low non-marital birth shares saw a larger increase in marital births than non-marital births associated with fracking production, but places with above-median rates of non-marital birth shares saw similarly large impacts on both marital and non-marital births.

We follow the approach of Black et al. in estimating an IV regression of birth rates as a function of earnings, but define earnings to be per capita and look specifically at PUMA-level births to women age 18 to 34. For the sake of comparison, we estimate an analogous IV regression of the relationship between PUMA-level earnings and birth and marriage in the context of the localized fracking periods using simulated fracking production as a source of exogenous variation in earnings (using all earnings, not just those recorded for jobs held by men).

The IV specification requires the strong exclusion restriction assumption that the only channel through which a localized fracking boom affected marriage and birth outcomes is through earnings, an assumption which is almost surely violated. Two obvious threats to this assumption are changes in house prices and changes in sex ratios. But, recall that these variables are controlled for in our model, so the IV estimate is conditional on those variables. Results reported in Appendix Table A2 confirm that the inclusion or exclusion of those variables does not alter the reduced form estimated effect of fracking production in birth outcomes, which suggests that these factors are not the mechanism through which fracking production is affecting family formation outcomes. Other potential threats to the validity of the exclusion restriction are changes in unearned income, changes in female earnings as opposed to male earnings, and changes in the male/female wage ratio. Conceptually, all of these variables could have a direct effect on marriage and birth outcomes and are affected by local fracking booms. Without additional instruments, we are unable to adequately control for these confounding factors. The estimated effects from this IV analysis should thus be interpreted with caution. Caveats notwithstanding, we report the results of this IV analysis because it provides a useful benchmark estimate that can be compared to estimates from the context studied by Black et. al (2013).

To conduct this analysis we use PUMA-level per capita earnings, which are aggregated up from county-level measures provided by the U.S. Bureau of Economic Analysis (BEA). The QWI data used in our earlier analyses only extends back to the 1990s and hence cannot be used to revisit the context of the Appalachian Coal Boom. To implement the IV strategy, we use coal reserve and price measures provided by Black et al. (2013) to instrument for the natural log of earnings per capita. By looking at the natural log of our outcome measures and the natural log of earnings we are able to interpret the coefficient as an elasticity estimate.

The estimates reported in Table 8, Panel A, Column (2) imply that a 10 percent increase in earnings led to a 7.5 percent increase in the marital birth rate among women age 18 to 34. This is comparable in magnitude to the estimated relationship during the fracking boom, as reported in Panel B, Column (2). The IV estimate of the relationship between per capita earnings associated with fracking and marital births implies that a 10 percent increase in earnings led to a 12.4 percent increase in the marital birth rate.

The non-marital birth response is very different between contexts. As reported in Table 7, Panel A, Column (3), the estimated relationship between earnings and non-married births during the Appalachian coal boom is negative and significant. A 10 percent increase in earnings associated with the coal boom led to a 25.5 percent reduction in the non-marital birth rate. Off of a mean of 12.8, this represents a reduction of 3.3 non-marital births per 1000 women. This contrasts sharply with the estimated relationship identified from variation in fracking production in the 2000s. As reported in Panel B, the IV results imply that a 10 percent increase in earnings associated with fracking production led to a 12.4 percent increase in non-marital births. The data also indicate a very different marriage response in the earlier and later periods. Column (4) reports the results of estimating the IV model for the dependent variable "share of women age

15-34 never married". The data suggest that in the earlier period, a 10 percent increase in per capita earnings was associated with a 9.6 percent reduction in the share of women age 15 to 34 who were never married.[13] The point estimate from the later period implies a statistically insignificant 1.7 percent increase for a 10 percent increase in per capita earnings.

To make a more direct comparison, we estimate these regressions for a common set of Appalachian PUMA in both periods. Unfortunately sample size limitations precludes a useful analysis along these lines. The coefficients from estimating the regression model on the three states from the Black et al. (2013) sample, suggest that during the fracking boom a 10 percent increase in earnings increased the marital birth rate by an insignificant 5.4 percent (s.e. = 3.7); increased the non-marital birth rate by a marginally significant 12.5 percent (s.e. = 6.3); and had no impact on the share never married. The results from estimating the model for the sample of Appalachian counties during the fracking boom are imprecisely estimated, but are consistent with an increase in non-marital births and no increase in marriage during this more recent context. (The point estimates on simulated new production for the key set of outcomes are as follows: 0.54 (s.e. 0.37) for the marital birth rate; 1.25 (s.e. 0.63) for the non-marital birth rate; and 0.03 (s.e. 0.59) for the female share never married.)

The contrast of findings between the context of the cool boom and bust of the 1970s and 1980s and the fracking boom of the 2000s is consistent with the notion that social context matters. In the earlier period, when non-marital births were still far from the norm, couples responded to the increase in earnings with increased rates of marriage and increased marital births, but no increase in births outside of marriage. In the later period, both marital and non-marital births increased significantly in response to the positive economic shock. And unlike

[13] We use this modified age range because in the 1970 census there are two age groups available: 15-24 and 25-34.

during the Appalachian coal boom, there is no discernible increase in marriage in response to the positive local economic shock associated with fracking.

Although this evidence is consistent with changing social norms, it is not definitive. There are other potential explanations for the differential responses observed between periods. For instance, the coal boom and bust particularly impacted male earnings (per Black et al, 2013). But, as we saw in Table 2, fracking increased the potential earnings of women as well, albeit to a lesser extent than for men. An increase in female earnings could mute the positive effect of male earnings on marriage rates, leading to the null effect found in the later period, and also make it more financially feasible for an unmarried woman to have a child without a spouse. It might also be the case that fracking jobs are particularly onerous, so that an increase in earnings for men directly employed in fracking-related jobs would make them less desirable marriage partners than would an increase in earnings for men employed in other jobs. We thus view the comparisons and contrasts of results between Panels A and B in Table 6 interesting, but our interpretation of them is necessarily speculative.

V. CONCLUSION

The fracking boom of the post 2005 period led to sizable improvements in the earnings potential of non-college educated men in counties located over geological shale plays. We use this context as a rare opportunity to investigate whether an increase in the employment and earnings potential of less educated men leads to an increase in marriage rates and a corresponding decrease in non-marital births. Our analysis is motivated by an interest in testing whether the reverse of the marriageable men hypothesis holds – *as the economic prospects of less educated men improves, are couples more likely to marry and are women less likely to have* *a child outside a marital union?* The data suggest that at least in the short term, if we extrapolate from the experience of the fracking boom, the answer is likely no.

The results of our analysis suggest that local area fracking production led to an increase in both marital and non-marital births, and no increase in marriage rates. The finding of a positive birth response is consistent with a positive income effect on births as found in the previous literature, and yields additional insight by considering marital and non-marital births separately in the same context. The data also suggest that the marital birth response is larger than the non-marital birth response in areas with relatively low non-marital birth shares at baseline, while the marital and non-marital responses are similar in areas with high non-marital birth shares at baseline. This would be consistent with a role of social norms in driving the family formation response to a local economic shock.

To further investigate the possibility that social context matters, we have compared the family formation response to the fracking boom of the 2000s to the family formation response to the Appalachian coal boom and bust of the 1970s and 1980s. The data indicate that the increased earnings associated with the coal boom during those earlier decades led to an increase in marriage rates and marital births, and no increase in non-marital births. In contrast, the increase in earnings associated with fracking in more recent years led to an increase in both marital and non-marital births, and no increase in marriage rates. This contrast is consistent with the notion that the family formation response to economic circumstances depends on social context. The patterns of results we observe might suggest that as non-marital births become increasingly common, individuals are more likely to respond to increased income with increased fertility, but not necessarily with marriage.

In conclusion, we find no evidence to support the proposition that as the economic prospects of less educated men improves, couples are more likely to marry before having children, as a reverse marriageable men hypothesis might predict. Crucially, this does not imply that the decline in the economic position of men in certain communities and demographic groups over the past four decades has not been a primary driver of the increase in non-marital birth rates and decrease in marriage rates in these communities and groups. It is quite possible that the reduction in male marriageability among less educated men in earlier decades was the driving force that led to a decline in marriage and a corresponding rise in non-marital childbearing, but now that non-marital childbearing has become so commonplace, a new social norm has been set and an increase in male economic prospects does not have the same effect as it would have in a different time or place. In other words, economics might have led to a new social structure such that now we are in a new paradigm. The proposition that individuals respond to economic circumstances in ways that are shaped by prevailing social norms warrants further empirical examination.

REFERENCES

Allcot, Hunt & Daniel Keniston. 2014. "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America," Working Paper No. 20508, National Bureau of Economic Research, Cambridge, MA.

Amato Paul and Brett Beattie. 2011. "Does the Unemployment Rate Affect the Divorce Rate: An Analysis of State Data 1960-2005." *Social Science Research* 40(3):704-715

Autor, David, & David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review* 103(5): 1553-1597.

Autor, David, David Dorn, and Gordon Hanson. 2017. "When Work Disappears: Manufacturing Decline and the Falling Marriage-Market Value of Men," NBER working paper 23173, February.

Bartik, Alexander, Janet Currie, Michael Greenstone, and Christopher Knittel. 2017. "The Local Economic and Welfare Consequences of Hydraulic Fracturing," NBER working paper 23060. January.

Becker, Gary. 1960. "An Economic Analysis of Fertility," in Becker, Gary, ed., *Demographic and Economic Change in Developed Countries.* Princeton: Princeton University Press.

Becker, 1974. "A Theory of Marriage" in Theodore Shultz, ed., *Economics of the Family: Marriage, Children, and Human Capital.* Chicago: University of Chicago Press.

Black, Dan, Natalia Kolesnikova, Seth Sanders, & Lowell J. Taylor. 2013. "Are 'Children Normal'?" *Review of Economics and Statistics*, 95(1): 21-33.

Browning, M., F. Bourguignon, P.-A. Chiappori, and V. Lechene (1994): "Income and Outcomes: A Structural Model of Intrahousehold Allocation," *Journal of Political Economy*, 102: 1067–96.

Cascio, Elizabeth & Ayushi Narayan. 2015. "Who Needs a Fracking Education? The Educational Response to Low-Skill Biased Technological Change," Working Paper No. 21359, National Bureau of Economic Research, Cambridge, MA.

Cherlin, Andrew, David Ribar, & Suzumi Yasutake. 2016. "Non-marital first births, marriage, and income inequality," *American Sociological Review*: pp. 1-22.

Charles, Kerwin Kofi, and Ching-Ming Luoh. 2010. "Male incarceration, the marriage market, and female outcomes." *Review of Economics and Statistics*, 92(August): 614–627.

Dettling, Lisa and Melisa S. Kearney. 2014. "House Prices and Birth Rates: The Impact of the Real Estate Market on the Decision to Have a Baby," *Journal of Public Economics* 110: 1-166.

Edin, Kathryn, & Maria Kefalas. 2005. *Promises I Can Keep: Why Poor Women Put Motherhood Before Marriage,* Berkeley: University of California Press, 1 edition ed.

Eliason, Paul and Christopher Timmins. 2014. "Measuring the Employment Impacts of Shale Gas Development," Working Paper. Duke University.

Fetzer, Thiemo. 2014. "Fracking Growth," Working Paper. Last accessed 11/21/2016: http://www.trfetzer.com/wp-content/uploads/fracking-local.pdf.

Feyrer, James, Erin Mansur, & Bruce Sacerdote. 2015. "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution," Working Paper No. 21624, National Bureau of Economic Research, Cambridge, MA.

Gold, Russell. 2014. "The Boom: How Fracking Ignited the American Energy Revolution and Changed the World." Simon & Schuster Paperbacks. New York , NY.

Gopalakrishnan, Sathya and H. Allen Klaiber. 2014. "Is the Shale Boom a Bust for Nearby Residents? Evidence from Housing Values in Pennsylvania," *American Journal of Agricultural Economics,* 96(1): 43-66.

Greenstone, Michael and Adam Looney. 2012. "The Marriage Gap: The Impact of Economic and Technological Change on Marriage Rates," A Hamilton Project Economic Analysis: http://www.hamiltonproject.org/assets/legacy/files/downloads_and_links/020312_jobs_greenstone_looney.pdf.

Hellerstein, Judith and Melinda Morrill. 2011. "Booms, Busts, and Divorce". *The B E Journal of Economic Analysis & Policy* 11(1):54-54.

Kearney, Melissa, and Phil Levine. 2016. "The Economics of Non-Marital Childbearing and the "Marriage Premium for Children," *Annual Review of Economics*, Forthcoming.

Lindo, Jason M. 2010. "Are children really inferior goods? Evidence from displacement driven income shocks," *Journal of Human Resources* 45(2), 301–327.

Lovenheim, Michael and Kevin Mumford. 2013. "Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market," Review of Economics and Statistics (2013) 95:2, 464-475.

Maniloff, Peter & Ralph Mastromonaco. 2014. "The Local Economic Impacts of Hydraulic Fracturing and Determinants of Dutch Disease," Division of Economic and Business Working Paper Series, Colorado School of Mines.

McLanahan, Sara & Tara Watson. 2011. "Marriage Meets the Joneses: Relative Income, Identity, and Marital Status," *Journal of Human Resources*, University of Wisconsin Press, vol. 46(3), pages 482-517.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The housing market impacts of shale gas development," *American Economic Review* 105(12), 3633-3659.

Ruggles, Steve, Katie Genadek, Ronald Geoken, Josiah Grover, & Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0 [dataset]. Minneapolis: University of Minnesota.* http://doi.org/10.18128/D010.V6.0.

Schaller, Jessamyn. 2012. "For Better if Not for Worse: Marriage and Divorce Over the Business Cycle," *Journal of Population Economics*.

Shenhav, Na'ama. 2016. "What Women Want: Family Formation and Labor Market Responses to Marriage Incentives," Job Market Paper. U.C. Davis Mimeo.

U.S. Census Bureau (Census). 2008. "A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know," U.S. Government Printing Office, Washington, DC.

U.S. Census Bureau (Census) 2014. Quarterly Workforce Indicators Data. Longitudinal-Employer Household Dynamics Program http://lehd.ces.census.gov/data/#qwi.

U.S. Census Bureau (Census). 2016. American FactFinder, http://factfinder2.census.gov.

Wilson, Riley. 2016. "Moving to Economic Opportunity: the Migration Response to the Fracking Boom," Available at SSRN: https://ssrn.com/abstract=2814147.

Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City the Underclass, and Public Policy*, Chicago: University of Chicago Press.

Wilson, William Julius and Kathryn Neckerman. 1986. "Poverty and Family Structure: The Widening Gap between Evidence and Public Policy Issues." In *Fighting Poverty: What Works and What Doesn't*," ed. Sheldon Danziger

Table 1. Summary Statistics for Analysis Sample Public Use Microdata Areas, Year 2000 (pre-fracking boom)

	Non-Fracking PUMA	Fracking PUMA
Marriage and Birth Outcomes		
Total births per 1000, women age 18-34	103.9	104.1
Non-marital birth share, women 18-34	34.2	33.4
Marital births per 1000, women age 18-34	68.6	69.3
Non-marital births per 1000, women 18-34	35.3	34.7
Percent women 18-34 never married	56.5	55.3
Percent women 18-34 married	37.6	38.6
Percent women 18-34 divorced	4.1	4.4
Percent women 18-34 cohabitating	7.6	7.5
Labor Market Characteristics		
Average Male Non-College Earnings (QWI)	43,295	42,930
Average Female Non-College Earnings (QWI)	27,241	26,560
Average Male College Earnings (QWI)	79,736	79,803
Average Female College Earnings (QWI)	45,997	45,436
Population Characteristics		
Percent 18-34 year olds, White-NH	62.7	64.1
Percent Men 18-34 with Coll. Degree	18.0	15.5
Percent Women 18-34 with Coll. Degree	21.5	18.4
Number of Public Use Microdata Areas (PUMA)	1,440	604

Sources: Data on births comes from the Vital Statistics natality files. Population counts used in the denominator come from the 2000 Decennial Census. Marital shares come from the 2000 Census. Wage data is obtained from Quarterly Workforce Indicators (QWI). County population characteristics obtained from U.S. Census Bureau using data from the 2000 Census and SEER population estimates, which are derived from the U.S. Census Bureau's population estimates. See data appendix for details.

Table 2. Effect of PUMA-level Simulated New Production on Labor Market Measures, 1997 to 2012

	Men			Women		
	Ln(Average Earnings)$_{py}$		Natural Log Jobs/Pop$_{py}$	Ln(Average Earnings)$_{py}$		Natural Log Jobs/Pop$_{py}$
	All Industries	Exclude Oil & Gas Extraction		All Industries	Exclude Oil & Gas Extraction	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All Workers						
Sim. New Production$_{py}$	0.038***	0.026**	0.052***	0.019***	0.016***	0.012
(Thousands 2010$ per Capita)	(0.013)	(0.011)	(0.020)	(0.006)	(0.005)	(0.010)
Dependent Means (in levels)	54372	54129	52.53	34183	34170	46.86
Observations	29,471	29,471	29,471	29,471	29,471	29,471
Panel B. Non-College Workers						
Sim. New Production$_{py}$	0.044***	0.029**	0.048**	0.023***	0.018***	-0.004
(Thousands 2010$ per Capita)	(0.014)	(0.012)	(0.019)	(0.007)	(0.006)	(0.011)
Dependent Means (in levels)	43626	43365	51.15	28531	28518	44.09
Observations	29,471	29,471	29,471	29,471	29,471	29,471
Panel C. College Workers						
Sim. New Production$_{py}$	0.006	0.001	0.048**	-0.014**	-0.013**	0.038**
(Thousands 2010$ per Capita)	(0.012)	(0.011)	(0.023)	(0.006)	(0.006)	(0.019)
Dependent Means (in levels)	80785	80586	61.25	48004	47991	60.47
Observations	29,466	29,471	29,466	29,469	29,471	29,469

Notes: Earnings and job data are from the U.S. Census Bureau Quarterly Workforce Indicators file (QWI). The unit of analysis is the Public Use Microdata Area (PUMA)/year; the sample consists of the 1,985 PUMA in our analysis sample that participate in the QWI. All monetary values are inflation-adjusted to 2010 dollars, using the BEA Personal Consumption Expenditure (PCE) index. To be consistent with regressions estimated in the subsequent birth and marriage models, all regression models reported in this table include controls for the female/male 18-34 sex ratio, the natural log of the average house price, gender by race shares for 18-34 year olds, gender by education shares for18-34 year olds, and stateXyear and PUMA fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1.

Table 3. Reduced Form Effect of Simulated New Production on PUMA-level Birth Outcomes for Women 18-34

	Percent Births Non-Marital$_{py}$ (1)	Total Birth Rate$_{py}$ (2)	Marital Birth Rate$_{py}$ (3)	Non-Marital Birth Rate$_{py}$ (4)
Sim. New Production$_{py}$	-0.11	5.96***	3.57***	2.39***
(Thousands 2010$ per Capita)	(0.41)	(0.96)	(1.00)	(0.60)
Dependent Means	38.42	101	62.33	38.68
Observations	32,704	32,704	32,704	32,704

Notes: Birth data from U.S. Vital Statistics Natality Files. The unit of analysis is the public use microdata area (PUMA)/year, including 2,044 PUMA from 1997 to 2012. In 2000, approximately 76 percent of births to women ages 18-34 were to women with less than a college degree. Sixty seven percent of marital births and 94 percent of non-marital births were to women with less than a college degree. The impact of simulated new production on marital and non-marital birth rates is not statistically different. All monetary values are inflation-adjusted to 2010 dollars, using the Personal Consumption Expenditure (PCE) index. All regression models include controls for the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, and state X year and PUMA fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

Table 4. Reduced Form Effect of Simulated New Production on PUMA-level Marriage Outcomes for Women Age 18 to 34

	Percent Never Married$_{py}$ (1)	Percent Married$_{py}$ (2)	Percent Newly Married$_{py}$ (3)	Percent Divorced$_{py}$ (4)	Percent Cohabitating$_{py}$ (5)
Sim. New Production$_{py}$	1.17	-1.37	-0.22	0.11	0.25
(Thousands 2010$ per Capita)	(0.86)	(0.85)	(0.28)	(0.29)	(0.39)
Dependent Mean	65.90	29.33	0.954	3.338	8.242
Observations	16,334	16,334	14,287	16,334	16,334

Notes: All outcomes are at the PUMA level and are constructed from the 2000 Decennial Census and the 2005-2011 5-year ACS public use microdata because county is only available for large counties in the public use census data. The unit of analysis is the PUMA/year, including 2,044 PUMA from 2000 and 2005-2011. The share newly married cannot be constructed from the 2000 Decennial Census. All monetary values are inflation-adjusted to 2010 dollars, using the PCE index. All regression models include controls for the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, PUMA and stateXyear fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

Table 5. Reduced Form Effect of Simulated New Production by Baseline Non-Marital Birth Share for Women Age 18 to 34

	Above median in 2000			Below median in 2000		
	Marital Birth Rate$_{py}$ (1)	Non-Marital Birth Rate$_{py}$ (2)	Percent Never Married$_{py}$ (3)	Marital Birth Rate$_{py}$ (4)	Non-Marital Birth Rate$_{py}$ (5)	Percent Never Married$_{py}$ (6)
Sim. New Production$_{cy}$ (Thousands 2010$ per Capita)	2.26***	2.52***	0.63	3.84**	2.21*	2.07
	(0.79)	(0.71)	(1.05)	(1.79)	(1.31)	(1.30)
Dependent Means	54.26	45.73	67.83	70.40	31.59	62.80
Observations	16,320	16,320	8,142	16,336	16,336	8,168

Notes: Birth data from U.S. Vital Statistics Natality Files. The unit of analysis is the public use microdata area (PUMA)/year including 2,044 PUMA from 1997-2012 in columns (1)-(3) and (5)-(7) and from 2000 and 2005-2011 in columns (4) and (8). PUMA are defined as having a high baseline non-marital birth share if the non-marital birth share in 2000 was above the median of 33.6 All monetary values are inflation-adjusted to 2010 dollars, using the Personal Consumption Expenditure (PCE) index. All regression models include controls for the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, and stateXyear and PUMA fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

Table 6. Long Difference Reduced Form Effect of Simulated New Production on Birth and Marriage Outcomes for Women 18 to 34

	Outcomes: 2000 to 2011 Long Difference					
	Marital Birth Rate$_p$ (1)	Non-Marital Birth Rate$_p$ (2)	Percent Never Married$_p$ (3)	Percent Married$_p$ (4)	Percent Divorced$_p$ (5)	Percent Cohabitating$_p$ (6)
Total Sim. New Production$_{p2000\text{-}2011}$	0.813***	0.229*	0.376**	-0.303	-0.046	0.041
(Thousands 2010$ per Capita)	(0.219)	(0.122)	(0.190)	(0.196)	(0.067)	(0.078)
Dependent Mean in 2000	68.81	35.15	56.13	37.92	4.201	7.589
Observations	2,044	2,044	2,041	2,041	2,041	2,041

Notes: Birth data from U.S. Vital Statistics Natality Files. Puma-level marriage outcomes are constructed from the 2000 Decennial Census and 2005-2011 American Community Survey (ACS) microdata. The unit of analysis is the long difference between 2000 and 2011 for 2,044 Public Use Microdata Areas (PUMA). In the ACS three PUMA in New Orleans were disbanded after Hurricane Katrina. The share cohabitating is defined (imperfectly) as the share of women 18-34 who are either recorded as an unmarried partner, or as the household head with an unmarried partner present. The impact of total simulated production on marital and non-marital birth rates is statistically different (p=.02). All monetary values are inflation-adjusted to 2010 dollars, using the PCE index. All regression models include controls for the change in the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, and state fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Robust standard errors are reported *** p<.01, ** p<.05, * p<.1. ***

Table 7. Reduced Form Effects for Non-College and College Educated Women Separately

	Marital Birth Rate$_{py}$ (1)	Non-Marital Birth Rate$_{py}$ (2)	Percent Never Married$_{py}$ (3)	Percent Married$_{py}$ (4)	Percent Newly Married$_{py}$ (5)	Percent Divorced$_{py}$ (6)	Percent Cohabitating$_{py}$ (7)
Panel A. All Women 18-34							
Sim. New Production$_{py}$	3.49***	2.56***	1.23	-1.29	-0.20	0.09	0.28
(Thousands 2010$ per Capita)	(1.13)	(0.61)	(1.02)	(1.01)	(0.26)	(0.32)	(0.43)
Dependent Means (in Levels)	63.19	38.94	65.02	30.09	0.856	3.422	8.074
Observations	22,288	22,288	11,144	11,144	11,144	11,144	11,144
Panel B. Non-College Women 18-34							
Sim. New Production$_{py}$	3.97***	2.26***	1.59	-1.59	-0.13	0.14	0.12
(Thousands 2010$ per Capita)	(1.25)	(0.81)	(1.05)	(1.04)	(0.26)	(0.34)	(0.45)
Dependent Means (in Levels)	53.42	46.71	68.08	26.77	0.772	3.545	8.168
Observations	22,288	22,288	11,144	11,144	11,144	11,144	11,144
Panel C. College Women 18-34							
Sim. New Production$_{py}$	0.86**	-0.09*	-4.85	3.89	-1.14	-0.21	1.30
(Thousands 2010$ per Capita)	(0.37)	(0.05)	(3.26)	(3.00)	(0.71)	(0.92)	(1.23)
Dependent Means (in Levels)	27	2.211	50.28	46.20	1.135	2.755	6.875
Observations	22,288	22,288	11,133	11,133	11,133	11,133	11,133

Notes: Birth data from U.S. Vital Statistics Natality Files. The unit of analysis is the Public Use Microdata Area (PUMA)/year, including 1,393 PUMA from 1997 to 2012 for the 31 states (plus the District of Columbia) that have mother's education available for all years. Marriage outcomes are only observed in 2000, and 2005-2011. All monetary values are inflation-adjusted to 2010 dollars, using the Personal Consumption Expenditure (PCE) index. All regression models include controls for the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, and stateXyear and PUMA fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

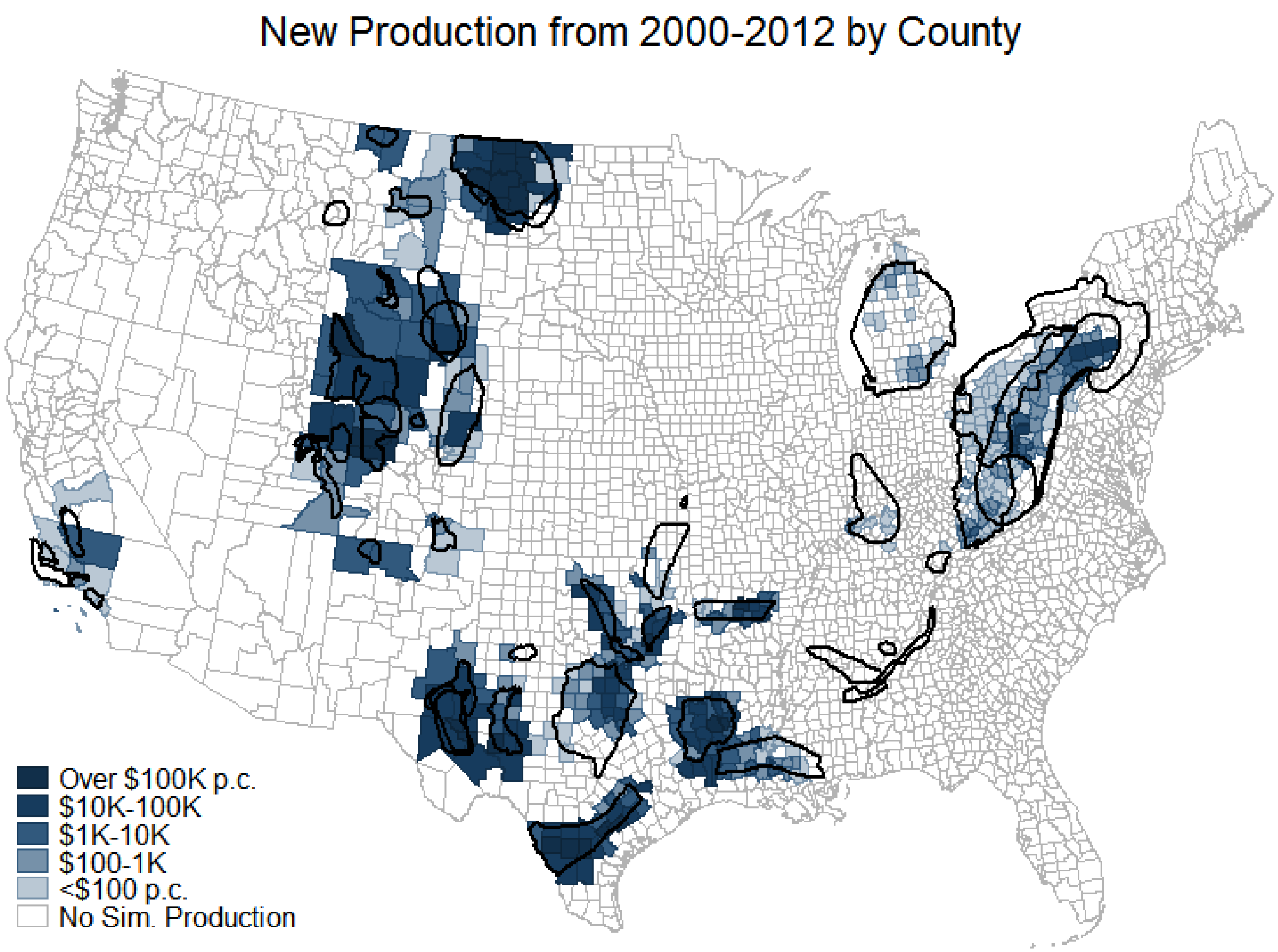
Table 8: Comparison to the Appalachian Coal Boom Context: IV Estimates of the Effect of Earnings on Marital Births, Non-Marital Births, and Share of Women Never Married

Panel A.	Appalachian Coal Boom (1969-1987)			
	Women 18-34			Women 15-34
Natural Log of -	Percent Births Non-Marital	Marital Birth Rate	Non-Marital Birth Rate	Percent Never Married
	(1)	(2)	(3)	(4)
Ln Earnings Per Capita$_{py}$	-3.03***	0.75***	-2.55***	-0.96***
	(0.55)	(0.23)	(0.47)	(0.18)
First Stage F-statistic	56.3	56.3	56.3	34.0
Dependent Means (in Levels)	13.94	81.63	12.81	51.69
Observations	2,546	2,546	2,546	268

Panel B.	Fracking Oil and Gas Boom (1997-2012)			
	Women 18-34			Women 15-34
Natural Log of -	Percent Births Non-Marital	Marital Birth Rate	Non-Marital Birth Rate	Percent Never Married
Ln Earnings Per Capita$_{py}$	-0.11	1.24***	1.24***	0.17
	(0.25)	(0.43)	(0.43)	(0.21)
First Stage F-statistic	11.8	11.8	11.8	8.6
Dependent Means (in Levels)	38.42	62.33	38.68	59.82
Observations	32,704	32,704	32,704	4,088

Notes: All birth data from Vital Statistics. Share never married is calculated from the 1970 and 1980 censuses (provided by National Historical Geographic Information System) as well as the 2000 decennial census and 2011 5-year ACS. We use coal reserve values to instrument for the natural log of total earnings per capita (from the BEA) for the Coal Boom, similar to Black et al. (2013). Data on coal reserves for 1969-1988 are constructed using data provided by Black et al., (2013). Sample for the Coal Boom period includes 134 Public Use Microdata Areas (PUMA). Sample for the fracking boom period includes 2,044 PUMA. All regression models include controls for the female population 18-34, male/female sex ratio age 18-34, gender by race shares, and stateXyear and county fixed effects. In Panel B we also control for the natural log average house price. This measure is not available for the earlier period. Estimates are weighted by the total number of births to 18-34 year olds in 1970 and 2000 respectively. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

Figure 1. Total Oil and Gas Production from Wells During First Year of Production, 2000-2012



Notes: Shale play boundaries outlined in black. Source: Constructed by the authors from DrillingInfo.

Appendix Tables

Table A1. Summary Statistics for Fracking Public Use Microdata Areas (PUMA)

	2004	2005	2006	2007	2008	2009	2010	2011	2012
Number of PUMA with Sim. New Production	331	393	367	443	476	523	553	558	581
	Simulated Annual New Production Value by Percentile (2010$ per Capita)								
10th	0.12	0.07	0.13	0.04	0.06	0.03	0.13	0.11	0.20
50th	1.46	1.42	1.72	1.43	2.45	1.48	2.51	2.88	3.05
90th	59.10	58.89	93.51	72.94	96.43	61.86	66.41	75.20	71.01
	Average Simulated Annual New Production Value (2010$ per Capita)								
All Producing PUMA	35.67	37.95	39.38	40.04	59.54	41.12	53.89	66.43	65.66
Top 10%	316.75	338.57	338.17	357.86	538.72	373.01	512.66	606.53	592.69

Notes: The statistics in this table refer to counties in the analysis sample (consisting of 2,044 PUMA, per Table A1) that have any simulated new production between 2004 to 2012.

Table A2. Robustness and Specification Sensitivity, Women Age 18 to 34

	Baseline (1)	Exclude Sex Ratio and House Price Controls (2)	Un-Weighted (3)	Year Fixed Effect (4)	Shale Play by Year Fixed Effects (5)	Natural Log of Outcome (6)
			Marital Birth Rate$_{py}$			
Sim. New Production$_{py}$	3.571***	3.419***	3.442***	2.155**	3.113***	0.053***
(Thousands 2010$ per Capita)	(1.001)	(1.099)	(0.831)	(1.046)	(1.155)	(0.013)
Dependent Means	62.33	62.33	62.33	62.33	62.33	62.33
Observations	32,704	32,704	32,704	32,704	32,704	32,704
			Non-Marital Birth Rate$_{py}$			
Sim. New Production$_{py}$	2.388***	2.713***	2.324***	3.516***	2.459***	0.039***
(Thousands 2010$ per Capita)	(0.601)	(0.723)	(0.506)	(0.784)	(0.832)	(0.009)
Dependent Means	38.68	38.68	38.68	38.68	38.68	38.68
Observations	32,704	32,704	32,704	32,704	32,704	32,704

Notes: Birth data from U.S. Vital Statistics Natality Files. The unit of analysis is the Public Use Microdata Area (PUMA)/year, including 2,044 PUMA from 1997 to 2012. All monetary values are inflation-adjusted to 2010 dollars, using the Personal Consumption Expenditure (PCE) index. Column (1) provides baseline estimates from Table 3. Column (2) excludes the female to male sex ratio and natural log average house price. Column (3) provides unweighted estimates. Column (4) includes year effects, rather than stateXyear, comparing PUMA across states. Column (5) includes shale playXyear effects, rather than stateXyear, comparing PUMA within a shale play. Column (6) estimates the relationship when the outcome is measured in logs, to approximate the percent change. All regression models include controls for the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, and stateXyear and PUMA fixed effects unless otherwise specified. Estimates are weighted by the total number of births to 18-34 year olds in 2000 unless otherwise specified. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

Table A3. Heterogeneity by Demographics, Women age 18 to 34 (Unless otherwise specified)

	Women 35-44 (1)	N.H. White (2)	N.H. Black (3)	Hispanic (4)	N.H. Other (5)	First Births (6)	Higher Parity Birth (7)
				Marital Birth Rate$_{py}$			
Sim. New Production$_{py}$	0.79**	4.35**	1.49	0.08	-18.99	1.17***	2.40***
(Thousands 2010$ per Capita)	(0.34)	(2.11)	(1.59)	(2.06)	(17.75)	(0.36)	(0.82)
Mean Number of Births	242.6	708.2	74.34	234.9	89.25	347.1	759.5
Dependent Means	21.23	65.70	32.61	69.07	84.65	19.43	42.91
Observations	32,704	32,704	32,701	32,704	32,704	32,704	32,704
				Non-Marital Birth Rate$_{py}$			
Sim. New Production$_{py}$	-0.37**	4.14***	-0.49	0.59	-9.40	1.26***	1.13**
(Thousands 2010$ per Capita)	(0.15)	(0.69)	(2.01)	(1.62)	(6.31)	(0.46)	(0.49)
Mean Number of Births	51.95	253	189.6	215.3	28.31	254.4	431.9
Dependent Means	4.656	23.26	66.95	61.20	28.38	14.35	24.33
Observations	32,704	32,704	32,701	32,704	32,704	32,704	32,704

Notes: Birth data from U.S. Vital Statistics Natality Files. The unit of analysis is the public use microdata area (PUMA)/year, including 2,044 PUMA from 1997 to 2012. All monetary values are inflation-adjusted to 2010 dollars, using the Personal Consumption Expenditure (PCE) index. All regression models include controls for the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, and state X year and PUMA fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

Table A4. Reduced Form Effect of Actual New Production on Birth and Marriage Outcomes for Women Age 18 to 34

	Marital Birth Rate$_{py}$ (1)	Non-Marital Birth Rate$_{py}$ (2)	Percent Never Married$_{py}$ (3)	Percent Married$_{py}$ (4)	Percent Newly Married$_{py}$ (5)	Percent Divorced$_{py}$ (6)	Percent Cohabitating$_{py}$ (7)
Actual New Production$_{py}$	0.426***	0.197**	-0.038	0.060	-0.011	-0.029	0.044
(Thousands 2010$ per Capita)	(0.145)	(0.086)	(0.106)	(0.107)	(0.029)	(0.044)	(0.044)
Dependent Means	62.33	38.68	65.31	29.87	0.852	3.362	8.101
Observations	32,704	32,704	16,334	16,334	16,334	16,334	16,334

Notes: Birth data from U.S. Vital Statistics Natality Files. The unit of analysis is the Public Use Microdata Area (PUMA)/year, including 2,044 PUMA from 1997 to 2012. Marriage outcomes are only observed in 2000, and 2005-2011. The impact on marital and non-marital birth rates is not statistically different. All monetary values are inflation-adjusted to 2010 dollars, using the Personal Consumption Expenditure (PCE) index. A regression of actual new production on simulated new production suggests that an additional one thousand dollars of simulate production per capita is associated with approximately 5,400 dollars of actual production per capita, suggesting that the impacts of actual production are approximately half the size. All regression models include controls for the natural log average house price, the male/female sex ratio age 18-34, gender by race shares for 18-34 year olds, gender by education shares for 18-34 year olds, and stateXyear and PUMA fixed effects. Estimates are weighted by the total number of births to 18-34 year olds in 2000. Standard errors are adjusted for clustering at the PUMA level. *** p<.01, ** p<.05, * p<.1. ***

Data Appendix

Fracking Production Data

To construct measures of fracking production, we overlay the latest shale play boundary shapefiles available at the time from 2015 (provided by the Energy Information Administration (EIA)) onto county boundary shapefiles (provided by the U.S. Census Bureau). ArcGIS software is used to determine if a shale play and county intersect.[14] Although there are 48 unique shale plays in the US, some of these plays are small and cover only a few counties. We combine small shale plays that cover less than nine counties into an “other” category so that total play level production is not driven by any one county. We first generate simulated production at the county level, as there are many counties in each play. We then aggregate both actual new production and simulated new production up from the county level to the PUMA level. For the most part, a county will be completely contained within a PUMA. Some densely populated counties, are divided into multiple PUMA, in which case we use the share of the county population in 2000 in each PUMA to assign production values. Only 35 counties with production are partitioned into two or more PUMA.

These geographic measures are then combined with quarterly well level production data provided through a restricted access agreement with DrillingInfo, a private firm that collects lease, permit, and production data on wells drilled in the United States. These data are proprietary, and obtained through an academic use agreement with DrillingInfo, available through their academic outreach initiative. Oil and gas production are reported in barrels and thousands of cubic feet respectively. We use annual national prices provided by the Energy Information Administration (EIA) to convert output metrics into dollar amounts, and we then use the Personal Consumption Expenditure (PCE) price index calculated by the United States Bureau of Economic Analysis to adjust all dollar amounts to year 2010 dollars. DrillingInfo does not indicate if a well is a fracking well, as fracking is a means of stimulating production, but it does report whether a well is horizontally or vertically drilled. We assign non-vertical wells drilled in counties that intersect with shale plays as fracking wells and then aggregate this production to the PUMA level.

Jobs and Earnings Data

We use the Quarterly Workforce Indicators (QWI) as our primary source of wage data. The QWI is aggregated from the Longitudinal Employer-Household Dynamics (LEHD) micro-level data collected from unemployment insurance earnings data from participating states and several other sources (U.S. Census, 2014). The QWI is aggregated to the county level, and can be tabulated by firm characteristics (industry, size) or worker characteristics (gender, age, education). When tabulating by worker characteristics, only two levels of tabulation are feasible (gender by age or gender by education). For this reason, we consider gender by education level

[14] A special thanks to Lisa Boland and Michael Bender for their help with ArcGIS.

measures for all workers over the age of 25. The QWI data is constructed through a state sharing process, and as such, only states that have made agreements with the Census have reported data. Many of the states began participating in 2000 with most participating by 2003.

The main measure we use is the beginning of quarter earnings for all jobs. This measures the quarterly earnings for all jobs that existed at the beginning of the quarter. Other options (not used) include stable jobs (spanning multiple quarters) or total jobs (employed at any time during quarter). We take the implied average annual earnings across all four quarters, accounting for differences in employment levels by weighting by the reported quarter specific job count. To do this we aggregate up the total number of jobs as well as the total earnings from the county level to the PUMA level, and then calculate the earnings per worker. We also construct a jobs to population ratio by taking the total job count, averaged across quarters, and dividing it by the estimated population 18 and older from the SEER population estimates.

Because the QWI is constructed from firm employment, all measures are constructed for the job count. This means that average quarterly earnings are the average earnings of all jobs in a given quarter. Individuals who are unemployed are not considered, and individuals who hold two jobs will be treated as two separate individuals. In general, average earnings levels in the QWI are higher than those calculated elsewhere, as it records average earnings conditional on working. Also, because some workers might hold jobs for less than the full year, the average annual earnings constructed from the QWI will be higher, because the construction implicitly assumes the job lasts the entire year. This measure of earnings is relevant as it captures the "potential earning" opportunity available to a man.

Alternative labor market measures could be constructed from the Quarterly Census of Employment and Wages (QCEW) or the 2000 Census and ACS microdata. The QCEW provides county level employment and wage measures, but it does not separately identify jobs by the gender or education of workers. Because we are particularly interested in looking at non-college educated workers, we rely on the QWI. The ACS microdata would allow us to construct PUMA level measures based on the individual, rather than the job, and more closely corresponds to realized earnings rather than potential earnings. The ACS also presents several data constraints. Consistent PUMA geo-codes are only available in 2000 and 2005-2011, and the ACS is a one percent survey of households, which likely would yield less precise estimates than the QWI which draws from the near universe of jobs.

Birth Outcomes

We use restricted access Vital Statistics data provided by the National Center of Health Statistics to construct county level measures of birth outcomes, which are then aggregated to the PUMA level. The restricted files provide the universe of births between 1998 and 2013 with county identifiers. We focus our analysis on mothers between 18 and 34. In theory, we are interested in fertility decisions of less educated women. However, a limitation of the Vital Statistics data is its measurement of mother's educational attainment. Starting in 2009, the Vital Statistics required states to updated measures of maternal education to the new 2003 classification. States that did not comply did not have maternal education recorded. As a result, in 2009, 20 states are missing measures of mother's education. Starting in 2010, some of these states' education measures were included again. As educational attainment is potentially endogenous, and to maintain all states in our analysis we focus on births to women 18 to 34, in 2000, 76 percent of births to women 18 to 34 were to women with less than a four year college degree.

We convert births to conceptions leading to a live birth by subtracting the length of gestation from the fifteenth of the month of birth. This shifts the timing to the time of conception. We do this to consider the time the pregnancy decision was made. Our analysis sample thus consists of live births that were conceived between 1997 and 2012. To construct age specific birth rates, we divide the number of births by the number of women 18-34 in each PUMA, constructed from the SEER gender by age population estimates, which are derived from the intercensal population estimates. We then multiply this by 1000 to be interpreted as birth per 1000 women in the age group of interest.

Marriage Outcomes

PUMA level marriage outcomes are constructed from the 2000 Census and 2005-2011 ACS microdata (Flood, 2015). Over these years consistent PUMA geo-codes are available, allowing us to estimate the share of 18 to 34 year old women who are never married, married, divorced, cohabitating, and married within the past year (share newly married). We use the population weights provided to do this. Never married, married, and divorced are created using the marital status variable. To construct the share cohabitating, we assign all women who are (1) the head of household with an unmarried partner present in the household, or (2) listed as unmarried partners in the household roster. The variable indicating if an individual was married in the past year is not available in the 2000 Census, so the share married within the past year is only available from 2005-2011.