# MOVING TO JOBS: THE ROLE OF INFORMATION IN MIGRATION DECISIONS*

Riley Wilson

December 31, 2018

Click here for latest version

**Abstract**

This paper exploits county-level variation in exposure to news about labor markets impacted by fracking to show that access to information about employment opportunities affects migration. Exposure to newspaper articles about fracking increased migration to areas mentioned in the news by 2.4 percent on average. Commuting also increased, positive news has larger effects than negative news, and TV news also has an impact. Google searches for the term fracking and the names of states specifically mentioned spike after news broadcasts about fracking. Counties experiencing weak labor markets are the most responsive, suggesting these areas see large benefits to information provision.

**Keywords:** geographic mobility, migration, information, news, fracking

**JEL Codes: J61, D83, R23, Q33**

*Brigham Young University, Department of Economics, 130 Faculty Office Building, Provo UT, 84602. Email: riley_wilson@byu.edu. I am grateful to Melissa Kearney, Judy Hellerstein, Katharine Abraham, Sergio Urzua, Ethan Kaplan, Nolan Pope, Joe Price, Cody Tuttle, Fernando Saltiel, as well as individuals at the University of Maryland, Brigham Young University, NC State, University of Arizona, University of Utah Finance Department, Federal Reserve Board, Atlanta Federal Reserve Bank, New York Federal Reserve Bank, Upjohn Institute, Southern Economic Association, and the Chicago Labor Dynamism Conference for useful comments. I gratefully acknowledge funding through the Economic Club of Washington Vernon E. Jordan Jr. doctoral fellowship.

# I Introduction

Migration is traditionally viewed as a human capital investment that provides access to more favorable labor market opportunities (Shultz, 1961; Sjastaad, 1962). However, people are unlikely to move away from areas experiencing negative labor demand shifts (Monras, 2015), even though the substantial geographic heterogeneity in employment prospects would suggest that many could encounter more favorable labor markets if they moved to a different state or county (BLS, 2017). These patterns have led many to wonder why populations in weak labor markets appear unlikely to move to better labor market opportunities.[1] As the previous literature recognizes, low migration might be the outcome of optimal decision-making, but might also be the result of market frictions such as credit constraints or incomplete information.

While incomplete information introduces uncertainty and increases the risk associated with the migration "investment", information provision may increase perceived returns and change migration outcomes. However, the role of information in migration decisions has frequently been overlooked in the empirical work because it is difficult to measure and identify a causal relationship. In this paper, I evaluate the role of information in migration decisions by asking, does information in the news about potential, lucrative employment opportunities in other labor markets induce people to move to those markets?

By focusing on national news about labor markets impacted by fracking, I can isolate one particular source of information transmission and estimate its causal impact on origin-destination specific migration flows. Over the last 10 to 15 years, the combination of two technologies, horizontal drilling and hydraulic fracturing, has led to localized "fracking booms" and sudden, large increases in both local employment and earnings. These booms have not

[1]Not only is out-migration from negatively affected areas low, but it has decreased over time (Dao, Furceri, & Loungani, 2017). This concern is evident in the news, such as the *New York Times'* "Fewer Americans Strike Out for New Jobs, Crimping the Recovery" by Patricia Cohen (2016) and "How to Get Americans Moving Again" by Arthur Brooks (2016).

only affected workers in oil and gas extraction, but have created large, persistent effects across industries (Feyrer, Mansur, & Sacerdote, 2017; Maniloff & Mastromonaco, 2014), resulting in perceived net benefits (Bartik et al., 2017). The "gold rush" style flurry of economic activity associated with fracking, and later the environmental and safety concerns, have led to numerous newspaper articles and television news broadcasts touting the economic impacts or debating the adverse side effects. As such, fracking created plausibly exogenous, positive labor market shifts that were talked about in the news. Because fracking was so novel it also introduced new words and vocabulary, making it easy to parse newspapers and TV news to see which sources were talking about fracking, what places they were talking about, what they were saying, and when they were saying it. This allows me to determine how exposed people were to news about fracking in a specific destination by measuring to what extent that news was circulated in their local area.

Although these labor market shifts were plausibly exogenous, the news and exposure to the news about fracking was not necessarily. Local newspapers make content decisions to cater toward their local geographic preferences (Gentzkow & Shapiro, 2010), and individuals can choose which news sources they consume. By combining location specific newspaper content from LexisNexis and TV news content from the Vanderbilt Television News Archive from national news outlets (whose content decisions are plausibly exogenous to local migration trends), with proprietary pre-fracking newspaper-specific county circulation rates from the Alliance for Audited Media (AAM) and TV channel viewership rates from the Television and Cable Factbook (to eliminate endogenous changes in penetration), I am able to isolate a plausibly exogenous component of exposure to news about fracking in a specific place.

Even though national newspapers do not adjust content to each county's idiosyncratic trends and preferences, they will likely produce more content about places that are more impacted by fracking. This is potentially problematic for identification if these characteristics also directly affect migration outcomes. As such, I control for time invariant origin/destination pair specific characteristics (like distance) and then compare migration flows

from different origins – with different levels of exposure– to the same destination. For example, in 2012 the *USA TODAY* published six articles about fracking in Pennsylvania. I test to see if origins that had historically higher circulation of the *USA TODAY* (and thus higher exposure) saw larger increases in migration to Pennsylvania fracking counties when this news is distributed. This specification will hold constant any characteristic of the destination that might be changing over time, and relies on variation across origins in historic circulation to identify the effect of news exposure on migration.

I estimate that when counties are exposed to this national newspaper news about a particular destination state (e.g., North Dakota or Pennsylvania) the annual number of migrants to fracking counties in that state increase by 2.4 percent on average. Exposure to news about fracking also increases cross county commute flows by 6.6 percent on average. This translates into approximately one to two additional migrants and four additional commuters to the fracking destination from each origin on average. Although this response is small, it is economically significant given the scope of the "treatment" and the aggregate effects at the destination. If there had been no news about local fracking booms in 2012, migration flows to fracking counties would have been 4.2 percent lower and commute flows would have been 11.7 percent lower on average.

One concern with this identification strategy is that the pre-fracking level differences in circulation – which generate the identifying variation – might be correlated with other origin level characteristics that are changing over time and affect migration. For example, counties with high readership of the *USA TODAY* might be more affluent and see larger income growth over time. If this additional growth in income affects migration decisions, the estimates would be biased. The data allow me to evaluate this concern in several ways. First, I find that areas with high historic circulation and low historic circulation follow similar trends in migration to fracking areas in the pre-period, and only diverge once they are "treated" with this news. Second, the estimates are insensitive to the inclusion of time-varying origin level controls. Finally, because, there are 16 states involved in fracking, there

are 16 potential destinations for each origin. This allows me to include an origin by year fixed effect and account for any observable or unobservable characteristics of the origin that are changing over time and affect migration. When making this comparison, the estimated effect of national news exposure is virtually the same, suggesting that when counties are exposed to this news there is an increase in migration to the destinations being discussed, and it is not driven by unobserved characteristics of the destination or origin that are changing over time. The estimates are also robust to controls for local news exposure, various functional forms, sample restrictions, and an alternative strategy comparing neighboring counties on either side of a local newspaper's distribution market.

Given the robustness of this result, I conduct additional analyses to better understand how the news influences migration behavior. Looking simultaneously at both newspaper exposure and TV news exposure, I find responses to both sources of information. The magnitude of the effects are similar, despite the fact that typical TV viewership is significantly higher than typical national newspaper circulation rates. Using linguistic parsing techniques, I find that news articles that are more "positive" discussing things like, jobs, booms, or growth, have a larger positive effect on migration than "negative" news articles, discussing contamination, pollution or earthquakes. The effect of negative news is still positive, suggesting "any news is good news" as it might provide information about where fracking is occurring. Positive and negative news affect cross-county commuting similarly, consistent with the theory that non-resident workers mostly experience the benefits of fracking while not incurring many of the costs (e.g., potential home water contamination). The migration response is largely driven by the *USA TODAY*, rather than the *New York Times* or *Wall Street Journal*, which are likely less accessible to the relevant population. The effect of news exposure on both migration and commuting also varies with distance, peaking for counties 400 to 1,000 miles away from the potential fracking destination, consistent with people being aware of nearby opportunities, but lacking information about distant labor market opportunities. In the days following a news broadcast about fracking, Google search interest in

the term "fracking" and the names of specific states mentioned in the news spike, consistent with people going online to seek more information.

The data suggest that the effect of newspaper exposure is over twice as large in origin counties with weak labor markets as it is in stronger labor markets, even though they face similar levels of exposure. This would suggest that providing labor market information can be a way of increasing geographic mobility, and might be particularly effective if targeted toward weak labor markets where the returns to migration are plausibly the largest and where we have also observed non-responsiveness in the past.

# II Information in Migration Decisions

## II.A Related Empirical Evidence

The propensity to migrate varies significantly by demographics, educational attainment, and geographic region (Molloy, Smith, & Wozniak, 2011), and some demographic groups are more likely to move in response to local labor market conditions than other groups (Bound & Holzer, 2000; Wozniak, 2010). Differential migration responses might be optimal if, for example, individuals are differentially affected by labor market shocks (Notowidigdo, 2013), or if strong labor markets also have high costs of living, resulting in a small or negative net return to migration for low wage workers (Ganong & Shoag, 2017). However, there is also credible evidence that liquidity constraints, credit constraints, and other market frictions impact the migration decision (Kling, Liebman, & Katz, 2007; Bryan, Chowdhury, & Mobarak, 2014). One potential frictions is a lack of information.

It has long been recognized that information will affect migration decisions, but most of the empirical work has been limited to focusing on the role of networks or linguistic and cultural enclaves.[2] Although there is not much work that speaks directly to the impact of labor market information on migration decisions, several recent studies have explored

[2]See for example Greenwood (1975), Winters, de Janvry, & Sadoulet (2001), Munshi (2003), McKenzie & Rapoport (2007, 2010), and Hanson & McIntosh (2010).

somewhat related topics. The Moving to Opportunity (MTO) experiment, which provided guidelines and information about local neighborhood poverty along with housing vouchers, induced treated households to move to more affluent neighborhoods, suggesting this type of information can change migration behavior (Kling et al., 2007). Although this did not improve economic outcomes for treated adults, recent work has found positive long-run effects for the young children who were treated (Chetty, Hendren, & Katz, 2015). Malamud and Wozniak (2012) exploit variation in the Vietnam draft and find that college attendance causally increased the incidence of migration. They suggest that exposure to other areas, and peers from other areas, provide information about alternative labor market opportunities. Using a structural model similar to the conceptual framework presented here, Kaplan and Schulhofer-Wohl (2017) propose a framework, where information helps people learn about amenities in different locations. This information updating process can also be applied to expectations about labor market opportunities.

In a randomized controlled trial in Bangladesh, households that received information about potential labor market opportunities and a conditional cash transfer were more likely to migrate, while households that were only given the information were not (Bryan et al., 2014). This suggests that relaxing credit constraints and information barriers together could increase migration. Farre and Fasani (2013), show that as villagers in Indonesia gain access to more TV stations, they become less likely to move. The authors propose that this is because media access corrects overly optimistic expectations of the return to migration. However, it is difficult to generalize the results from Bangladesh and Indonesia to the United States.[3] The question of how to conceptualize the role of labor market information in migration decisions still remains.[4]

[3]For example, the conditional round-trip transfer in Bangladesh was only equal to $8.50 (about one weeks work), suggesting these people are highly credit constrained (Bryan et al,. 2014).

[4]An early related literature explores how things like the risk of unemployment (Todaro, 1969) and uncertainty about the future affect migration and human capital investments more generally (see Becker, 1962; Greenwood 1975, 1985; Langley, 1974; O'Connell, 1997). Under uncertainty, different states of the world occur with some known probability. Under incomplete information,

### II.B Conceptual Framework: Information and Migration

In this section I present a conceptual framework to show how information frictions affect migration behavior. Although I do not estimate the structural parameters of this model, the intuition helps inform the empirical strategy. In the canonical migration choice model, the decision to move is an investment in human capital (Sjaastad, 1962). An individual will move if the lifetime utility derived from moving minus the fixed costs of moving exceeds the utility of staying at the original location. The individual observes the fixed utility cost $c_{od}$ associated with migrating from $o$ to $d$ as well as the real returns $y_d(t)$ and $y_o(t)$ for each period in each location, which are defined to account for earnings, cost of living, local amenities, and idiosyncratic fit. The location specific returns can vary over time, and are discounted by $\beta$. Assume the individual is risk averse and has monotone preferences (the utility function is strictly increasing and concave). The decision to move from $o$ to $d$ ($m_{od}$) is characterized as follows

$$m_{od} = \begin{cases} 1 & if \quad \sum_{t=0}^{T} \beta^t u(y_d(t)) - c_{od} \geq \sum_{t=0}^{T} \beta^t u(y_o(t)) \\ 0 & else \end{cases} \tag{1}$$

But individuals likely face uncertainty about conditions in the potential destination such that $y_d(t)$ is a random variable, where $y_d(t) \sim G(y;\theta)$.[5] The individual will thus decide to migrate if

$$\sum_{t=0}^{T} \beta^t \big( Eu(y_d(t)) - u(y_o(t)) \big) - c_{od} \geq 0 \tag{2}$$

where the $E$ operator is the expected value at time zero. Changes in the parameters $\theta$ will

potential destinations, possible states of the world, and the true probabilities are potentially unobserved.

[5]The model implications are similar if the individual is also uncertain about conditions at the origin.

affect the outcome of this decision. For example, define

$$c^*_{od} = \sum_{t=0}^{T} \beta^t \big( Eu(y_d(t)) - u(y_o(t)) \big). \tag{3}$$

The value $c^*_{od}$ is the threshold moving cost at which the individual is indifferent between staying and moving. If $y_d(t)$ is distributed normally with a mean ($\mu_d$) and variance ($\sigma^2$), the nature of $u$ implies that

$$\frac{\partial c^*_{od}}{\partial \mu_d} = \sum_{t=0}^{T} \beta^t \frac{\partial Eu(y_d(t))}{\partial \mu_d} > 0 \text{ and } \frac{\partial c^*_{od}}{\partial \sigma^2} = \sum_{t=0}^{T} \beta^t \frac{\partial Eu(y_d(t))}{\partial \sigma^2} \leq 0. \tag{4}$$

Intuitively, as the mean increases, less weight is placed on low values of $y_d$ and expected utility rises. This increases the threshold moving cost, and the individual will be willing to pay a larger cost to move. An increase in the variance, holding all else equal, represents a mean preserving spread which results in weakly lower expected utility because the individual is risk averse (Rothschild & Stiglitz, 1970).[6] The increase in variance increases risk, and the individual's moving cost threshold becomes smaller, as she must be compensated by a lower cost to move.

People might have incomplete information about the parameters that govern the distribution of $y_d(t)$, and this additional uncertainty will also affect migration decisions.[7] For example, if the individual's prior belief is that the return to migration is low, she will be less willing to move. Similarly, if her prior is diffuse and the investment in migration appears more risky, she will also be less willing to move. Receiving additional information over time

[6]If the distribution of $y_d(t)$ is governed by more than just locational parameters this is not necessarily true (Tobin, 1965; Dionne & Harrington, 1991). More generally, if $\hat{\sigma}^2$ is a mean preserving spread of $\sigma^2$, then Rothschild and Stiglitz (1970) prove that $Eu(y_d(t); \sigma^2) \geq Eu(y_d(t); \hat{\sigma}^2)$. If instead the utility is linear and individual is risk neutral, changes in the dispersion that preserve the mean will not affect the cost threshold.

[7]This type of uncertainty is prevalent. Even among highly educated medical students in the residency match process there is substantial heterogeneity in their ability to accurately predict the expected cost of living and earnings rank in their top two ranked locations (Bottan & Perez-Truglia, 2017).

can change migration outcomes as individuals update their beliefs about these parameters. Specifically, exposure to news stories that credit fracking with creating local booms, fueling local economic growth, or raising wages in potential destinations might change people's perceptions of the distribution of the returns in the fracking destinations mentioned; even negative news can provide information about where fracking is occuring and change people's beliefs.[8] For example, individuals exposed to numerous newspaper articles and TV news broadcasts touting the local economic benefits of fracking in Texas might adjust their original mean or dispersion beliefs about the returns to migrating to a Texas fracking county. This news information does not necessarily need to be correct, as long as the individual believes it contains truthful information.

When an individual receives information in the news about fracking in a specific destination, she can update her prior beliefs following a process like Bayes' Rule. Under Bayesian updating, sample moments from the new information are used to update posterior beliefs. Although individuals might not perform exact Bayesian updating to incorporate new information, using sample moments from information in the news seems reasonable at an intuitive level.[9] Observing a large sample mean (news that the return is high) will increase the posterior belief about the mean, but the magnitude of this increase will depend on how precise or diffuse the prior belief is. Observing more information provides a larger sample and reduces uncertainty about the parameters. However, the marginal impact of information becomes smaller as she gets more information and her beliefs converge to the true distribution.

As the individual incorporates new information about the parameters, she better understands the distribution of $y_d(t)$ and can compute the likelihood of observing the return $y$

[8]Up through 2012, the last year of my sample, about 60 percent of adults were familiar with fracking, and over half of this population was in favor of fracking (Pew Research, 2013a). For someone that views fracking favorably, even a negative news story could provide information about where fracking is occurring, and result in updated beliefs.

[9]Wiswall and Zafar (2015) show that when college students receive information about the distribution of earnings, they update their beliefs, but often do not strictly follow a Bayesian updating process.

given her information set. The effect of additional information on the perceived mean of $y_d(t)$ will depend on her prior beliefs. If she initially believed the average return in a potential destination was lower than the news suggested (a likely case given the coverage about fracking jobs and booms), the information will increase her perception of $\mu_d$. This in turn increases $c^*_{od}$, meaning she is more willing to move (see equation (4)). The converse is also true. Similarly, receiving more information will reduce uncertainty, which makes moving less risky all else equal. If the individual is uncertain about the distribution of potential returns in a destination, receiving information in the news that portrays $y_d(t)$ as larger than she initially believed will increase her propensity to migrate.[10] Although I do not estimate the structural parameters of this individual level model, the predictions provide motivation for how geographic differences in news exposure might affect migration behavior.

## III Data Sources

Fracking provides a unique setting to explore the impact of news exposure on migration outcomes. Fracking began quite suddenly in the mid-2000s and by 2012 had affected oil and gas production in 252 counties in 16 states. These local fracking booms increased economic activity and improved labor markets in those counties (Feyrer et al., 2017). States in all

[10]Appendix Figure A1 presents results from a simulation that illustrates how new information provision can change migration decisions. Two scenarios are presented for three types of people. Individual 1 has a diffuse prior over the expected return ($\mu_d$) and incorrectly believes $\mu_d$ is lower than the true mean. Individual 2 has a precise belief that $\mu_d$ is low. Individual 3 has a diffuse prior, but correctly predicts $\mu_d$. In both scenarios the true parameters are the same, the only difference is that individual are exposed to more information in scenario 1 than in scenario 2. The perceived distribution of both $\mu_d$ and $y_d$ are plotted for each individual in each scenario over two iterations of receiving more "news". If initial beliefs about the expected return are low, new information shifts up the beliefs about $\mu_d$ and $y_d$. Additional information also reduces uncertainty about $\mu_d$ and $y_d$, which increases expected utility and the probability of migrating. Updating is more drastic when there is more information, and changes in the probability of moving will be the largest among people or areas that are exposed to more new information. Initial draws of information are very beneficial, but the marginal value of additional information becomes smaller.

four census regions have been affected and many people were unaware of exactly where these fracking booms were occurring. Both positive and negative aspects of fracking have been highly publicized through newspapers and TV news, and many of these news stories reference specific locations affected by fracking. Because fracking is a novel term, I am able to parse news content to identify which sources discuss fracking, which places they talk about, and what aspects of fracking were discussed. By linking this with measures of news penetration, I am able to estimate how geographic differences in exposure to news about fracking affect migration flows. This estimation requires detailed data on fracking production, migration flows, news content, and news circulation. In this section I briefly describe each data source and highlight key strengths and limitations with a full description in the online data appendix.

*Fracking Production Data.* Oil and gas extraction data is obtained through a proprietary data agreement with the private company DrillingInfo. DrillingInfo provides well-level information on drilling and quarterly production. The two technologies that characterize fracking, horizontal drilling and hydraulic fracturing, have drastically increased the productivity of thick layers of dense shale rock, known as shale plays. Using the well characteristics, I define a fracking county as any county with positive oil or gas extraction from a non-vertical well in a drilling formation that corresponds to a shale play. This amounts to fracking regions in 16 states: Arkansas, California, Colorado, Louisiana, Michigan, Mississippi, Montana, New Mexico, North Dakota, Ohio, Oklahoma, Pennsylvania, Texas, Utah, West Virginia, and Wyoming as seen in Figure I.

*Migration Data.* Migration data is obtained from the Internal Revenue Service (IRS) Statistics of Income (SOI). Using tax documentation, such as Tax Form 1040, the IRS tracks the number of households that filed their taxes in one county in one year and in a different county the next year. Most filing occurs between February and April, so annual migration flows capture moves from approximately March or April from one year to the

next.[11] This measure is then aggregated up to the county to determine the approximate flow of households (returns) and individuals (tax exemptions) from one county to another. These two numbers are provided for pairs of counties in the United States, but is censored for county pairs with fewer than 10 returns for privacy purposes. In 2013, the censoring threshold increased from 10 to 20 returns, leading to much higher levels of suppression. For this reason I restrict my analysis to migration between 2000 and 2012. The IRS data only provide a raw count and do not provide information about demographic characteristics. Although this is perhaps the most comprehensive data on internal migration in the United States, it might under-represent a subset of the extremely poor (who fall below mandatory tax filing thresholds and do not file for other benefits such as the Earned Income Tax Credit) as well as a small subset of the extremely wealthy (who are more likely to be granted filing extensions for complex returns).

*Newspaper Circulation Data.* Proprietary newspaper readership data is obtained from the Alliance for Audited Media (AAM). The AAM conducts regular newspaper circulation audits for national, regional, and most local newspapers in the United States. This includes the number of copies sold on the audit date and the number of copies as a percent of households for each county with over 25 copies. Counties with fewer than 25 copies sold are assigned a zero value. For some newspapers, these measures are only available at the Designated Market Area (DMA) level. Historic circulation rates from 2005 through 2008 are scraped from pdf files.

*TV Viewership Data.* TV viewership data is calculated from the Television and Cable Factbook using Nielsen viewership data. For my analysis, I use viewership rates from both the 2008 and 2016 Factbook. Between 2007 and 2009, TV stations were transitioning from analog to digitally transmitted broadcasts on a market-by-market basis. When a market transitioned, viewers were required to obtain digital reception equipment, and it is unclear

[11] For example, migrants who moved between March/April of 2011 and March/April 2012 will be assigned the year 2011. This introduces a slight lag relative to the measurement of news (from January to December).

how this affected viewership and if 2008 viewership is representative of later years.[12] For this reason I also examine the most recent viewership rates from 2016. TV viewership is reported at the DMA level for each TV station and includes viewership from both cable and non-cable households. These data are available at the station-level and are not specific to news programming. The viewership rate is constructed by dividing total weekly viewership by the total number of households in the DMA.

*Newspaper Content Data.* Newspaper content is obtained through the LexisNexis database, which provides access to articles from over 2,600 news sources, including *USA TODAY*, the *New York Times*, and the *Wall Street Journal*. First, I preserve all articles since 1999 that include any of the search terms "frack∼", "fracing", or "hydraulic fractur∼" anywhere in the text. I then linguistically parse each article to exclude spurious keyword references such as "frick and frack", unrelated acronyms, and the last names of people. Most of my analysis is restricted to three national news sources: *USA TODAY*, the *New York Times*, and the *Wall Street Journal*.[13] The in depth news coverage of fracking begins in 2009, and dramatically increases each year. In these three newspapers there were 562 news articles related to fracking between 1999 and 2012. The first two articles in the national news were in 2002 and 2003 in the *New York Times*, which briefly reference court cases about patents related to hydraulic fracturing. There was then one article in 2006, five in 2008, 20 in 2009, 48 in 2010, 198 in 2011, and 288 in 2012. Next, I linguistically parse the entire text of each of these articles to determine which of the 16 fracking states listed above each article discusses.[14] I also parse each article for specific keywords such as "growth", "boom", "contaminat∼", and "earthquak∼" to determine the positive and negative content of each article (discussed in

[12]A special thanks to Matt Long from Warren Communication News for finding out how the viewership rates for the 2008 Factbook were constructed, and to Colin Wick for transcribing viewership rates from the 2008 Factbook.

[13]When including local news coverage, I restrict the sample to news articles from domestic US newspapers.

[14]Not every article mentions a specific state. I have also parsed each article for city names from the U.S. Postal Service's registry of city names, but find that local jurisdictions are referenced far less frequently.

detail later). These statistics are reported in Appendix Table A1. Articles that mention specific states are more likely to refer to things like jobs, booms, and growth and there is also heterogeneity across states in how frequently these "positive" effects of fracking as well as "negative" effects such as pollution, danger, and earthquakes are cited.

*TV News Content Data.* TV news content is obtained from the Vanderbilt Television News Archive (VTNA). The VTNA database contains TV news recordings and transcript abstracts for nightly news broadcasts from the three major news networks (ABC, CBS, and NBC) and the cable news channels CNN and Fox News. The database only includes one hour of programming each day for both cable news outlets. Because the available content of cable news is limited, and viewership rates are only available for the TV networks, I restrict the sample to TV broadcasts from the three major news networks. I parse the transcript abstracts for search terms such as "fracking" and "shale" as well as which state is being discussed. Between 1999 and 2012 there is far less coverage of fracking on the nightly news than in the newspaper. The VTNA database only records 17 news broadcasts, with one in 2006, two in 2008, three in 2010, four in 2011, and seven in 2012.

*Cross-County Commute Data.* In addition to migration, I also explore impacts on workers who live in one county but work in another. This captures both long distance commuting and temporary relocation, such as moving to the job site for several weeks at a time but maintain the same permanent address. This data is available through the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES). This data provides statistics on the number of jobs for each home and work census block pair. I aggregate up these pairs to the county level to determine how many workers live in one county but work in another. This data is available for all years since 2002, and also provides statistics by broad age groups (under 30, 30-54, over 54), monthly earnings (under $1,250, $1,250-3,333, over $3,333), and industry (goods, trade/transportation, other). This allows me to explore heterogeneous commute responses across different groups.

*County Characteristics Data.* County level economic and population characteristics are

obtained from a range of sources. Economic outcomes such as employment to population ratios, unemployment rates, and average earnings are obtained from the Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW). County-level age and racial demographics are obtained from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program population data and are constructed from the U.S. Census Bureau's Population Estimates Program. Other county level characteristics are obtained from the 2000 Census and ACS through the American Factfinder.

# IV Empirical Strategy

To estimate the impact of labor market news on migration behavior, I exploit county-level differences in exposure to news about fracking in a specific destination state. To capture this variation, consider the following measure

$$newspaper\ exposure_{oSt} = \sum_{n \in N} \Big(\text{total articles in } n \text{ about fracking in } S\Big)_t * circ.\ rate_{ont} \quad (5)$$

where $N$ is the set of all domestic newspapers. $Newspaper\ exposure_{oSt}$ is the weighted sum of news articles that mention fracking in destination state $S$ in year $t$, where each newspaper is weighted by its circulation rate in the origin county $o$ (ranging from zero to one). News exposure varies at both the origin and destination level, and increases when there are more news articles, but the magnitude of this increase is weighted by circulation rates. To identify the causal impact of labor market news exposure on migration, consider the hypothetical relationship between newspaper exposure and $Y_{oSt}$, a measure of migration flows from origin county $o$ to fracking counties in destination state $S$ in year $t$:

$$Y_{oSt} = f(newspaper\ exposure_{oSt}; \theta) + \phi_{oS} + \eta_{oSt}. \quad (6)$$

Migration flows are composed of three parts. The first part is a flexible function of newspaper exposure. The second part is an origin/destination pair specific level effect that accounts for time invariant origin/destination specific characteristics, such as distance or industry ties. The third part is an error term, $\eta_{oSt}$, composed of two systematic components that vary over time: destination state characteristics ($\psi_{St}$) and origin county characteristics ($\lambda_{ot}$), as well as an idiosyncratic origin-by-destination pair specific error term ($\nu_{oSt}$) as follows

$$\eta_{oSt} = \psi_{St} + \lambda_{ot} + \nu_{oSt}. \tag{7}$$

If *newspaper exposure*$_{oSt}$ is correlated with destination specific, origin specific, or pair specific characteristics that are changing over time, the estimated effect of *newspaper exposure*$_{oSt}$ will be a biased estimate of the causal effect of labor market news on migration.

Consider the following thought experiment: if a set of origin counties are randomly assigned different levels of exposure to news about fracking in Texas (or any other fracking state), do counties that were more exposed to this information see larger increases in migration to fracking counties in Texas? By random assignment, news exposure will be uncorrelated with the unobserved origin component ($\lambda_{ot}$) and the origin/destination idiosyncratic error ($\nu_{oSt}$). By comparing migration flows to the same destination, everything about fracking counties in the destination state is held constant, allowing the effect of news exposure on migration to be isolated. The relationship over multiple destinations can be explored in a regression framework by stacking the estimation as follows

$$Y_{oSt} = f(newspaper\ exposure_{oSt}; \theta) + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \tag{8}$$

The level of observation is the annual migration flow from an origin county to all fracking counties in a destination state. An origin/destination pair fixed effect, $\phi_{oS}$, is included to control for time invariant characteristics of the pair that affect migration, like distance. A destination state-by-year fixed effect, $\psi_{St}$, is also included to control for destination specific

characteristics that are changing over time, and makes this a comparison of migration flows to the same destination state from origin counties that have different levels of news exposure. Importantly, this fixed effect captures changes in fracking production, labor market characteristics, and amenities which might directly affect migration behavior and lead to higher news exposure.

Unlike the thought experiment, actual exposure to news about fracking in a specific region is not exogenously assigned and endogenous to decisions of both news providers and consumers. This endogenous variation in news exposure might be correlated with both origin ($\lambda_{ot}$) and origin/destination ($\nu_{oSt}$) characteristics and could lead to biased estimates of the causal impact of news exposure on migration in equation (8). To obtain exogenous variation in news exposure, two major concerns about endogeneity must be addressed.

The first concern is the endogeneity of news content arising from news providers' decisions. For example, if people from Franklin County, Ohio start moving to Alleghany County, Pennsylvania, the local *Columbus Dispatch* might produce more content about fracking in Pennsylvania, raising concerns about reverse causality. This concern seems particularly relevant for local newspapers, where content decisions strongly respond to consumer preferences in local geographic markets (Gentzkow & Shapiro, 2010). However, large, national newspapers such as *USA TODAY*, the *New York Times*, and the *Wall Street Journal* do not have a well-defined geographic market and operate differently (Gentzkow & Shapiro, 2010).[15] These newspapers are read throughout the country and make content decisions to cater to the nation as a whole. While content decisions of the *Dearborn County Register* might be endogenous to the number of people moving from Dearborn County, Indiana to fracking counties in Ohio, content decisions of the *New York Times* are likely driven by aggregate trends rather than idiosyncratic patterns. Counties across the country are exposed to the

[15]Gentzkow and Shapiro (2010) also list the *Christian Science Monitor* as a national newspaper, however, circulation for this newspaper is only available at the state level, so it is excluded from all analysis.

same national news, regardless of their idiosyncratic deviation from the national trend.[16]

Although the same national news is available across the country, counties will be differentially exposed to newspaper articles about fracking because they vary in their reading habits (circulation rates). This raises the second concern of endogenous news consumers' decisions. Over time, residents of Franklin County, Ohio might come to view fracking more positively (or negatively), which might affect both their readership of the *New York Times* and migration to fracking destinations, leading to omitted variable bias. However, because fracking began quite suddenly there is a clear pre-period when circulation was not a function of preferences toward fracking. The largest expansions in oil and gas production due to fracking started after 2008 and national news about fracking only began in earnest in 2009. By using pre-2009 circulation rates, I can isolate variation in exposure due to pre-fracking differences in circulation rather than changes in exposure that arise from changing preferences toward fracking.

Exploiting variation in news exposure due to national news content and pre-fracking circulation can mitigate concerns about endogenous decisions of both news producers and news consumers in this modified definition of $newspaper\ exposure_{oSt}$ originally presented in equation (5)

$$newspaper\ exposure_{oSt} = \sum_{n \in N'} \Big(\text{total articles in } n \text{ about fracking in } S\Big)_t * Pre09\ circ.\ rate_{on}. \quad (9)$$

The set of newspapers is now restricted to national news sources, $N' = \{$*USA TODAY, New York Times, Wall Street Journal*$\}$, and the number of articles is weighted by the pre-2009 circulation rate, which is the average circulation between 2005 and 2008. As such, an

[16]National newspapers might report more about destinations that see large changes in labor markets or migration (nationwide trends). However, including destination by year specific effects compares migration flows from different origins to the same destination, eliminating destination specific differences that might drive news coverage. It could be argued that readers in and around New York City have a large effect on the content decisions of the *New York Times*. As a precaution, I exclude counties in the New York City DMA from the analysis. In Column (1) of Appendix Table A5 I show that the migration response is still significant if the New York City DMA is included.

additional news article in a national newspaper will increase newspaper exposure, but this increase will be largest in counties that had high readership prior to the fracking boom. This strategy is similar to previous work using variation in circulation exposure to explore the impact of media and news on other outcomes.[17]

When using this new definition of newspaper exposure to estimate equation (8), the effect of national news exposure is identified by variation across origin counties in pre-fracking circulation rates. This variation is potentially problematic if pre-fracking circulation is correlated with changes over time in other local characteristics that affect preferences to move to fracking, captured in $\lambda_{ot}$. For example, if baseline readership of the *USA TODAY* was higher in more affluent counties and the distribution of income became more dispersed over time, this differential increase in income would introduce omitted variables bias if it is correlated with both baseline circulation and residents' propensity to migrate to fracking. Because fracking was such a new, unknown technology, it is not always clear what direction these characteristics might bias the estimates.[18] As seen in Figure II, there is no obvious, strong geographic correlations in pre-2009 circulation of the *USA TODAY* and even within close regions there is significant variation. As seen in Table I, areas with low and high circulation of the *USA TODAY* were similar in 2000 on average.

Counties with below median circulation of the *USA TODAY* had slightly lower employment, lower median income, higher poverty, and an older population. If any of these characteristics that are correlated with circulation affect migration or moving costs in a time

[17]For example, Gentzkow (2006) examines TV introduction on voter turnout, DellaVigna and Kaplan (2007) examine Fox News introduction on Republican vote shares, Jensen and Oster (2009) examine Indian cable introduction on women's status, Chong and La Ferrara (2007) and La Ferrara et al. (2012) examine Brazilian soap opera introduction on divorce and fertility, Garthwaite and Moore (2012) examine exposure to Oprah Winfrey content on votes for Barack Obama after her endorsement, Kearney and Levine (2015a; 2015b) examine exposure to the MTV series "16 and Pregnant" on teen births, and exposure to Sesame Street on grade completion.

[18]For example, even the prominent environmental organization, the Sierra Club, went from supporting fracking and natural gas extraction (as a cleaner alternative to coal) to later condemning it (Gold, 2014).

invariant way, it will be controlled for by the origin county by destination state fixed effects. Of more concern to causal identification are changes over time that are correlated with pre-fracking circulation. Columns (3) and (4) in Table I suggest that between 2000 and 2010 both low and high circulation counties followed parallel trends in migration. Other origin county characteristics evolved similarly in low and high circulation counties, although high circulation counties saw relatively larger decreases in employment and the percent white, and larger increases in unemployment, median income, and poverty, slightly closing the gap that existed in 2000 between low and high circulation counties. However, these differences are never more than one or two percentage points.

Column (5) formally tests if pre-2009 *USA TODAY* circulation rates predict differential changes in origin county characteristics between 2000 and 2010. Pre-2009 circulation rates do not predict changes in migration or the employment to population ratio, but some of the other local characteristics are statistically different. However, these differences are quite small: an increase in readership from the 25th to the 75th percentile of *USA TODAY* circulation (1.88 percentage points) predict a 0.09 percentage point reduction in the unemployment rate, a $290 increase in median income, a 0.51 percentage point increase in the poverty rate, a 0.17, 0.29, and 0.41 percentage point increase in the percent black, Hispanic and other race respectively, and a 0.27 percentage point decrease in the population 35 to 64. *New York Times* and *Wall Street Journal* pre-2009 circulation rates predict similarly small changes in county characteristics (Appendix Table A2).[19] These differential trends in origin county characteristics are small and unlikely to have large effects, but it is still possible that they, or other unobserved characteristics of the origin, might bias the estimated effect of news exposure. However, as explained below, I can also include origin county by year fixed effects to control for characteristics of the origin that affect migration and are changing over time.

[19]Readership of the *New York Times* and the *Wall Street Journal* are highly correlated, and the predicted effects are similar. The one characteristic that varies the most across newspapers is median household income. The *New York Times* and *Wall Street Journal* have higher readership in large urban areas that saw larger increases in earnings.

To evaluate if these potential threats to identification are valid concerns, I estimate the effect of national news exposure on migration in a series of progressively more conservative regressions. I first estimate the model corresponding to the thought experiment as follows

$$Y_{oSt} = \beta_1 newspaper\ exposure_{oSt} + \beta_2 newspaper\ exposure^2_{oSt} + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \quad (10)$$

The main outcome of interest is the inverse hyperbolic sine of the number of migrants from origin county $o$ to fracking counties in state $S$ in year $t$. The inverse hyperbolic sine approximates a natural log transformation but is defined for flows with zero migrants, allowing me to measure the percent effect of news exposure. Origin county by destination state fixed effects are included to control for time-invariant pair specific characteristics, and destination state by year fixed effects are included to account for changing characteristics of the fracking destinations. To account for correlated shocks across geography, the standard errors are adjusted for clustering at the origin DMA level (203 clusters), a geographic measure meant to capture media markets. Observations are equally weighted.[20] As suggested by the theoretical framework, I include news exposure quadratically to capture decreasing marginal returns to information, although the relationship is robust to different functional forms. I begin with the specification in equation (10) because the identifying variation is highly transparent: origin counties experience different exposure to news about a specific destination because they have different pre-fracking circulation of national newspapers.

I then progressively adjust this baseline specification to address potential concerns associated with this identifying variation. First, I include a vector of time varying origin county labor market controls, including the employment to population ratio, the unemployment rate, and average earnings (in 2010$). These controls account for observable changes in the origin labor market that might be correlated with news exposure and affect migration. Second, I include origin county by year fixed effects which account for both observed and

[20]If I instead weight by the origin county population in 2000, the impact from equation (10) is about 1.5 times as large and significant.

unobserved components of $\lambda_{ot}$. This is possible because I observe migration flows to 16 different fracking states from each origin county/year pair. Including origin county by year fixed effects will control for changing characteristics of the origin county that affect all of these migration flows. For example, if counties with higher circulation rates, and thus higher newspaper exposure, are changing over time in ways that reduce migration costs (e.g., becoming younger, more educated, or more wealth), these omitted variables might affect decisions to move to fracking areas in general. Origin county by year fixed effects will absorb these and other changes over time and exploit variation in news exposure across potential destinations from the same origin. This specification tests to see if, for a given origin county, destination states that had more news exposure also experienced larger increases in migration flows. In this specification, confounding omitted variables must be origin/destination pair specific and vary over time (contained in $\nu_{oSt}$).

For example, if local and national news exposure are strongly correlated and local news is endogenous to migration preferences, omitting local news will bias the coefficient on national newspaper exposure.[21] I combine the content of all domestic newspapers available through LexisNexis with circulation rates, and construct an analogous measure of local newspaper exposure to test and see if local newspaper exposure changes the estimated effect of national newspaper exposure.[22] I also conduct placebo tests and alternative estimation strategies designed to test if the observed relationship is actually driven by unobserved origin/destination specific changes over time.

In these specifications, the sample is restricted to exclude origin counties involved in fracking as information in the news might effect the decisions of people originally living in

[21]The actual correlation between national and local newspaper exposure is 0.12.

[22]Many local news sources provide free access to content online, which is not captured by this measure of local news exposure. National and regional news sources often provided limited free access, but ultimately require a paid subscription. The AAM circulation data includes digital replica newspapers, but not necessarily individual browsing behavior. To the extent that online exposure is positively correlated with print exposure, the estimates will simply represent the response to total news exposure (where print exposure is used as a proxy).

fracking counties differently. For example, residents of Bradford County, a fracking county in Pennsylvania, are likely affected differently by news about fracking in Pennsylvania than residents of non-fracking Adams County, Pennsylvania. However non-fracking origin counties in states with fracking are still included.[23]

As *newspaper exposure*$_{oSt}$ is a weighted sum, it is not immediate how to interpret a one unit increase in this measure. From equation (9), if the pre-2009 county circulation rate is one (meaning every household receives the newspaper) an additional news article will increase newspaper exposure by one unit. In reality, newspaper circulation rates are significantly lower than one hundred percent. To make exposure more readily interpretable, I divide *newspaper exposure*$_{oSt}$ by 0.05, such that a one unit increase is equivalent to one additional news article in a newspaper with a five percent circulation rate.[24] This level of circulation is comparable to a county with high readership of *USA TODAY*.[25] Conveniently, when using this scaling average news exposure among treated observations is 0.99, suggesting a one unit increase also approximates the mean effect.

# V Main Results

## V.A Graphical Analysis

Before estimating the regression in equation (10), I explore pre-trends and present event study graphical evidence of the impact of national newspaper exposure on migration. This specification can verify that origin counties that will eventually be highly exposed to news

[23] As noted earlier, origin counties in the New York City DMA are also excluded. Relaxing these sample restrictions do not significantly impact the results (see Appendix Table A5).

[24] This scaling is similar to a continuous versions of the persuasion rate (DellaVigna & Kaplan, 2007). The effect of one additional news article is scaled by the exposed population and in this case, the population available to persuade is approximately one. However, as I do not observe individual exposure, I cannot construct a direct comparison.

[25] For reference, *USA TODAY* circulation ranges from 0 to 27.8 percent, with a mean of 1.2 percent; *New York Times* circulation ranges from 0 to 3.3 percent, with a mean of 0.51 percent; and the *Wall Street Journal* circulation ranges from 0 to 6.4 percent, with a mean of 1.2 percent.

do not have higher migration flows in years prior to actual exposure to this news, relative to origins that will be less exposed. There are various ways to measure this exposure "treatment", but I will focus on differences in exposure due to initial differences in circulation rates of national newspapers that will eventually report on fracking.[26] This tests to see if origin/destination specific news exposure is correlated with other unobserved characteristics that evolve over time and affect migration ($\nu_{oSt}$). For each origin county I collapse the pre-fracking circulation rates of the *USA TODAY*, *New York Times*, and *Wall Street Journal* to a single weighted average, where the weights are the share of the total national news articles about fracking in destination $S$ in each newspaper. This measure captures the extent to which an origin will eventually be exposed to news about fracking in the destination state. I interact this measure with year indicators between 2001 and 2012 (omitting 2000 as the reference year), and then regress the inverse hyperbolic sine of the number of migrants on this set of interactions to trace out the impact of pre-fracking circulation on migration over time as follows:

$$Y_{oSt} = \sum_{\tau=2001}^{2012} \theta_\tau Circulation_{oS} * 1\{t = \tau\} + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \tag{11}$$

I include origin-destination pair fixed effects as well as destination-by-year fixed effects to exploit the same variation used in the main analysis. The coefficients on these year interactions are interpreted as the marginal effect of a one percentage point increase in the pre-fracking circulation rate on migration flows in that given year, and are plotted with 95 percent confidence intervals in Figure III. For reference, a bar graph of the average number of articles about fracking in a specific state is superimposed, to show when news content about fracking was published.

Between 2000 and 2007, specific destination states were only mentioned in the one 2006

[26]This measure is ideal for testing that different levels of initial circulation do not follow differential trends. The figure is almost identical when looking at alternative measures of treatment, such as the total newspaper exposure summed over all years.

*New York Times* article, otherwise, there were no national newspaper articles referencing fracking destinations. There are small increases in the number of articles about fracking between 2008 and 2010, with a large jump in 2011 and 2012. Prior to 2010, migration fluctuates around zero, with only one statistically significant, negative estimate in 2003. Starting in 2006 there is a slight, statistically insignificant upward trend, but overall it appears that origins that would eventually be highly exposed to news about fracking followed similar trends in migration. Since treatment also starts during this time, it is not clear this is evidence of non-parallel pre-trends.[27] In 2010 the effect on migration becomes significant, and discontinuously jumps in 2011, when news content increased dramatically. The data suggest that a one percentage point increase in the pre-fracking circulation rate did not increase migration prior to news exposure, but was associated with a 2.5 percent increase in migration in 2011 and 2012, precisely when there was intense news coverage of fracking.[28]

### V.B Impact of Newspaper Exposure on Migration

I now formalize this relationship by estimating the regression in equation (10). These estimates are provided in Column (1) of Table II. Given the absence of news in early years, I interpret effects as changes from zero to one. For an origin county with a five percent circulation rate, one additional newspaper article about fracking in a specific state increased migration flows to fracking counties in that state by 2.4 percent on average $(0.025 * 1 - 0.001 * 1)$.[29] As average news exposure is also approximately one, this would suggest the mean effect of news exposure on migration was 2.4 percent as well. The most news was published in 2012 (average news exposure was 1.8) suggesting that in 2012, news about fracking increased

[27] This pattern could also be consistent with origin destination pairs with high circulation following an upward trend that was suppressed during the Great Recession (2007-2010), only to rebound in 2011. This pattern is essentially unchanged if I include origin county labor market controls.

[28] As seen in Appendix Figure A2, commuting responds similarly, although the increase is larger (8-12 percent) and begins earlier in 2009.

[29] These estimates are not just statistically significant due to a large sample. As seen later, the significance remains when estimated over much smaller subsamples.

migration flows to fracking counties by 4.2 percent on average.

I next adjust the baseline specification as outlined above to determine if changing characteristics of the origin bias the estimates. In Column (2) I include the annual origin county-level labor market measures, and the estimated impact of national newspaper exposure is virtually unchanged at 2.4 percent. In Column (3) I include origin by year fixed effects. This absorbs the labor market measures included in Column (2) as well as any other unobserved characteristic of the origin that is changing over time and affects migration behavior. In this specification the effect of one additional newspaper article is 2.5 percent, and not statistically different from the baseline estimates. Finally, in Column (4) I include the origin by year fixed effects and control for local newspaper exposure. The effect of one additional national newspaper article remains 2.4 percent. For completeness, I repeat the same estimation using the number of migrants in levels as the outcome. In each of these specifications the marginal impact ranges from 1.4 to 1.7 and is not statistically distinguishable.[30] For the remainder of the paper, I will estimate the model corresponding to Column (2), which includes controls for labor market conditions at the origin, although the results are not sensitive to this choice of specification.

The data suggest that increased national news coverage of fracking increased migration to the fracking counties in states publicized. Although these estimated impacts are small, they are both statistically and economically significant. These estimates imply that news about fracking increased migration flows to fracking counties by 2.4 percent on average, and that the 36 articles in the *New York Times* in 2011 that discussed fracking in Pennsylvania led to an 8.2 percent average increase in migration flows to Pennsylvania fracking counties. When considering the levels specification, exposure to news about fracking in a particular state led to one to two additional migrants from each origin on average. For a given origin this

[30]An increase of 1.4 migrants represents a much larger effect at the mean than captured by the inverse hyperbolic sine specification. This appears to be driven by origin counties with large migrant flows. If the sample is restricted to origin/destination pairs with non-zero flows, the two specifications yield similar percent effects at the mean.

effect is small, but becomes large when aggregated up for a given destination. This would suggest that, at the margin, relaxing information constraints and providing information about potentially lucrative labor market opportunities elsewhere will increase migration to those destinations.[31]

### V.C Impact of Newspaper Exposure on Cross-County Commuting

To avoid the monetary, psychic, and amenity costs that might accompany a move, an individual can choose to commute rather than migrate. In a companion paper, I show that many people took advantage of the earnings gains associated with fracking by taking up jobs in fracking counties and commuting, rather than migrating (Wilson, 2017). Information revealed through newspaper exposure might also affect aggregate behavior at this margin. In Table III I report the impact of newspaper exposure on the total number of workers who live in one county but work in a fracking county in the state mentioned in the newspaper article. In addition, I report the number of jobs for three pre-defined age groups: workers under 30, workers 30 to 54, and workers 55 and older. These data are obtained from the LEHD Origin-Destination Employment Statistics (LODES). For an origin county with a five percent circulation rate, one additional news article about fracking in a specific state increased the number of workers commuting to fracking counties in that state by approximately 6.6 percent. The impact on commuting is nearly three times as large as the migration response, which is not surprising as commuters avoid many of the fixed costs associated with moving.

[31]Unfortunately the IRS migration data does not provide demographic information. In theory, the microdata from the American Community Survey could be used (Ruggles et al., 2015). However, because the annual ACS is only a one percent sample of households, the probability of observing a mover to a specific state is very low, leading to measurement error and attenuation. Often these flows are constructed from less than five people. Using the 2005 through 2011 ACS, I have estimated the corresponding models at the PUMA level by gender, age, race, education, and family status. Consistent with attenuation bias due to measurement error, the estimates are about one tenth the size and insignificant. The largest group effects are for men (0.23 percent, p=0.13), 18-34 year olds (0.16 percent, p=0.19), non-Hispanic other (0.16 percent, p=0.15), and college graduates (0.15 percent, p=0.16).

When looking across age groups, the response for one additional news article for 30 to 54 year olds is 5.2 percent and statistically larger than the response of both younger workers (3.1 percent) and older workers (3.6 percent).[32] These differential responses among workers under 30 and between 30 and 54 also correspond to patterns of newspaper readership. Throughout the 2000s, 30 to 49 year olds were more likely to report the newspaper as a main source for national news than 18 to 29 year olds (Pew Research, 2013b). In Appendix Table A4, I also report differences by job level earnings and broad industry. Consistent with people commuting for fracking jobs that pay well, workers in higher paying jobs are more responsive, but there is also an increase in commuting among low paying jobs. Non-production or trade workers in the "other" industry, which includes oil and gas extraction, are the most responsive. However, workers in goods production and trade and transportation also significantly respond, suggesting that increased news exposure not only induced people to commute for oil and gas extraction, but for other jobs that were also affected by the labor market shock.

The effects of newspaper exposure on both migration and commuting are robust to different ordered polynomials of newspaper exposure (see Appendix Table A4), excluding each destination state one at a time, and insensitive to sample restrictions (Appendix Table A5).[33] The estimated migration relationship is also robust when I account for censoring in the IRS migration data (Appendix Table A6).[34]

[32]The reader will notice that the percentage effect is larger for all workers than for any of the three subgroups. This is in part because the pooled specification constrains the controls and fixed effects to be the same for each group. If this specification is run in levels, the effect for all workers is the sum of the effects for each subgroup, as expected.

[33]There are only two years in the end of the sample with high levels of news exposure, making dynamic effects less relevant. In specifications including one and two year lagged exposure, the effect of concurrent newspaper exposure on migration is nearly identical and the first lagged effect is small and marginally significant. For commuting the effect of concurrent exposure is slightly smaller, but not statistically different, and the lags have small, significant effects.

[34]The effect of newspaper exposure is similar in the extreme lower bound case where all zero county-to-county flows are changed to 9 (the highest possible censored value). Newspaper exposure increases the probability of not being censored by over 15 percent (0.005/0.03) at the mean, as well

### V.D News about Fracking in Another State

News about fracking in a specific state can affect migration decisions by providing general information about the labor market shifts associated with fracking, or by providing specific information about where these labor market shifts are occurring.[35] To separate these channels of effects, I evaluate how migration flows to fracking counties in a particular destination state respond to news about fracking in a different state. For example, observing that the migration response to news about fracking in a different state is smaller than the response to news about fracking in the destination state would suggest the news provides some locational signal.

In practice, I randomly assign all observations indexed by $S$ the fracking news exposure of $S'$, one of the other 15 fracking states. For example, all observations for the destination Arkansas might be randomly assigned the news exposure of North Dakota, while the observations of North Dakota might be randomly assigned the news exposure of Pennsylvania. I then estimate the regression similar to equation (10), but replace *News Exposure*$_{oSt}$ with the randomly assigned *News Exposure*$_{oS't}$, and calculate the marginal impact of a one unit increase in *News Exposure*$_{oS't}$ (i.e., one news article in a county with a five percent circulation rate). I repeat this regression 200 times to plot the distribution of potential impacts. This

as the number of migrants for the severely limited subset of origin destination pairs that always report positive flows.

[35]This is related to the advertising literature, suggesting information could either lead to more migration overall (expansion) or shift people from other destinations (share stealing). In Appendix Table A7 I include the total exposure to news about any of the 16 destination states within an origin and year, to determine if news about fracking in general has an effect. This has no effect on migration, but a small, significant effect on commuting (0.5 percent). I also estimate a separate specification including the news exposure for the state that received the highest exposure within an origin and year. This effect is positive and areas with the most news did not reduce flows to other fracking destinations, suggesting news exposure led to expansion, rather than shifting. See for example, Garthwaite (2014) examining book sales after celebrity endorsements. It is also possible that migrants to fracking areas simply shifted from moving to other, non-fracking areas. However, when looking at the impact of total news exposure about fracking on migration to non-fracking areas, the coefficient is small, but positive.

histogram of potential impacts is plotted in Panel A of Figure IV and the estimated effect using actual news exposure from Table II is indicated for reference. Importantly, I do not consider this a placebo test, because general information can plausibly affect the outcome.

The estimated effect from Table II is larger than the estimates from all but 5 of the repetitions (2.5 percent), suggesting that at least part of the migration response is due to locational signaling. The distribution of effects using randomly assigned news exposure is centered around 0.013, suggesting news about fracking in a different state has some positive predictive power. However, this 1.3 percent effect cannot be strictly attributed to general information about fracking. National newspapers report about multiple destinations, and news exposure is positively correlated across potential destination states. Among the 200 regressions, the average correlation coefficient between actual news exposure and randomly assigned news exposure was 0.44. To some degree, randomly assigned news exposure will proxy for actual news exposure, which might drive the estimated 1.3 percent effect.

For this reason I adjust the regression specification to include origin county by year fixed effects, and repeat the process 200 times. This specification absorbs average changes in news exposure at the origin county level and exploits destination specific deviations from the origin average. This specification looks to see if, for example, an origin that had unusually high exposure to news about fracking in North Dakota saw larger increases in migration to fracking counties in Arkansas. Randomly assigned news exposure no longer proxies for actual news exposure, but provides a test to determine if destination specific fluctuations in news exposure impact the corresponding migration flows. If the effect size from these regressions were comparable to the estimates in Table II, then we would be concerned that the results are driven by things correlated with news exposure and pre-fracking circulation rates in general, not a causal effect of news content. This histogram of potential impacts is plotted in Panel B of Figure IV, along with the estimated effect from the origin county by year fixed effects specification from Table II. The distribution of effects are centered around zero and are all smaller than the estimated effect in Table II, suggesting that destination

specific deviations in news exposure largely affect migration flows to the destination that is being mentioned. The information about fracking conveyed in the news affects migration decisions primarily by providing information about where fracking is occurring.

## VI Alternative Strategy: Newspaper Market Border Comparison

Although origin level characteristics and local news exposure do not appear to introduce bias into the estimates in Table II, it is possible that news exposure is correlated with other unobserved origin destination pair characteristics captured by $\nu_{oSt}$. For example, counties that have higher circulation of national news, and thus higher news exposure might, for unobserved reasons, also have a higher propensity to migrate to fracking areas when a boom hits. This could happen if, for example, origin counties more tied to the oil and gas industry also had higher readership of national newspapers, and thus higher exposure when these booms happened. To address this potential concern, I employ an alternative strategy that exploits variation in news exposure among neighboring counties. Using all domestic newspapers that had at least one article about fracking between 1999 and 2012 and had circulation data available, I construct newspaper geographic markets. This is the set of counties in the newspaper's distribution network. For many local newspapers, distribution costs inhibit broad distribution and these markets are composed of a small group of adjacent counties around a central hub.[36] I then identify counties on the border of this distribution network as well as contiguous counties that do not receive the newspaper and compare the effect of news articles, specific to that newspaper, on migration and commuting for counties on both side of the market border. This is done in a stacked regression as follows

$$\begin{aligned} Y_{oSt} = \gamma_1 Articles_{nSt} * InMarket_{on} + \gamma_2 Articles^2_{nSt} * InMarket_{on} \\ +\gamma_3 InMarket_{on} + X'_{ot}\Gamma + \phi_{oS} + \psi_{nSt} + \varepsilon_{oSt}. \end{aligned} \tag{12}$$

[36] Over 90 percent of these newspapers distribute to 40 counties or less.

The outcome in equation (12) is the same as in previous specifications. The variable *Articles* is the number of articles in newspaper $n$ published in year $t$ about fracking in state $S$ (in units of ten), while *InMarket* is an indicator variable that equals one if the origin county $o$ is in the market for newspaper $n$. To be specific $n$ uniquely identifies each newspaper and the corresponding market border. So counties that do not receive newspaper $n$ but are on the market border will also be assigned to $n$. Time-varying origin controls and an origin-destination pair fixed effect are included. A newspaper-by-destination state-by-year fixed effect is also included, making this a comparison of flows to the same destination among counties along the same market border. The identifying assumption is that counties on either side of the newspaper's market border would evolve similarly, but for the news coverage about fracking. Because counties are being compared to other local counties, similar preferences and propensities among these neighboring counties captured in $\nu_{oSt}$ will be differenced out. This will identify the causal effect of news coverage as long as propensities to migrate to fracking during booms is local, but not county specific.

These results are reported in Table IV. Relative to no articles, ten news articles significantly increased migration by 5.6 percent in counties that received the newspaper, relative to their neighbors. There was also a similar significant effect on cross-county commuting (4.9 percent). In this sample, the average circulation rate among in-market border counties was slightly higher than the benchmark five percent at 5.5 percent, making it easy to compare the magnitude of this effect to the estimates from the previous specification. In the average in-market county with a newspaper circulation rate slightly higher than five percent, one additional local news article increases migration by approximately 0.6 percent.[37]

[37]The point estimates are similar if I exclude national newspapers (as these borders are less local). Counties might appear multiple times in the stacked regression in equation (12). However, if I restrict the sample to only include one newspaper market border per county to ensure that counties only appear once, the pattern is similar. To do this I take the set of newspaper market borders each county belongs to, and restrict the sample to only include the newspaper market border that had the highest number of articles about fracking among these market borders.

# VII Additional Explorations

## VII.A Impact of TV News Exposure

In a 2013 Pew Research report, 69 percent of adults cite television as one of their main sources for news. This rate has only slightly fallen from 74 percent in 2001. Meanwhile, the share of adults citing newspapers as a main source of news has fallen from 45 percent to 28 percent. Also during this time period, the internet has become an increasingly important source of news going from 13 in 2001 to 50 percent in 2013. Data constraints prevent me from comparing internet news exposure to traditional news sources, but I am able to compare migration and commute responses to television and newspaper news exposure.

Using abbreviated news transcripts from the three major TV news networks (ABC, CBS, and NBC) available through Vanderbilt Television News Archive (VTNA), I construct a measure of TV news exposure similar to the measure of national newspaper news as follows

$$TV\ news\ exposure_{oSt} = \sum_{c \in C} \big(\text{broadcasts on } c \text{ about fracking in } S\big)_t * pre09\ viewrate_{oc}. \quad (13)$$

The set $C = \{ABC, CBS, NBC\}$ and captures TV news coverage from the major national news networks. As with $newspaper\ exposure_{oSt}$, $TV\ news\ exposure_{oSt}$ captures variation in national news, which is weighted by the channel's pre-fracking Nielsen's viewership rates obtained through the 2008 Television and Cable Factbook. During this period, TV stations were transitioning from analog to digital broadcasts on a market-by-market basis, and new digital equipment was needed to receive the transmission. This might introduce bias in the viewership rates as only some markets had transitioned by the time data was collected. For this reason, I also run specifications using ratings from the latest 2016 Factbook, after these updates were fully made.[38] Nielsen ratings are only available at the DMA-level, which

[38]Using 2016 ratings potentially introduces endogeneity if viewership is responding to migration and commute behavior. However, circulation rates are highly persistent, suggesting this bias might be small.

is a mutually exclusive set of similar, nearby counties that represent a media market. To conduct this analysis, I aggregate all data up from the county to the DMA-level, including migration flows, newspaper circulation and exposure, and labor market measures. Typical viewership of ABC, CBS, and NBC was approximately 50 percent during this time period, so I scale TV news exposure such that a one unit increase represents the effect of one additional news broadcast from a TV network with 50 percent viewership.

This measure of TV exposure does not capture news exposure through cable news channels, such as CNN or Fox News. The VTNA only collects one hour of news broadcast data from these channels, and cable circulation is measured differently than traditional TV. If $TV\ news\ exposure_{oSt}$ is negatively correlated with cable news exposure (i.e., if network and cable news are substitutes) and both sources of news lead to more migration, than these estimates will be biased downward. If instead network and cable news are complements, network news could be interpreted as a proxy for total TV news.

DMA-level estimates are presented in Table V. I first report the effects of newspaper exposure on migration as the level of analysis has changed. One additional news article in a DMA with a five percent circulation rate increased migration to the fracking state mentioned by 5.0 percent. This point estimate is twice as large as the county-level estimate, but is not statistically different. Column (2) reports the estimated effects for TV news exposure. Using 2008 viewership rates, TV news exposure had no impact on migration behavior. In Column (3) I regress migration on both newspaper exposure and TV news exposure, to determine which news source is more closely associated with migration.[39] The coefficients for both news sources remain similar and the effect for newspaper exposure is significant. In Column (4) I conduct the same analysis, but use a measure of TV news exposure using 2016 viewership rates, as all markets had fully transitioned to digital TV by this time. In this specification the effect of newspaper circulation is 4.9 percent, while the effect of TV news exposure jumps to 7.9 percent (the p-value on the first order effect is 0.11). If only TV news exposure is

[39]Newspaper exposure and TV news exposure are moderately, positively correlated ($\rho = 0.36$).

included, the magnitude of the effect is similar and significant, suggesting TV news exposure also affects migration, but the relationship is much weaker than for newspaper news.

When looking at commute behavior, the DMA-level point estimate on newspaper exposure is smaller than the county-level estimate in Table III, but not statistically different.[40] The effect of TV exposure is large and significant. Relative to zero TV news exposure, one news broadcast in a DMA with a 50 percent circulation rate increased the number of people commuting to the fracking area discussed in the news report by 10.6 percent. The point estimates are largely unchanged when both news sources are included or when using 2008 or 2016 TV circulation rates. Many more people cite TV as a news source (Pew Research, 2013b), and circulation rates are an order of magnitude higher. However, there are far fewer TV broadcasts about fracking than newspaper articles, and these news clips are only 1-5 minutes. Observing similar impacts for both newspaper and TV exposure would suggest that when providing information about potential labor market opportunities, both the intensity of content and penetration influence the magnitude of the effect.

### VII.B Positive versus Negative News

Many newspaper articles that mentioned specific states also referenced positive characteristics such as jobs, booms, or growth. However, there were also many articles that discussed negative aspects such as pollution, health, dangers, and earthquakes (see Appendix Table A1). Although not necessary, it is possible that individuals are more responsive to positive news than negative news. It could also be the case that if people have preconceived beliefs about the value of fracking, and are only uncertain about where it is occurring, people might move even in response to "bad news". I parse the text of each news article to determine how many times positive and negative aspects are mentioned. I then classify an article as positive if it has at least two positive mentions and has more positive mentions than negative. Nega-

[40]When aggregating to the DMA-level, many neighboring counties fall into "fracking" DMAs that are excluded from the sample of origin DMAs. This likely attenuates the estimated impact on commuting.

tive articles are similarly defined. I then estimate the separate effect of positive and negative newspaper exposure on migration and cross-county commuting in Table VI.[41] Relative to no newspaper exposure in a county with a five percent circulation rate, one positive article significantly increased migration by 4.0 percent. This effect is one and a half times as large as the effect of one negative article, and significantly different. The effect of negative news is positive and significant, suggesting that "any news is good news", but people are more responsive to positive news. As changing origin characteristics that affect migration might be correlated with either positive or negative news rather than total news, I include origin county by year fixed effects and the effects are not statistically different.

In contrast to migration, positive and negative newspaper exposure affect commuting similarly. Relative to no exposure, one additional news article, positive or negative, leads to 7-9 percent higher cross-county commute flows. Including origin county by year fixed effects reduces the size of these effects but they remain significant and statistically indistinguishable. Unlike migrants, long distance commuters do not bear some of the drawbacks associated with fracking. such as potential home water contamination, earthquakes, or noise on residential streets. Negative news might provide a location signal, in addition to informing potential migrants about amenity costs that might be associated with moving; workers looking to commute might only value the location signal provided by the news.

In addition to exploring differences by source and content, I have also explored differences by newspaper and distance. The readership of the *New York Times* and *Wall Street Journal* is on average more-educated, higher income, and older than the typical migrant to fracking areas (Wilson, 2017). In Appendix Table A8, I show that the estimated effects on migration and commuting are largest and most significant for news in the *USA TODAY*, with smaller effects from the *New York Times*, and very imprecise, insignificant effects from the *Wall Street Journal*. To explore difference across distance, I estimate equation (10) for origin

[41]Exposure to neutral articles with less than two positive and two negative keywords are not included in this regression. Specifications including all three levels are similar but less precise.

county by destination state pairs in one hundred mile bins and plot the total marginal effect of news exposure for each distance in Appendix Figure A3. The effect climbs to about 6 percent between 400 and 1,000 miles and then gradually falls, consistent with information provision having an effect because people are aware of nearby opportunities, but lack information about distant potential opportunities.[42]

### VII.C The Role of Origin County Labor Market Conditions

In recent years there has been considerable concern about decreasing labor market fluidity and mobility, especially when it appears that people in weak labor markets could encounter more abundant opportunities elsewhere (Molloy, Smith, Trezzi, & Wozniak, 2016).[43] Previous evidence from the MTO experiment and educational differences in migration are consistent with information constraints playing a role (Malamud & Wozniak, 2012; Chetty, et al., 2015). I next explore heterogeneity in the migration response to newspaper exposure by the labor market strength of the origin to understand if providing information is particularly impactful for people in weak economic areas. To do this, I estimate a variant of my main specification as follows

$$\begin{aligned} Y_{oSt} = {} & \beta_1 newspaper\ exposure_{oSt} + \beta_2 newspaper\ exposure^2_{oSt} \\ & + \beta_3 newspaper\ exposure_{oSt} * emp/pop_{ot-1} + \beta_4 newspaper\ exposure_{oSt} * emp/pop^2_{ot-1} \\ & + \beta_5 newspaper\ exposure^2_{oSt} * emp/pop_{ot-1} + \beta_6 newspaper\ exposure^2_{oSt} * emp/pop^2_{ot-1} \\ & + \beta_7 emp/pop_{ot-1} + \beta_8 emp/pop^2_{ot-1} + X'_{ot}\Gamma + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \end{aligned} \tag{14}$$

Where *emp/pop* is the lagged county employment to population ratio, for the adult population. I then calculate the total effect of one unit of newspaper exposure, which is allowed to vary quadratically with the employment to population ratio and use the delta method to obtain standard errors (the corresponding coefficients are provided in Appendix Table A9).

[42] The effects of news on commuting for different age groups is similar (Appendix Figure A4).

[43] This topic has also come up in the popular press (Brooks, 2016; Cohen, 2016).

The effects on both migration and commuting are plotted in percentage points in Figure V for county employment to population ratios between 60 and 85 percent (approximately the 15th to 90th percentile). Both the migration and commute responses are larger for weaker economic areas. A one unit increase in newspaper exposure led to a 2.8 percent increase in migration from counties with a low employment to population ratio, but had a small, one percent impact on migration from counties with a high employment to population ratio. Low employment counties saw commute flows increase by nearly 8 percent for an additional news article, while counties with high employment saw increases closer to 2 percent. Exposure to news about fracking in distant, potential labor markets had a larger impact on migration flows from economically weak areas, suggesting informational constraints might be a contributing factor to differences in migration behavior.[44]

Using this specification, I estimate the implied impact of various news exposure treatments at different points in the origin employment to population ratio distribution as reported in Table VII. These simulations highlight the heterogeneous impacts of different potential policy interventions. Even increasing the exposure level by one unit (one additional article in a county with a five percent readership rate) has substantial impacts across the distribution. The impacts are the largest for counties with employment to population ratios below the mean of 70.9 percent (effects between 2.7-2.8 percent), and monotonically decrease as the local labor market conditions improve. As exposure increases to five or ten the impacts become very large for counties with weak labor markets and the differential impact becomes substantial. Increasing exposure to ten increases migration flows from counties with weaker than average labor markets by 19-21 percent, 20.5 percent for counties near the mean, and 11.6 percent for counties with very strong labor markets. Once exposure reaches 15 (close to the top one percent of exposure in 2012) effect sizes start to plateau at most employment to population levels, suggesting the additional benefits of exposure at this level are small. For reference the average actual exposure level in 2012, along with the implied impact of this

[44]The pattern is more flat, but still downward sloping if instead the unemployment rate is used.

level of exposure are also reported. Average actual exposure is between 1.5 and 2.5 for all levels of labor market strength, although it does peak slightly around the mean. For origins near the mean employment to population level, migration flows would have been 6.4 percent lower if there had been no news about these local fracking booms.

From equation (14), I solve for the value of exposure that maximizes the impact for each point in the employment to population distribution. For counties with employment to population ratios below the average at 65 or 70 percent, the maximizing level of exposure is very high, at 18.4 or 19.6. This level decreases, falling to only 15.6 for counties with employment at 80 percent. The maximum implied impacts also vary greatly, going from 27 percent in areas below the average to only 13.3 percent at the top. The patterns are similar when looking at commute behavior although the impacts are larger and more heterogeneous, while the maximizing level are more uniform (see Appendix Table A10).

Heterogeneous impacts by origin labor market strength could result from differential exposure to new or heterogeneous returns to the information. As differences in actual exposure are small it is likely that only a small part of the heterogeneous impacts can be explained by differential exposure. Exposure at all levels of labor market strength is substantially lower than the maximizing level, suggesting all regions face limited information. However, the information is most valuable for people living in counties with weak labor markets, where the expected gains to moving are largest.

This has several implications for potential policy interventions. Information provision policies could increase geographic mobility, potentially resulting in more beneficial labor market transitions (Molloy et al., 2016) and higher economic mobility (Chetty & Hendren, 2016). Providing a modest amount of information about potential labor market opportunities in other parts of the country to all counties will significantly increase migration to those regions. However, a government facing limited resources would see the largest returns by focusing on providing information to weak labor markets. Not only would the migrant benefit, by encountering more favorable labor markets, but this might also generate positive

externalities for workers in the weak origin labor market, as the market becomes less slack.

## VIII Online Searches, a Potential Mechanism

The data indicate that when counties are more exposed to news coverage about fracking in a certain state, migration and commute flows to fracking areas in that state increase. This relationship posits that news coverage provides information about potential labor market opportunities, and affects migration through changing expectations and uncertainty. This channel, however, cannot be directly tested in the data. Rather than verify that people's expectations change, I am able to quantify how interest in fracking and the states mentioned changes after news is disseminated. Using Google Trends data, I next explore search interest before and after TV news broadcasts about fracking.[45] For a specified search term (i.e., "fracking"), Google Trends will provide a time-series of search intensity at the national, state, or DMA level. This time-series is an ordinal measure of intensity that equals 100 on the day with the highest number of searches per capita, and with every other day scaled as a percent of the maximum. For example, on a day that is assigned a value of 20, search intensity for the search term was only 20 percent the level from the maximum day. This measure facilitates comparisons within a region over time, but is not conducive to studying differences across both geography and time. As such, I will examine changes in search behavior a short period before and after a TV news broadcast, but cannot reliably determine if search intensity increased by more in areas with higher TV viewership rates.

For each of the 17 TV news broadcasts that mention "fracking" or "shale", I pull daily time-series for every DMA in the United States for 15 days before the broadcast, the day of the broadcast, and 14 days after for several search terms.[46] First I look at search intensity for the term "fracking", and then I look at search intensity for the name of any states

[45] Ideally I would also like to look at search behavior after newspaper articles are published. However, as there are over 560 articles, the pre- and post- windows for each article overlap extensively.

[46] A special thanks to Tanner Eastmond for help working through the Python code.

that are mentioned in the broadcast. States are only mentioned in 14 of the broadcasts. To identify the impact of the news broadcast on average search intensity I estimate the following regression

$$searchindex_{opt} = \sum_{\tau=-14}^{14} \delta_\tau * \mathbf{1}\{t \text{ is } \tau \text{ days from broadcast}\}_{op} + X_t'\Gamma + \phi_{op} + DOW_t + \varepsilon_{opt} \quad (15)$$

where $searchindex_{opt}$ is the search index on date $t$ in DMA $o$ relative to the search period $p$. The search period is the 15 days prior, the day of, and the 14 days after each broadcast, such that $op$ uniquely identifies each DMA/term pair, over which the relative search index is measured. The set of coefficients $\delta_\tau$ trace out the daily deviations in the search index from the omitted day ($\tau = -15$). I include a DMA by search period fixed effect in order to compare days from the same search that have comparable indices. Day of the week fixed effects are also included to account for differences in search behavior during different times in the week.

Several of the news reports were broadcast in close proximity to other high publicity events connected to either fracking or the states mentioned in the reports. For example, On January 24, 2012, four days prior to a news report about fracking in Pennsylvania, President Barack Obama discussed shale gas extraction and fracking in the State of the Union. Similarly, late on December 31, 2011, there was an earthquake in Youngstown, Ohio, that many linked to fracking. Four days later there was a news report on fracking in Ohio. There are also other high publicity state specific events (such as college football bowl games or school shootings) that occur during some of the search period windows. When looking at searches for fracking I include indicator controls for the three days after the two events related to fracking, and when looking at searches for specific state names I include a set of indicator controls for the local, high publicity events that are listed in Appendix Table A11. If I do not control for these events the series becomes more volatile, but there is still a significant spike directly after the broadcast. Figures that do not control for other high

publicity events are included in Appendix Figure A5.[47]

These effects are plotted in Figure VI for "fracking" and Figure VII for the state names. As seen in Figure VI, search intensity for "fracking" spikes the day of the broadcast and remains elevated for the next two days before falling back to the previous levels. Across all DMAs, search intensity for "fracking" jumps by nearly 2.5 points on average. Because the search index is a relative measure, this cannot be converted to back out how many additional searches were made. If I combine days into 3 day bins, for statistical power, I estimate a 2.5 point spike in days 0 to 2, followed by an statistically significant one point increase for the remainder of the days in the sample, suggesting that search interest remained elevated for some time (see Appendix Figure A6).

When looking at search interest in the names of states that were mentioned in the broadcast there is also a spike one day after the broadcast and interest remains elevated for the next five days. Although not directly comparable, the jump in search intensity is larger at nearly 6.5 points. For reference I also regress the search intensity for the same set of fracking states that are not mentioned in the news broadcasts. These estimates remain close to zero, with no spike or increase after the news broadcast, suggesting this is not capturing overall interest in fracking states. The Google Trends data indicate that following a TV news broadcast about fracking, people search more for fracking and for the states mentioned in the broadcast. Although this is not direct evidence that news coverage motivates people to move, it does suggest news coverage induces people to seek more information about the potential fracking destination, which might be due to an interest in moving or a desire to obtain more information.[48]

[47]In theory I would like to look at search behavior after newspaper publications as well. However, national news articles about fracking were published quite frequently, leading to large overlap in sample windows.

[48]I have also looked at search interest in moving specific terms such as "Uhaul" or "Uhaul rental". At both the DMA and state-level there appears to be a visual shift at the time of broadcast, however it is not statistically different. Search intensity for terms like "fracking jobs" or "oil jobs" are low and frequently suppressed by Google. These patterns are similar if the window is extended.

# IX Conclusion

Migration is an investment that can improve the types of labor markets individuals encounter, but many of those that appear to face the largest benefit do not move. This could be the result of optimal behavior, but could also be due to various constraints or market frictions, such as limited information. In this paper I evaluate the role of information in the decision to move to more favorable labor market opportunities. The current literature speaks very little to the effect of labor market information on migration behavior. I outline a conceptual framework for understanding how information will affect the migration decision and potentially change migration outcomes.

To estimate the effect of news on migration I exploit information disseminated through national news coverage of localized fracking booms. The technological and geological constraints associated with fracking have led to sudden, large labor market shocks in well-defined areas. The novelty of fracking also makes it straightforward to identify news coverage about fracking across different affected areas. I combine national news content with historic local circulation rates to construct a measure of news exposure that strips away endogenous changes in consumer readership and endogenous changes in producer content decisions.

The data suggest that for a county with a five percent circulation rate and no previous exposure, one news article about fracking in a specific state increased migration flows to fracking counties in that state by 2.4 percent. This estimate accounts for destination specific characteristics that are changing over time and does not change when controlling for origin-level shocks or local news exposure. Cross-county commute flows increase by 6.6 percent.

I also provide evidence that TV news exposure has an effect on commuting and potentially migration also. The magnitude of these effects are similar to the response to newspaper news. Migration flows are more responsive to exposure to positive news than negative news, though both lead to more migration. In contrast, commute flows respond similarly to positive and negative news, consistent with commuters not facing many of the negative costs associated

with fracking at their homes (e.g., water contamination, increased risk of earthquakes). As further evidence that news coverage increases interest in these fracking destinations, I find that, directly after a TV news broadcast about fracking, Google search interest in both the term "fracking" and the names of states mentioned significantly increases. News exposure induces people to seek more information on the internet, which might influence the migration decision.

Importantly, the migration response is largest from origin counties that have been experiencing weak labor market conditions, suggesting the benefit to news provision is largest in those areas. This has potential implications for people trying to understand why less-educated and low-income households in poor performing labor markets are unlikely to move, and if there are policies that can encourage more migration to better economic opportunity. Simulations suggest that providing more information about potential labor market opportunities in other areas would increase geographic mobility in all areas, with the most pronounced response in weak labor markets where the returns to migration are the largest.

## References


Bartik, Alexander, Janet Currie, Michael Greenstone, and Christopher Knittel. 2017. "The Local Economic and Welfare Consequences of Hydraulic Fracturing," Available at SSRN: https://ssrn.com/abstract=2692197.

Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 5(2): 9-49.

Bottan, Nicolas and Ricardo Perez-Truglia. 2017. "Choosing Your Pond: Revealed-Preference Estimates of Relative Income Concerns," Working Paper No. 23615, National Bureau of Economic Research, Cambridge, MA.

Bound, John and Harry Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes During the 1980s," *Journal of Labor Economics*, 18(1): 20-54.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: the Case of Seasonal Migration in Bangladesh," *Econometrica*, 82(5): 1671-1748.

Brooks, Arthur C. 2016. "How to Get Americans Moving Again," *The New York Times*, May 20, 2016. https://www.nytimes.com/2016/05/21/opinion/how-to-get-americans-moving-again.html.

Bureau of Labor Statistics, U.S. Department of Labor. 2015. The Economics Daily, State employment-population ratios in 2014 on the Internet at https://www.bls.gov/opub /ted/2015/state-employment-population-ratios-in-2014.html (visited July 18, 2017).

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2015. "The Effect of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 106 (4): 855-902.

Chetty, Raj, Nathaniel Hendren. 2016. "The Impact of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," Working Paper No. 23002, National Bureau of Economic Research, Cambridge, MA.

Chong, Alberto and Eliana La Ferrara. 2009, "Television and Divorce: Evidence from Brazilian "Novelas"," *Journal of the European Economic Association*, 7(2/3): 458-468.

Cohen, Patricia. 2016. "Fewer Americans Strike Out for New Jobs, Crimping the Recovery," *The New York Times*, May 24, 2016. https://www.nytimes.com/2016/05/21/opinion/how-to-get-americans-moving-again.html.

Dao, Mai, Davide Furceri, and Prakash Loungani. 2017. "Regional Labor Market Adjustments in the United States: Trend and Cycle," *the Review of Economics and Statistics*, 99(2): 243-257.

DellaVigna, Stefano and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 122(3): 1187-1234.

Dionne, Georges and Scott E. Harrington. 1991. An Introduction to Insurance Economics, in Foundations of Insurance Economics: *Readings in Economics and Finance*, Georges Dionne and Scott Harrington eds. (Boston, Mass: Kluwer Academic, 1991).

Farre, Lidia and Francesco Fasani. 2013. "Media exposure and internal migration – Evidence from Indonesia," *Journal of Development Economics*, 102: 48-61.

Feyrer, James, Erin Mansur, and Bruce Sacerdote. 2017. "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution," *American Economic Review*, 107(4): 1313-1334.

Ganong, Peter and Daniel Shoag. 2017. "Why has Regional Income Convergence in the U.S. Declined?" Working Paper No. 23609, National Bureau of Economic Research, Cambridge, MA.

Garthwaite, Craig and Timothy J. Moore. 2012. "Can Celebrity Endorsements Affect Political Outcomes? Evidence from the 2008 US Democratic Presidential Primary," *The Journal of Law, Economics, & Organization*, 29(2): 355-384.

Garthwaite, Craig. 2014. "Demand Spillovers, Combative Advertising, and Celebrity Endorsements," *American Economic Journal: Applied Economics*, 6(2): 76-104.

Gentzkow, Matthew. 2006. "Television and Voter Turnout," *Quarterly Journal of Economics*, 121(3): 931-972.

Gentzkow, Matthew, and Jesse Shapiro. 2010. "What Drives Media Slant? Evidence from U.S. Daily Newspapers," *Econometrica*, 70(1): 35-71.

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

Greenwood, Michael J. 1975. "Research on Internal Migration in the United States: A Survey," *Journal of Economic Literature*, 13(2): 397-433.

Greenwood, Michael J. 1985. "Human Migration: Theory, Models, and Empirical Studies," *Journal of Regional Science*, 25(4): 521-544.

Hanson, Gordon and Craig McIntosh. 2010. "The Great Mexican emigration," *the Review of Economics and Statistics*, 92(4): 798-810.

Jensen, Robert and Emily Oster. 2009. "The Power of TV: Cable Television and Women's Status in India," *Quarterly Journal of Economics*, 124(3): 1057-1094.

Kaplan, Greg and Sam Schulhofer-Wohl. 2017. "Understanding the Long-Run Decline in Interstate Migration," *International Economic Review*, 58(1): 57-94.

Kearney, Melissa and Phillip B. Levine. 2015a. "Early Childhood Education by MOOC: Lessons from Sesame Street," Working Paper No. 21229, National Bureau of Economic Research, Cambridge, MA.

Kearney, Melissa and Phillip B. Levine. 2015b. "Impact of MTV's *16 and Pregant* on Teen Childbearing," *American Economic Review*, 105(12): 3597-3632.

Kling, Jeffrey, Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1): 83-119.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap Operas and Fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 4(4): 1-31.

Langley, Paul C. 1974. "The Spatial Allocation of Migrants in England and Wales: 1961-66," *Scottish Journal of Political Economy*, 21: 259-277.

Malamud, Ofer and Abigail Wozniak. 2012. "The Impact of College Migration: Evidence from the Vietnam Generation," *Journal of Human Resources*, 47(4): 913-950.

Maniloff, Peter and 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.

McKenzie, David, and Hillel Rapoport. 2007. "Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence From Mexico," *Journal of Development*

*Economics*, 84(1): 1-24.

McKenzie, David and Hillel Rapoport. 2010. "Self-Selection patterns in Mexico-U.S. migration: the role of migration networks," *the Review of Economics and Statistics*, 92(4): 811-821.

Molloy, Raven, Christopher Smith, and Abigail Wozniak. 2011. "Internal Migration in the United States," *Journal of Economic Perspectives*, 25(2): 1-42.

Molloy, Raven, Christopher Smith, Riccardo Trezzi, and Abigail Wozniak. 2016. "Understanding declining fluidity in the U.S. Labor Market," :Brookings Paper in Economic Activity, Spring 2016.

Monras, Joan. 2015. "Economic Shocks and Internal Migration," *IZA DP No.* 8840.

Munshi, Kaivan. 2003. "Networks in the modern economy: Mexican migrants in the US labor market," *Quarterly Journal of Economics*, 118, 549-599.

Notowidigdo, Matthew. 2013. "The Incidence of Local Labor Demand Shocks," Working Paper No. 17167, National Bureau of Economic Research, Cambridge, MA.

O'Connell Paul G. J. 1997. "Migration Under Uncertainty: 'Try Your Luck' or 'Wait and See'," *Journal of Regional Science*, 37(2): 331-347.

Pew Research Center, Washington D.C. 2013a. "As Gas Prices Pinch, Support for Oil and Gas Production Grows, Those Aware of Fracking Favor Its Use," (March, 19, 2012). http://www.people-press.org/2012/03/19/as-gas-prices-pinch-support-for-oil-and-gas-production-grows/.

Pew Research Center, Washington D.C. 2013b. "Amid Criticism, Support for Media's 'Watchdog' Role Stands Out," (August, 8, 2013). http://www.people-press.org/2013/08/08/amid-criticism-support-for-medias-watchdog-role-stands-out/2-main-source-for-news/.

Rothschild, Michael, and Joseph Stiglitz. 1970. "Increasing risk: I. A definition," *Journal of Economic Theory*, 2(3): 225-243.

Ruggles, Steve, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. Integrated Public Use Microdata Series: Version 6.0. Minneapolis: University of Minnesota.

Shultz, Theodore W. 1961. "Investment in Human Capital," *American Economic Review*, 51 (1):1-17.

Sjaastad, Larry. 1962. "The costs and returns to human migration," *Journal of Political Economy*, 70(supplemental): 80-93.

Tobin, James. 1965. "Money and Economic Growth," *Econometrica*, 33(4): 671-684.

Todaro, Michael P. 1969. "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," *American Economic Review*, 59(1): 138-148.

Wilson, Riley. 2017. "Moving to Economic Opportunity: the Migration Response to the Fracking Boom," Available at SSRN: https://ssrn.com/abstract=2814147.

Winters, Paul, Alain de Janvry, and Elsabeth Sadoulet. 2001. "Family and Community Networks in Mexico-U.S. Migration," *the Journal of Human Resources*, 36(1): 159-184.

Wiswall, Matthew and Basit Safar. 2015. "How Do College Students Respond to Public Information about Earnings," *Journal of Human Capital*, 9(2): 117-169.

Wozniak, Abigail. 2010. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources*, 45(4): 944-970.

## Tables

Table I

County Characteristics by *USA TODAY* Pre-Fracking Circulation Rate

| | County Characteristics in 2000 | | Change from 2000 to 2010 | | Predicted Difference from 25th to 75th Percentile |
|---|---|---|---|---|---|
| | Below Median (1) | Above Median (2) | Below Median (3) | Above Median (4) | (5) |
| *Migrants to fracking areas (Percent of Population)* | 0.09 | 0.11 | -0.01 | -0.02 | 0.00 |
| *Employment to Population (16+)* | 55.68 | 59.14 | -0.78 | -1.43 | -0.10 |
| *Unemployment Rate* | 3.29 | 3.44 | 0.96 | 1.33 | -0.09** |
| *Median Household Income* | 31,805 | 38,834 | 8,485 | 8,763 | 290*** |
| *Percent in Poverty* | 15.82 | 12.29 | 0.83 | 2.18 | 0.51*** |
| *Percent White* | 85.16 | 84.32 | -0.83 | -2.26 | -0.58*** |
| *Percent Black* | 8.95 | 9.27 | -0.15 | 0.42 | 0.17*** |
| *Percent Hispanic* | 5.87 | 5.72 | 1.77 | 2.38 | 0.29*** |
| *Percent Other Race* | 5.89 | 6.41 | 0.98 | 1.85 | 0.41*** |
| *Percent Population 20-34* | 16.67 | 19.60 | -0.55 | -0.67 | 0.02 |
| *Percent Population 35-64* | 38.76 | 38.56 | 1.82 | 1.47 | -0.27*** |
| *Percent Population Over 64* | 15.96 | 13.64 | 1.20 | 1.15 | 0.00 |
| *Percent Households Renting* | 23.47 | 28.30 | 1.64 | 1.81 | 0.10** |
| *Number of Counties* | 1,420 | 1,418 | 1,420 | 1,418 | 2,838 |

Notes: Migration data from the IRS Statistics of Income. Other county characteristics obtained through American FactFinder from the 2000 Census and 2010 Census and 5-Year American Community Survey. *USA TODAY* circulation data from the Alliance for Audited Media. The county level median pre-2009 circulation rate of the *USA TODAY* was 0.83 percent, and ranged from 0 to 27.8 percent. Median Household Income is reported in current dollars. Column (5) reports the predicted change in the characteristic between 2000 to 2010 when pre-2009 circulation increases from the 25th to the 75th percentile. p<0.01 ***, p<0.05 **, p<0.1 *.

Table II

Impact of Destination State Specific National Newspaper Exposure on Migration to Fracking Counties in State

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | | | Number of Migrants$_{oSt}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *National Newspaper Exposure*$_{oSt}$ | 0.025*** | 0.025*** | 0.026*** | 0.025*** | 1.439*** | 1.491*** | 1.673*** | 1.644*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.419) | (0.437) | (0.534) | (0.532) |
| *National Newspaper Exposure*$^2_{oSt}$ | -0.001*** | -0.001*** | -0.001*** | -0.001*** | -0.036* | -0.038* | -0.036* | -0.038* |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.020) | (0.020) | (0.022) | (0.021) |
| *Local Newspaper Exposure*$_{oSt}$ | | | | 0.009** | | | | -0.671 |
| | | | | (0.004) | | | | (1.662) |
| *Local Newspaper Exposure*$^2_{oSt}$ | | | | -0.0001** | | | | 0.057 |
| | | | | (0.00004) | | | | (0.043) |
| *Origin Labor Market Controls* | | X | | | | X | | |
| *Origin by Year Effects* | | | X | X | | | X | X |
| *Origin/Destination Local News* | | | | X | | | | X |
| *Mean Number of Migrants* | 7.6 | 7.6 | 7.6 | 7.6 | 7.6 | 7.6 | 7.6 | 7.6 |
| *Observations* | 590,224 | 590,224 | 590,224 | 590,224 | 590,224 | 590,224 | 590,224 | 590,224 |

Notes: Data from the IRS Statistics of Income, LexisNexis newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. The level of observation is the origin county by destination state by year from 2000 to 2012. Origin counties with any fracking production or in the New York City designated market area are excluded. The variable *National Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. For origin/destination pairs with any news exposure, mean national newspaper exposure is 0.99. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Origin controls include the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$). Origin/destination specific local news is all destination state specific fracking news content listed in LexisNexis from non-national domestic newspapers. The variable *Local Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 40 percent circulation rate, approximately the 95th percentile of pre-fracking circulation among non-national newspapers with articles about fracking. The sample correlation between national news exposure and local news exposure is approximately 0.12. Origin county by year fixed effects control for time-varying characteristics of the origin county and account for potential changes in preferences toward fracking that might be correlated with newspaper readership and affect migration to fracking areas. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table III

Impact of Destination State Specific Newspaper Exposure on Cross-County Commuting to Fracking Counties in State

| | Inverse Hyperbolic Sine of the Number of Cross-County Commute Jobs$_{oSt}$ | | | |
|---|---|---|---|---|
| | | By Age | | |
| | All Jobs | Under 30 | 30-54 | Over 54 |
| | (1) | (2) | (3) | (4) |
| *National Newspaper Exposure*$_{oSt}$ | 0.068*** | 0.032*** | 0.053*** | 0.037*** |
| | (0.009) | (0.005) | (0.007) | (0.005) |
| *National Newspaper Exposure*$^2_{oSt}$ | -0.002*** | -0.001*** | -0.001*** | -0.001*** |
| | (0.0004) | (0.0002) | (0.0003) | (0.0002) |
| *Dependent Mean* | 31.4 | 8.6 | 18.0 | 4.9 |
| *Observations* | 499,440 | 499,440 | 499,440 | 499,440 |

Notes: Data from the LEHD Origin-Destination Employment Statistics (LODES), LexisNexis newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. The level of observation is the origin county by destination state by year from 2002 to 2012. LODES data is only available starting in 2002. Origin counties with any fracking production or in the New York City designated market area are excluded. The variable *National Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. For comparison, circulation of the *USA TODAY* was 4.5 percent at the 95th percentile. Commuting jobs are also examined by age groups, pre-defined in the LODES data. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. The effect of national newspaper exposure on commuting is significantly larger for workers aged 30-54 than the other two age groups. p<0.01 ***, p<0.05 **, p<0.1 *.

Table IV

Newspaper Market Cross Border Analysis: Impact of Newspaper Articles on Migration and Commuting

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | | Inverse Hyperbolic Sine of the Number of Cross-County Commute Jobs$_{oSt}$ | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *News articles*$_{nSt}$ $*$ *In-Market*$_{on}$ | 0.061*** | 0.056*** | 0.069*** | 0.054** | 0.042** | 0.082*** |
| *(in 10s of Articles)* | (0.013) | (0.012) | (0.017) | (0.023) | (0.021) | (0.032) |
| *News articles*$^2_{nSt}$ $*$ *In-Market*$_{on}$ | -0.005*** | -0.004*** | -0.005*** | -0.005** | -0.004** | -0.007** |
| *(in 10s of Articles)* | (0.001) | (0.001) | (0.002) | (0.002) | (0.002) | (0.003) |
| *In-Market*$_{on}$ | -0.001*** | -0.001*** | -0.00004** | -0.001** | -0.001* | -0.0001* |
| | (0.0002) | (0.0002) | (0.00002) | (0.0004) | (0.0005) | (0.00003) |
| *Exclude National Newspapers* | | X | | | X | |
| *Only One Border per Origin County* | | | X | | | X |
| *Dependent Mean* | 24.4 | 24.0 | 14.1 | 144.0 | 141.8 | 72.2 |
| *Observations* | 1,476,352 | 1,465,648 | 509,664 | 1,250,112 | 1,240,784 | 431,360 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. The level of observation is the origin county by destination state by year from 2000 to 2012 for the migration data and 2002 to 2012 for the commute data. Sample includes all counties on both sides of the border of a newspaper market for any of the 220 newspapers with an article about fracking and circulation data. News articles are newspaper and destination state specific and measured in units of ten. In-market is an indicator that equals one if the county is inside the newspaper's market area (i.e., has positive circulation). In all specifications origin/destination pair fixed effects are included to control for time-invariant differences across pairs. Newspaper market border by destination by year fixed effects are also included to control for characteristics of the local border/destination that vary over time, and make this a comparison of origin counties within the same newspaper market border. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. Average circulation among in-market counties across all newspapers was 5.5 percent. Columns (2) and (5) exclude national newspapers, as their market borders are not local. In Columns (3) and (6) the sample is restricted to only include one newspaper market border per county, and it is the border that had the highest number of articles about fracking among all of the borders the county belongs to. Standard errors adjusted for clustering at the origin designated market area are in parentheses. p<0.01 ***, p<0.05 **, p<0.1 *.

Table V

Source of News: Impact of Newspaper and TV News Exposure on Migration to Fracking Regions

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | | | Inverse Hyperbolic Sine of the Number of Cross-County Commute Jobs$_{oSt}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Newspaper Exposure*$_{oSt}$ | 0.052*** | | 0.052*** | 0.051** | 0.045** | | 0.055** | 0.045** |
| | (0.018) | | (0.018) | (0.018) | (0.019) | | (0.019) | (0.019) |
| *Newspaper Exposure*$^2_{oSt}$ | -0.002*** | | -0.002*** | -0.002*** | -0.001** | | -0.001** | -0.001** |
| | (0.001) | | (0.001) | (0.001) | (0.001) | | (0.001) | (0.001) |
| *TV News Exposure*$_{oSt}$ | | 0.040 | 0.029 | 0.100 | | 0.125** | 0.115* | 0.134* |
| | | (0.072) | (0.072) | (0.064) | | (0.062) | (0.061) | (0.069) |
| *TV News Exposure*$^2_{oSt}$ | | -0.010 | -0.011 | -0.021 | | -0.019 | -0.019 | -0.018 |
| | | (0.017) | (0.017) | (0.013) | | (0.014) | (0.014) | (0.014) |
| *2008 TV Viewership Rates* | | X | X | | | X | X | |
| *2016 TV Viewership Rates* | | | | X | | | | X |
| *Dependent Mean (in Levels)* | 60.2 | 60.2 | 60.2 | 60.2 | 152.6 | 152.6 | 152.6 | 152.6 |
| *Observations* | 32,864 | 32,864 | 32,864 | 32,864 | 27,808 | 27,808 | 27,808 | 27,808 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. TV news circulation is only available at the Designated Market Area (DMA) level, from the 2008 Television Factbook, and all data is aggregated to that level. The level of observation is the origin DMA by destination state by year from 2000 to 2012. The variable *Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. For comparison, circulation of the *USA TODAY* was 4.5 percent at the 95th percentile. The variable *TV News Exposure*$_{oSt}$ is scales such that a one unit increase represents the impact of one additional TV news broadcast on a network with a 50 percent circulation rate, approximately the average circulation rate of ABC, CBS, or NBC. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Controls for the origin DMA unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. In 2008, there was significant transition to digital TV and full viewership ratings were not available, so Columns (4) and (8) use TV circulation from 2016 to construct TV news exposure. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table VI

Positive vs. Negative News: Impact of Newspaper Exposure on Migration and Commuting to Fracking Regions

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | Inverse Hyperbolic Sine of the Number of Cross-County Commuting Jobs$_{oSt}$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Positive Newspaper Exposure*$_{oSt}$ | 0.044*** | 0.047*** | 0.072*** | 0.056*** |
| | (0.010) | (0.011) | (0.018) | (0.020) |
| *Positive Newspaper Exposure*$^2_{oSt}$ | -0.004** | -0.004** | -0.007** | -0.003 |
| | (0.002) | (0.002) | (0.003) | (0.003) |
| *Negative Newspaper Exposure*$_{oSt}$ | 0.026*** | 0.025*** | 0.098*** | 0.051*** |
| | (0.005) | (0.005) | (0.014) | (0.011) |
| *Negative Newspaper Exposure*$^2_{oSt}$ | -0.001*** | -0.001*** | -0.006*** | -0.002*** |
| | (0.0005) | (0.0004) | (0.001) | (0.001) |
| *Origin by Year Fixed Effects* | | X | | X |
| *Dependent Mean* | 7.6 | 7.6 | 31.4 | 31.4 |
| *Observations* | 590,224 | 590,224 | 499,440 | 499,440 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. The level of observation is the origin county by destination state by year from 2000 to 2012 for the migration data and 2002 to 2012 for the commute data. *Exposure*$_{oSt}$ measures are scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. A positive news article is one that contains at least two positive phrases (referencing jobs, boom, or growth) and more positive than negative phrases (referencing pollution, health, danger, or earthquakes), while a negative article is the opposite. Some fracking destinations have many positive and negative articles, leading to a high correlation between *Positive Newspaper Exposure*$_{oSt}$ and *Negative Newspaper Exposure*$_{oSt}$ ($\rho = 0.70$). Controls include the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$). In all specifications origin/destination pair fixed effects and destination by year fixed effects, are included to control for time-invariant differences across pairs and characteristics of the destination and origin that vary over time. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table VII

Simulated Impacts of News Exposure Migration Flows by Origin Employment to Population Ratio

| | Employment to Population Ratio in $t-1$ ($\mu = 70.9$) | | | | |
|---|---|---|---|---|---|
| | 60 | 65 | 70 | 75 | 80 |
| *Exposure Level* | | | | | |
| *0* | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| *1* | 0.027*** | 0.028*** | 0.027*** | 0.023*** | 0.017*** |
| | (0.003) | (0.004) | (0.004) | (0.003) | (0.003) |
| *5* | 0.116*** | 0.125*** | 0.120*** | 0.103*** | 0.072*** |
| | (0.015) | (0.016) | (0.016) | (0.015) | (0.014) |
| *10* | 0.189*** | 0.210*** | 0.205*** | 0.174*** | 0.116*** |
| | (0.025) | (0.027) | (0.027) | (0.026) | (0.026) |
| *15* | 0.218*** | 0.256*** | 0.255*** | 0.214*** | 0.133*** |
| | (0.031) | (0.034) | (0.035) | (0.035) | (0.040) |
| *Mean Exposure in 2012* | 1.57 | 2.08 | 2.48 | 2.44 | 1.89 |
| *Implied Impact* | 0.041*** | 0.057*** | 0.064*** | 0.054*** | 0.030*** |
| | (0.005) | (0.007) | (0.008) | (0.008) | (0.006) |
| *Maximizing Exposure* | 15.9 | 18.4 | 19.6 | 18.9 | 15.6 |
| *Implied Impact* | 0.219*** | 0.265*** | 0.270*** | 0.223*** | 0.133*** |
| | (0.032) | (0.038) | (0.041) | (0.043) | (0.042) |

Notes: Simulated impacts are obtained for each combination of origin employment to population ratio and exposure level from equation (14), where the outcome is the inverse hyperbolic sine of the number of migrants. The corresponding coefficients are reported in Appendix Table A9. The maximizing exposure is obtained by setting the first derivative of equation (14) with respect to newspaper exposure equal to zero and solving for the maximizing exposure for the specified employment to population ratio. This value is rounded down to the nearest whole number. The implied impact is the corresponding effect of the maximizing exposure level. For reference, the mean exposure level in 2012 for origin counties with employment to population ratios within 2.5 percent of the specified threshold. An employment to population ratio of 60 roughly corresponds to the 15th percentile while a ratio of 80 corresponds to roughly the 85th percentile. p<0.01 ***, p<0.05 **, p<0.1 *.

## Figures

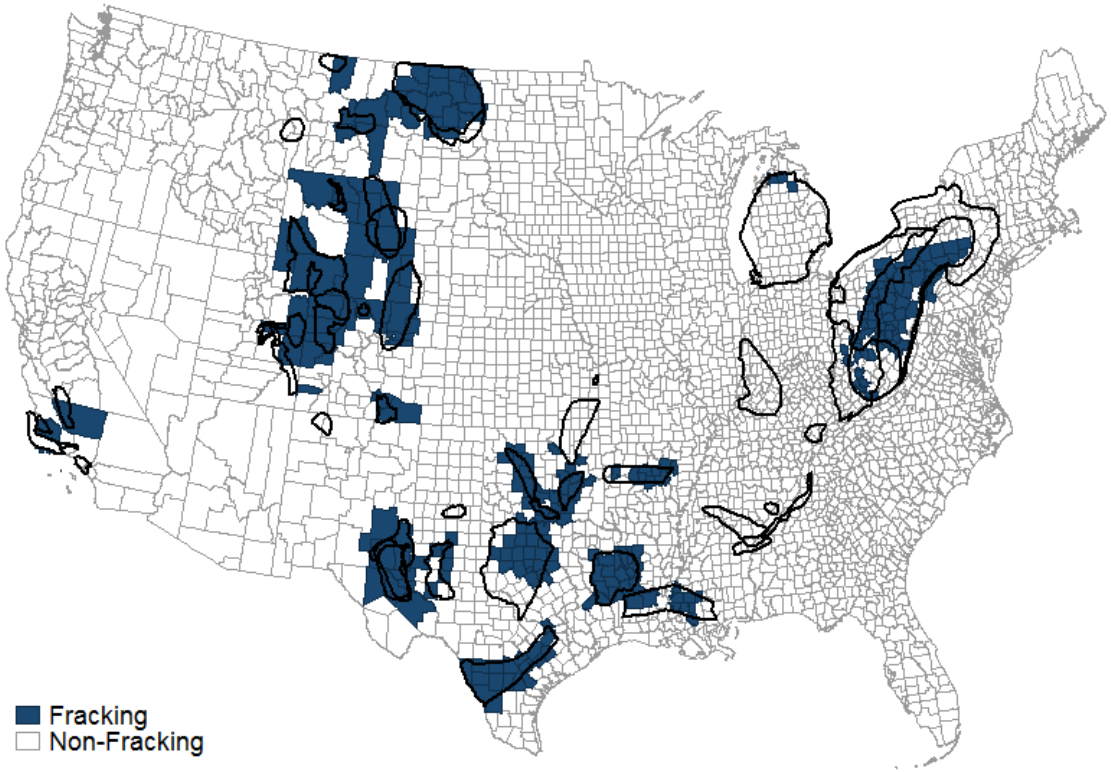


Figure I

Fracking Counties and Shale Plays

Notes: Any county with production from fracking wells between 2000 and 2012 is labeled as a fracking county. Shale play boundaries are outlined in black.

Source: Author's calculations constructed from DrillingInfo well level data. Shale play boundaries are from the EIA.

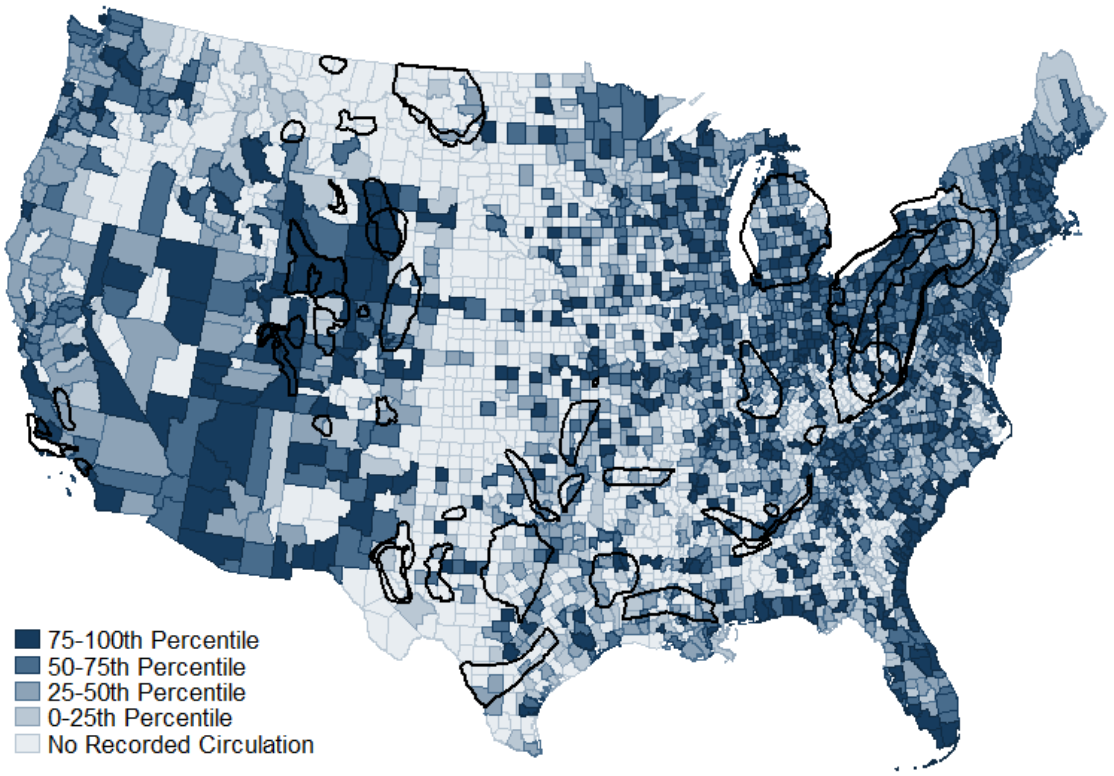


Figure II

County-level Circulation of USA TODAY between 2005 and 2008

Notes: Location of shale plays outlined in black.

Source: Author's calculations using annual county-level circulation rates averaged between 2005 and 2008 obtained from the Alliance of Audited Media.

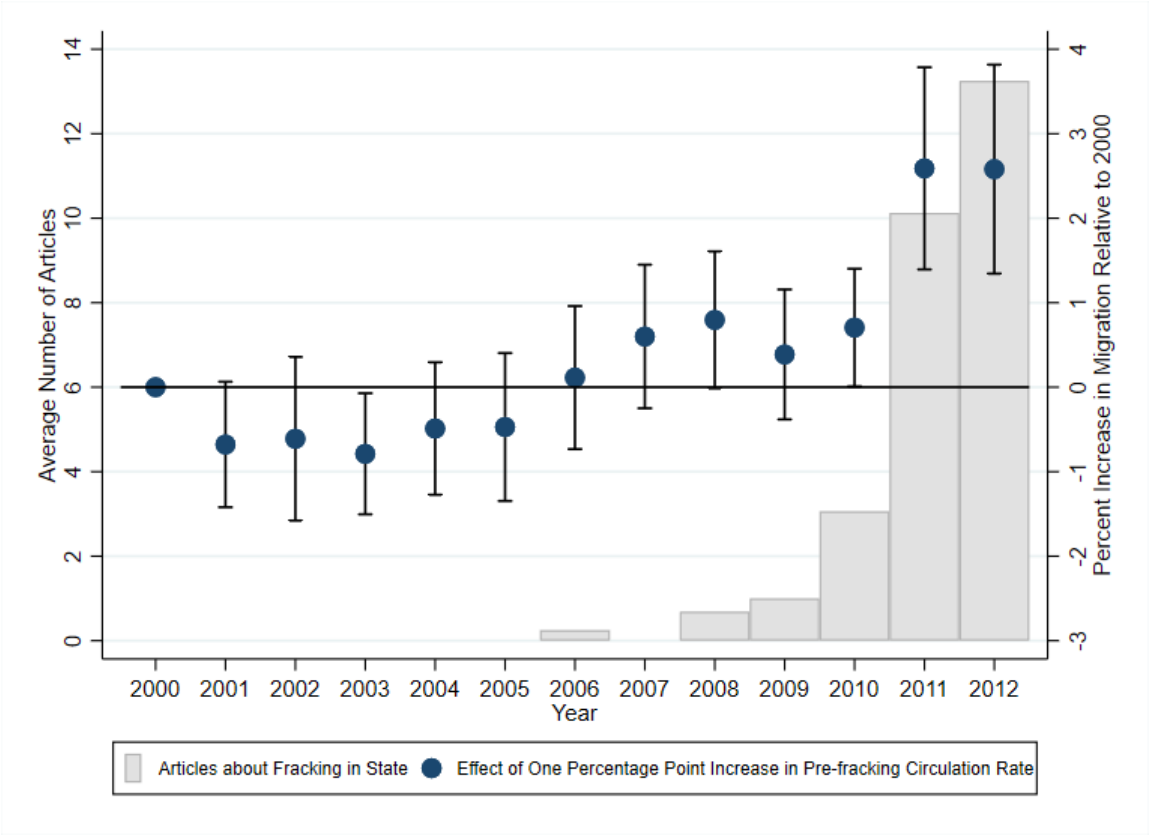


Figure III

Trends in Migration by Pre-fracking Circulation

Notes: For each origin, the pre-fracking circulation rate is the weighted average of the pre-fracking circulation of the *USA TODAY*, *New York Times*, and *Wall Street Journal*, where weights are the share of the total articles about fracking in the destination state in each newspaper. This measure captures the extent to which an origin will eventually be exposed to fracking news. This measure is then interacted with year indicators. The year 2000 is treated as the base year. The inverse hyperbolic sine of the number of migrants is then regressed on this set of interactions along with origin-destination pair effects and destination-by-year fixed effects, as in the main specification, to trace out the effect of a one percentage point increase in the pre-fracking circulation rate on migration, as a percent. The coefficients on these year interactions are interpreted as the marginal effect of a one percentage point increase in the pre-fracking circulation rate on migration flows in that given year and are plotted for each year on the right axis, to look at trends by differences in eventual exposure. Standard errors from the regressions are corrected for clustering at the origin DMA level. For reference, the average number of articles about fracking in each state is also plotted for each year in bars on the left axis.

Source: Author's calculations using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and migration flows from the IRS SOI.

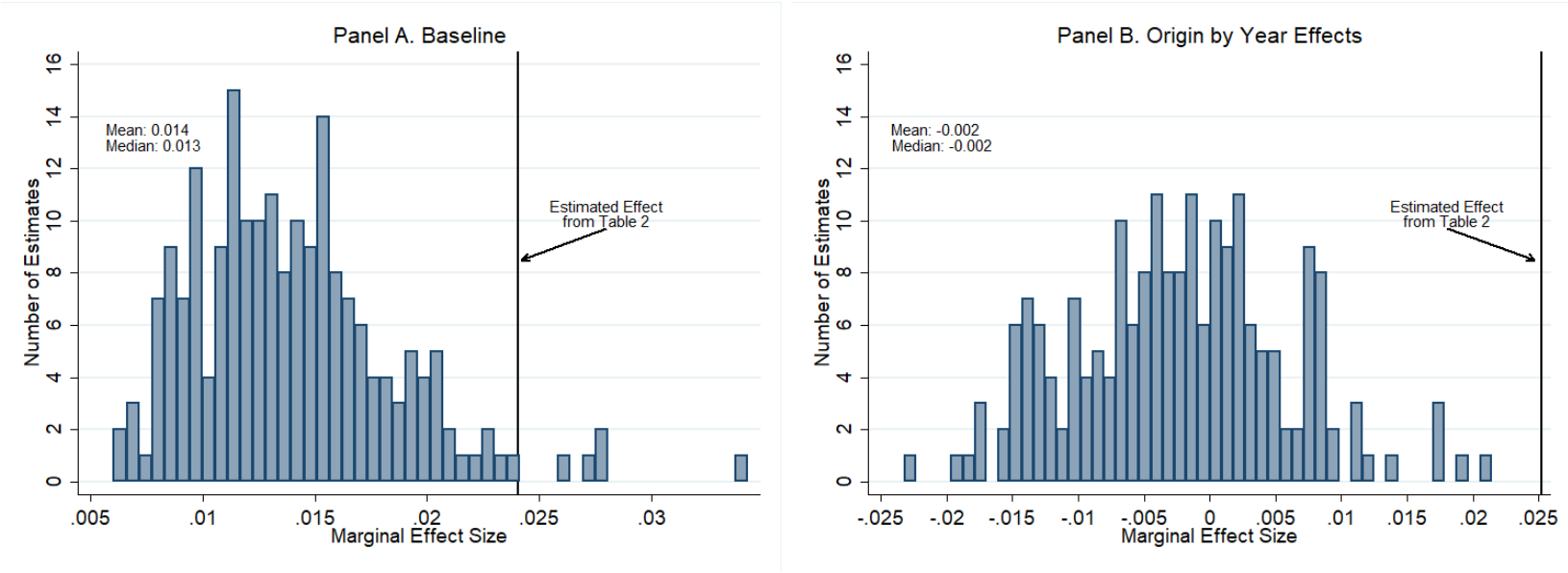


Figure IV

Migration Response to Randomly Assigned News about Fracking in a Different State

Notes: Each state is randomly assigned the fracking news exposure of a different state, and then the inverse hyperbolic sine of migration is regressed on a quadratic of this randomly assigned news exposure, similar to the baseline regression in equation (10). The histogram of estimated effects from the baseline model for 200 regressions are plotted in Panel A. For some states the trends in news exposure are similar, and across all 200 regressions the average correlation between actual news exposure and randomly assigned news coverage was 0.44. Panel B. repeats the same 200 regressions but includes origin by year effects. This exploits variation in news coverage across destinations within an origin, relying on destination state specific deviations in news exposure.

Source: Author's calculation from 200 regressions of randomly assigned news exposure on the inverse hyperbolic sine of migration using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and county to county migration flows from the IRS SOI.

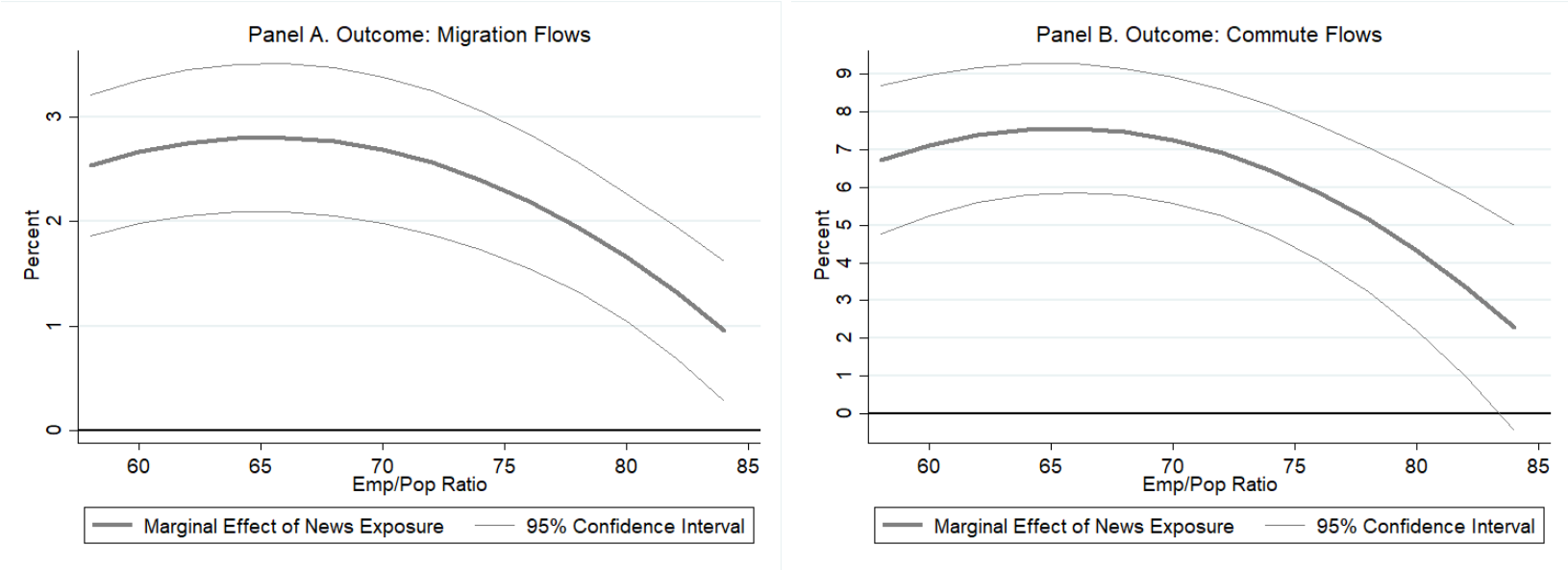




Figure V

Heterogeneous Impacts of Newspaper Exposure by Origin Employment to Population Ratio in $t-1$

Notes: Marginal impact of newspaper exposure calculated by interacting a quadratic in newspaper exposure and a quadratic of lagged employment to population ratio at the origin. Approximately the 10th to 90th percentile of the employment to population ratio are plotted. Standard errors are calculated using the delta method.

Source: Author's calculations using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, migration flows from the IRS SOI, and county employment to population ratio constructed from BLS QCEW data.

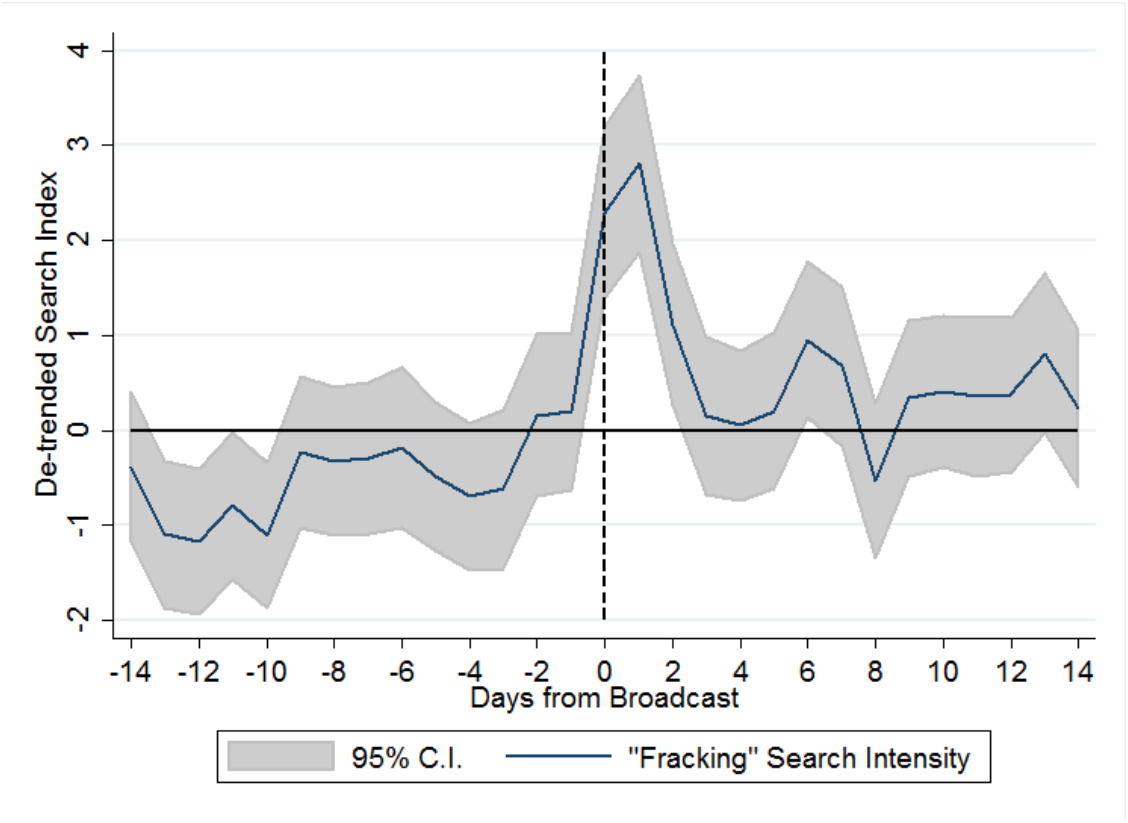


Figure VI

Google Search Interest in "Fracking" After TV News Broadcasts

Notes: Plot depicts the average daily search index for the term "fracking" by DMA before and after 17 TV broadcast mentioning fracking or shale gas between 2006 and 2012 as recorded by the Vanderbilt Television News Archive. Search intensity is de-trended by removing day of week and search (DMA by four week publication window) specific effects. To be consistent with other analysis in the paper, one broadcast from CNN and one broadcast from Fox News are excluded. Four days prior to a news broadcast on January 28, 2012, President Barack Obama mentioned shale gas exploration due to fracking in the State of the Union Address. Four days prior to a news broadcast on January 4, 2012, there was an earthquake in Ohio that reporters linked to fracking. For both of these event I include indicator variables for the next four days. Excluding these controls does not significantly change the daily average search index time series (see Figure A5). Standard errors are clustered at the search level.

Source: Source: Author's calculations using daily search indices from Google Trends.

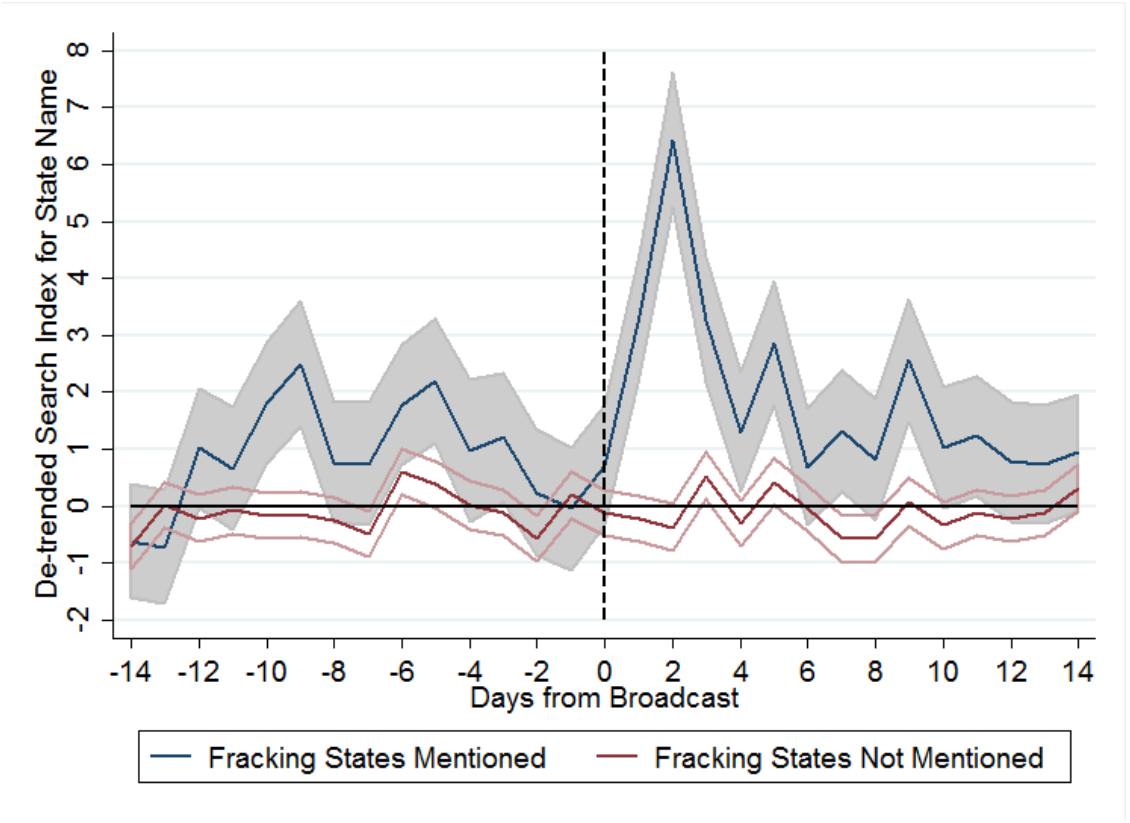


Figure VII

Google Search Interest in the Names of Fracking States Mentioned in TV News Broadcasts

Notes: Plot depicts the average daily search index for the name of the state by state before and after 14 TV broadcast mentioning fracking or shale gas and a specific state between 2006 and 2012 as recorded by the Vanderbilt Television News Archive. Search intensity is de-trended by removing day of week and search (DMA by four week publication window) specific effects. To be consistent with other analysis in the paper, one broadcast from CNN and one broadcast from Fox News are excluded. Additional control indicators are also included for specific high publicity state-specific events that fall in the search period window, such as the earthquakes, wildfires, special elections, and major sporting events. Excluding these controls does not significantly change the daily average search index time series (see Figure A5). For reference, the search intensity for fracking states *not* mentioned in the news broadcast is also plotted with 95 percent confidence intervals. Standard errors are clustered at the search level.

Source: Source: Author's calculations using daily search indices from Google Trends.

## For Online Publication: Appendix A. Additional Tables and Figures

Table A1

Content of Newspaper Articles

| | Share of Articles that Mention | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Jobs References*[1] | *"boom"* | *"growth"* | *Pollution References*[2] | *"health"* | *"danger"* | *"earthquake"* | *Total Articles* |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *All Articles* | 0.17 | 0.24 | 0.14 | 0.39 | 0.20 | 0.12 | 0.06 | 562 |
| *Mention State:* | | | | | | | | |
| *Arkansas* | 0.29 | 0.50 | 0.0 | 0.64 | 0.43 | 0.07 | 0.36 | 14 |
| *California* | 0.28 | 0.41 | 0.34 | 0.41 | 0.31 | 0.13 | 0.03 | 32 |
| *Colorado* | 0.10 | 0.39 | 0.23 | 0.65 | 0.28 | 0.13 | 0.06 | 31 |
| *Louisiana* | 0.32 | 0.47 | 0.26 | 0.58 | 0.21 | 0.26 | 0.05 | 19 |
| *Michigan* | 0.57 | 0.43 | 0.43 | 0.29 | 0.0 | 0.14 | 0.14 | 7 |
| *Mississippi* | 0.0 | 0.50 | 0.0 | 0.75 | 0.75 | 0.50 | 0.0 | 4 |
| *Montana* | 0.67 | 0.67 | 0.67 | 0.17 | 0.17 | 0.0 | 0.0 | 6 |
| *New Mexico* | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2 |
| *North Dakota* | 0.29 | 0.57 | 0.25 | 0.46 | 0.29 | 0.11 | 0.04 | 28 |
| *Ohio* | 0.33 | 0.36 | 0.12 | 0.55 | 0.39 | 0.21 | 0.24 | 33 |
| *Oklahoma* | 0.28 | 0.60 | 0.36 | 0.72 | 0.32 | 0.28 | 0.12 | 25 |
| *Pennsylvania* | 0.30 | 0.41 | 0.14 | 0.66 | 0.30 | 0.23 | 0.07 | 91 |
| *Texas* | 0.22 | 0.42 | 0.27 | 0.41 | 0.24 | 0.11 | 0.09 | 105 |
| *Utah* | 0.20 | 0.40 | 0.60 | 0.40 | 0.60 | 0.20 | 0.20 | 5 |
| *West Virginia* | 0.25 | 0.46 | 0.13 | 0.63 | 0.54 | 0.13 | 0.04 | 24 |
| *Wyoming* | 0.04 | 0.36 | 0.14 | 0.68 | 0.25 | 0.21 | 0.04 | 28 |

Notes: Newspaper content for articles between 2008 and 2012 obtained through LexisNexis, for the *New York Times*, *USA TODAY*, and *Wall Street Journal*. Not all articles reference a state, and some articles reference multiple states. Search terms are truncated to include various tenses and included both capitalized and lower case. [1] Jobs References include the following search terms: "new job", "creat∼ + job", "low + unemploy∼", "hire/hiring". [2] Pollution References include the following search terms: "contaminat∼" and "pollut∼".

Table A2

County Characteristics by the *New York Times* Pre-Fracking Circulation Rate

| | Pre-2009 Circulation Rate of the *New York Times* | | | | |
|---|---|---|---|---|---|
| | County Characteristics in 2000 | | Change from 2000 to 2010 | | Predicted Difference from 25th to 75th Percentile |
| | Below Median (1) | Above Median (2) | Below Median (3) | Above Median (4) | (5) |
| *Migrants to fracking areas (Pct. of Population)* | 0.11 | 0.08 | -0.02 | -0.01 | 0.00 |
| *Employment to Population (16+)* | 55.62 | 59.22 | -0.6 | -1.62 | -0.15** |
| *Unemployment Rate* | 3.50 | 3.23 | 0.82 | 1.47 | 0.17*** |
| *Median Household Income* | 31,652 | 39,018 | 8,080 | 9,173 | 1,119*** |
| *Percent in Poverty* | 16.05 | 12.04 | 1.11 | 1.90 | 0.07 |
| *Percent White* | 84.49 | 85.0 | -1.26 | -1.84 | -0.41*** |
| *Percent Black* | 9.40 | 8.82 | 0.11 | 0.16 | 0.10*** |
| *Percent Hispanic* | 5.35 | 6.24 | 1.83 | 2.32 | 0.28*** |
| *Percent Other Race* | 6.11 | 6.19 | 1.15 | 1.68 | 0.31*** |
| *Percent Population 20-34* | 17.55 | 18.72 | -0.40 | -0.82 | -0.13*** |
| *Percent Population 35-64* | 38.28 | 39.04 | 1.65 | 1.64 | 0.03 |
| *Percent Population Over 64* | 15.65 | 13.95 | 1.08 | 1.28 | 0.07*** |
| *Percent Households Renting* | 25.37 | 26.40 | 1.75 | 1.69 | -0.21*** |
| *Number of Counties* | 1,426 | 1,412 | 1,426 | 1,412 | 2,838 |

Notes: Migration data from the IRS Statistics of Income. Other county characteristics obtained through American FactFinder from the 2000 Census and 2010 Census and 5-Year American Community Survey. Circulation data for the *New York Times* from the Alliance for Audited Media. The county level median pre-2009 circulation rate of the *New York Times* was 0.32 percent, ranging from 0 to 3.29 percent. Circulation of the *New York Times* and the *Wall Street Journal* are highly correlated ($\rho = 0.8$), and characteristics look similar by circulation of the *Wall Street Journal*. Median Household Income is reported in current dollars. Column (5) reports the predicted change in the characteristic between 2000 to 2010 when pre-2009 circulation increases from the 25th to the 75th percentile. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A3

Impact of Destination State Specific Newspaper Exposure on Cross-County Commuting to Fracking Regions

| | Inverse Hyperbolic Sine of the Number of Cross-County Commute Jobs$_{oSt}$ | | | | | |
|---|---|---|---|---|---|---|
| | By Monthly Earnings | | | By Broad Industry | | |
| | ≤$1,250 | $1,250–$3,333 | ≥$3,333 | Goods Producing | Trade and Transportation | Other Industry |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *National Newspaper Exposure*$_{oSt}$ | 0.034*** | 0.045*** | 0.045*** | 0.025*** | 0.029*** | 0.058*** |
| | (0.005) | (0.006) | (0.006) | (0.005) | (0.004) | (0.007) |
| *National Newspaper Exposure*$^2_{oSt}$ | -0.001*** | -0.001*** | -0.001*** | -0.001*** | -0.001*** | -0.001*** |
| | (0.0002) | (0.0003) | (0.0002) | (0.0002) | (0.0002) | (0.0003) |
| *Dependent Mean* | 0.7 | 11.2 | 11.7 | 6.1 | 8.1 | 17.2 |
| *Observations* | 499,440 | 499,440 | 499,440 | 499,440 | 499,440 | 499,440 |

Notes: Data from the LEHD Origin-Destination Employment Statistics (LODES), LexisNexis newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. The level of observation is the origin county by destination state by year from 2002 to 2012. LODES data is only available starting in 2002. Origin counties with any fracking production or in the New York City designated market area are excluded. The variable *National Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. For comparison, circulation of the *USA TODAY* was 4.5 percent at the 95th percentile. Earnings and Industry classifications are pre-defined in the LODES data. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A4

Robustness to Functional Form

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | | | Inverse Hyperbolic Sine of the Number of Cross-County Commute Jobs$_{oSt}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear (1) | Quadratic (2) | Cubic (3) | IHS (4) | Linear (5) | Quadratic (6) | Cubic (7) | IHS (8) |
| *National Newspaper Exposure*$_{oSt}$ | 0.014*** | 0.025*** | 0.022*** | | 0.037*** | 0.068*** | 0.099*** | |
| | (0.003) | (0.004) | (0.005) | | (0.007) | (0.009) | (0.011) | |
| *National Newspaper Exposure*$^2_{oSt}$ | | -0.001*** | -0.0002 | | | -0.002*** | -0.006*** | |
| | | (0.0001) | (0.001) | | | (0.0004) | (0.001) | |
| *National Newspaper Exposure*$^3_{oSt}$ | | | 0.00001 | | | | 0.0001*** | |
| | | | (0.00001) | | | | (0.00002) | |
| *Inverse Hyperbolic Sine of National Newspaper Exposure*$_{oSt}$ | | | | 0.046*** | | | | 0.155*** |
| | | | | (0.007) | | | | (0.019) |
| *Dependent Mean (in Levels)* | 7.6 | 7.6 | 7.6 | 7.6 | 31.4 | 31.4 | 31.4 | 31.4 |
| *Observations* | 590,224 | 590,224 | 590,224 | 590,224 | 499,440 | 499,440 | 499,440 | 499,440 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis, and the Alliance for Audited Media. The level of observation is the origin county by destination state by year. *National Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. The inverse hyperbolic sine approximates a natural log transformation, but is defined for values of zero. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A5

Sensitivity to Sample

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | | | Inverse Hyperbolic Sine of the Number of Cross-County Commute Jobs$_{oSt}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | Include NYC DMA (1) | Include Fracking Origins (2) | Exclude Top One Percent of Exposure (3) | Exclude Zero Exposure (4) | Include NYC DMA (5) | Include Fracking Origins (6) | Exclude Top One Percent of Exposure (7) | Exclude Zero Exposure (8) |
| *National Newspaper Exposure*$_{oSt}$ | 0.013*** | 0.025*** | 0.031*** | 0.022*** | 0.045*** | 0.067*** | 0.123*** | 0.041*** |
| | (0.004) | (0.003) | (0.006) | (0.003) | (0.006) | (0.009) | (0.013) | (0.007) |
| *National Newspaper Exposure*$^2_{oSt}$ | -0.0001** | -0.001*** | -0.0003 | -0.001*** | -0.0004*** | -0.002*** | -0.007*** | -0.001** |
| | (0.0001) | (0.0001) | (0.001) | (0.0001) | (0.0001) | (0.0003) | (0.001) | (0.0003) |
| *Dependent Mean (in Levels)* | 7.7 | 12.2 | 6.1 | 11.4 | 31.2 | 62.2 | 26.9 | 49.1 |
| *Observations* | 596,256 | 639,840 | 505,664 | 160,350 | 504,544 | 541,504 | 427,840 | 160,350 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis, and the Alliance for Audited Media. The level of observation is the origin county by destination state by year. *National Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A6

Accounting for Censoring: Impact of Destination State Specific Newspaper Exposure on Migration Flows

| | Inverse Hyperbolic Sine of the Number of Migrating Tax Units$_{oSt}$ | | Number of Migrating Tax Units$_{oSt}$ | | Over 10 Migrating Tax Units$_{oSt}$ | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ |
|---|---|---|---|---|---|---|
| | As Reported | Lower Bound: Replace 0 with 9 | As Reported | Lower Bound: Replace 0 with 9 | | Positive Flows in All Years |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *National Newspaper Exposure*$_{oSt}$ | 0.022*** | 0.001*** | 0.838*** | 0.732*** | 0.005*** | 0.018** |
| | (0.003) | (0.0003) | (0.222) | (0.209) | (0.001) | (0.007) |
| *National Newspaper Exposure*$^2_{oSt}$ | -0.001*** | -0.00002** | -0.019** | -0.016* | -0.0002*** | -0.0004** |
| | (0.0001) | (0.00001) | (0.010) | (0.009) | (0.00003) | (0.0002) |
| *Dependent Mean (in Levels)* | 4 | 145 | 4 | 145 | 0.03 | 348.7 |
| *Observations* | 590,224 | 590,224 | 590,224 | 590,224 | 590,224 | 12,092 |

Notes: Data obtained from the IRS Statistics of Income, LexisNexis Newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. The level of observation is the origin county by destination state by year from 2000 to 2012. The variable *National Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. For comparison, circulation of the *USA TODAY* was 4.5 percent at the 95th percentile. In Columns (1) and (2) the outcome is the inverse hyperbolic sine of migrating tax units (rather than migrants). Censored values are assigned a value of 0 in Column (1), and assigned a value of 9 in Column (2), to provide a lower bound. In Columns (3) and (4) the outcome is the number of migrating tax units in levels, to account for the fact that percentages are not comparable when censored values are reassigned a value of 9. The outcome in Column (5) is an indicator that equals one if there were over 10 migrating tax units. During the sample period, flows with less than 10 returns were censored, and this outcome captures transitions across the censoring threshold. The outcome in Column (6) is the inverse hyperbolic sine of migrants for a subsample of origin/destination pairs that reported positive flows in all years. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A7

Advertising Effects of News Exposure: Market Expanding or Share Stealing

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | Inverse Hyperbolic Sine of the Number of Cross-County Commute Jobs$_{oSt}$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *National Newspaper Exposure*$_{oSt}$ | 0.027*** | 0.027*** | 0.041*** | 0.049*** |
| | (0.004) | (0.003) | (0.007) | (0.007) |
| *National Newspaper Exposure*$^2_{oSt}$ | -0.001*** | -0.001*** | -0.001*** | -0.001*** |
| | (0.0001) | (0.0001) | (0.0003) | (0.0003) |
| *All States Newspaper Exposure*$_{ot}$ | 0.0003 | | 0.005*** | |
| | (0.0003) | | (0.001) | |
| *All States Newspaper Exposure*$^2_{ot}$ | -0.00001** | | -0.00003*** | |
| | (0.000003) | | (0.00001) | |
| *Max. State Newspaper Exposure*$_{ot}$ | | 0.001 | | 0.015*** |
| | | (0.001) | | (0.004) |
| *Max. State Newspaper Exposure*$^2_{ot}$ | | -0.0001** | | -0.0004*** |
| | | (0.0001) | | (0.0002) |
| *Dependent Mean (in Levels)* | 7.6 | 7.6 | 31.4 | 31.4 |
| *Observations* | 590,224 | 590,224 | 499,440 | 499,440 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis, and the Alliance for Audited Media. The level of observation is the origin county by destination state by year. *Newspaper Exposure*$_{oSt}$ is scaled such that a one unit increase represents the impact of one additional news story in a newspaper with a 5 percent circulation rate. *All States Newspaper Exposure*$_{ot}$ is the total news exposure for all 16 destination states within an origin year, to determine if news about fracking in general affects migration. *Max. States Newspaper Exposure*$_{ot}$ is the highest level of news exposure across all 16 destination state within an origin year, to determine if higher news exposure leads to shifting away from other fracking destinations. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A8

Heterogeneity by Newspaper: Impact of Newspaper Exposure on Migration and Commuting to Fracking Regions

| | Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$ | | Inverse Hyperbolic Sine of the Number of Cross-County Commuting Jobs$_{oSt}$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *USA TODAY Exposure*$_{oSt}$ | 0.027*** | 0.021*** | 0.081*** | 0.030*** |
| | (0.004) | (0.004) | (0.013) | (0.007) |
| *USA TODAY Exposure*$^2_{oSt}$ | -0.001*** | -0.001*** | -0.003** | -0.001 |
| | (0.0003) | (0.0002) | (0.001) | (0.001) |
| *New York Times Exposure*$_{oSt}$ | 0.007*** | 0.009*** | 0.019*** | 0.016*** |
| | (0.002) | (0.002) | (0.004) | (0.004) |
| *New York Times Exposure*$^2_{oSt}$ | -0.0001** | -0.0002*** | -0.001*** | -0.0004** |
| | (0.0001) | (0.0001) | (0.0002) | (0.0002) |
| *Wall Street Journal Exposure*$_{oSt}$ | 0.098* | 0.056 | 0.097 | 0.018 |
| | (0.055) | (0.046) | (0.130) | (0.086) |
| *Wall Street Journal Exposure*$^2_{oSt}$ | -0.027 | -0.006 | -0.062 | -0.033 |
| | (0.036) | (0.029) | (0.098) | (0.060) |
| *Origin by Year Effects* | | X | | X |
| *Dependent Mean* | 7.6 | 7.6 | 31.4 | 31.4 |
| *Observations* | 590,224 | 590,224 | 499,440 | 499,440 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis Newspaper transcripts, and newspaper circulation from the Alliance for Audited Media. The level of observation is the origin county by destination state by year from 2000 to 2012 for the migration data and 2002 to 2012 for the commute data. Each newspaper's exposure level is scaled to represent the impact of one additional news story in a county with circulation at the 95th percentile (3.9 percent for the *USA TODAY*, 1.9 percent for the *New York Times*, and 2.4 percent for the *Wall Street Journal*). Controls include the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$). In all specifications origin/destination pair fixed effects and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A9

Heterogeneous Impacts by Origin Employment to Population Ratio

| | Inverse Hyperbolic Sine | | Levels | |
|---|---|---|---|---|
| | Number of Migrants | Number of Commuters | Number of Migrants | Number of Commuters |
| | (1) | (2) | (3) | (4) |
| $Newspaper\ Exposure_{oSt}$ | 0.0270*** | 0.0731*** | 1.7029*** | 5.3651** |
| | (0.0037) | (0.0088) | (0.5103) | (2.2426) |
| $Newspaper\ Exposure^2_{oSt}$ | -0.0007*** | -0.0020*** | -0.0392* | -0.2425** |
| | (0.0001) | (0.0004) | (0.0234) | (0.1048) |
| $Newspaper\ Exposure_{oSt} * Emp/Pop_{ot-1}$ | -0.0006*** | -0.0017** | -0.0100 | -0.1350 |
| | (0.0002) | (0.0006) | (0.0221) | (0.0958) |
| $Newspaper\ Exposure_{oSt} * Emp/Pop^2_{ot-1}$ | -0.0001*** | -0.0002*** | -0.0040*** | -0.0189*** |
| | (0.000) | (0.000) | (0.0013) | (0.0066) |
| $Newspaper\ Exposure^2_{oSt} * Emp/Pop_{ot-1}$ | 0.000 | 0.000 | -0.0009 | 0.0058 |
| | (0.000) | (0.0001) | (0.0011) | (0.0043) |
| $Newspaper\ Exposure^2_{oSt} * Emp/Pop^2_{ot-1}$ | -0.000 | 0.000 | -0.000 | 0.0008** |
| | (0.000) | (0.000) | (0.0001) | (0.0004) |
| $Emp/Pop_{ot-1}$ | 0.0004** | 0.0002 | 0.0052 | -0.1376 |
| | (0.0002) | (0.0006) | (0.0192) | (0.1236) |
| $Emp/Pop^2_{ot-1}$ | 0.000 | -0.000 | 0.0006* | 0.0032 |
| | (0.000) | (0.000) | (0.0004) | (0.0021) |
| *Dependent Mean (in levels)* | 7.635 | 31.23 | 7.635 | 31.23 |
| *Observations* | 544,688 | 499,296 | 544,688 | 499,296 |

Notes: Data obtained from the IRS Statistics of Income, LEHD Origin-Destination Employment Statistics (LODES), LexisNexis, and the Alliance for Audited Media. The level of observation is the origin county by destination state by year. The origin county employment to population ratio (Emp/Pop) is obtained from the BLS, and lagged by one year. Emp/Pop is demeaned, such that the direct effect of newspaper exposure is the effect for a county at the mean employment to population ration (70.9 percent). Controls for the current origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are also included. In all specifications origin/destination pair fixed effects, and destination by year fixed effects are included to control for time-invariant differences across pairs and characteristics of the destination that vary over time. Standard errors are adjusted for clustering at the designated market area to account for correlation across geography and time. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A10

Simulated Impacts of News Exposure Cross-County Commute Flows by Origin Employment to Population Ratio

| | Employment to Population Ratio in $t-1$ ($\mu = 70.9$) | | | | |
|---|---|---|---|---|---|
| | 60 | 65 | 70 | 75 | 80 |
| *Exposure Level* | | | | | |
| *0* | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| *1* | 0.071*** | 0.076*** | 0.073*** | 0.062*** | 0.043*** |
| | (0.010) | (0.009) | (0.009) | (0.009) | (0.011) |
| *5* | 0.315*** | 0.336*** | 0.322*** | 0.274*** | 0.191*** |
| | (0.041) | (0.040) | (0.039) | (0.040) | (0.046) |
| *10* | 0.527*** | 0.566*** | 0.544*** | 0.461*** | 0.319*** |
| | (0.70) | (0.071) | (0.071) | (0.070) | (0.074) |
| *15* | 0.637*** | 0.689*** | 0.664*** | 0.562*** | 0.384*** |
| | (0.089) | (0.099) | (0.102) | (0.098) | (0.093) |
| *Mean Exposure in 2012* | 1.57 | 2.08 | 2.48 | 2.44 | 1.89 |
| *Implied Impact* | 0.110*** | 0.153*** | 0.172*** | 0.144*** | 0.079*** |
| | (0.015) | (0.018) | (0.020) | (0.021) | (0.020) |
| *Maximizing Exposure* | 17.9 | 18.3 | 18.5 | 18.4 | 17.7 |
| *Implied Impact* | 0.654*** | 0.713*** | 0.689*** | 0.582*** | 0.393*** |
| | (0.099) | (0.118) | (0.128) | (0.120) | (0.105) |

Notes: Simulated impacts are obtained for each combination of origin employment to population ratio and exposure level from equation (14), where the outcome is the inverse hyperbolic sine of the number of cross-county commuters. The corresponding coefficients are reported in Appendix Table A9. The maximizing exposure is obtained by setting the first derivative of equation (14) with respect to newspaper exposure equal to zero and solving for the maximizing exposure for the specified employment to population ratio. This value is rounded down to the nearest whole number. The implied impact is the corresponding effect of the maximizing exposure level. For reference, the mean exposure level in 2012 for origin counties with employment to population ratios within 2.5 percent of the specified threshold. An employment to population ratio of 60 roughly corresponds to the 15th percentile while a ratio of 80 corresponds to roughly the 85th percentile. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A11

State-level Events Controlled for in Google Trends State Specifications

| Date<br>(1) | Event<br>(2) | States<br>(3) |
|---|---|---|
| November 7-9, 2006 | Four-way Texas Gubernatorial Election | Texas |
| September 9-10, 2010 | San Bruno Pipeline Explosion | California |
| September 6-11, 2011 | 2011 Texas Wildfires | Texas |
| Dec. 31, 2011-Jan. 2, 2012 | 4.0 Earthquake in Eastern Ohio | Ohio |
| February 19-20, 2012 | Texas A&M v. Oklahoma State Basketball Game | Texas, Oklahoma |
| February 27-28, 2012 | Chardon High School shooting | Ohio |
| March 7, 2012 | Ohio Primary Elections | Ohio |
| May 21-22, 2012 | Arkansas Primary Elections | Arkansas |

Notes: High-level interest events that are closely tied to a specific state during the Google Trend search windows are controlled for to increase precision. Indicators that equal one for each of the listed dates for the destination state listed are included in the state name Google Trend analysis.

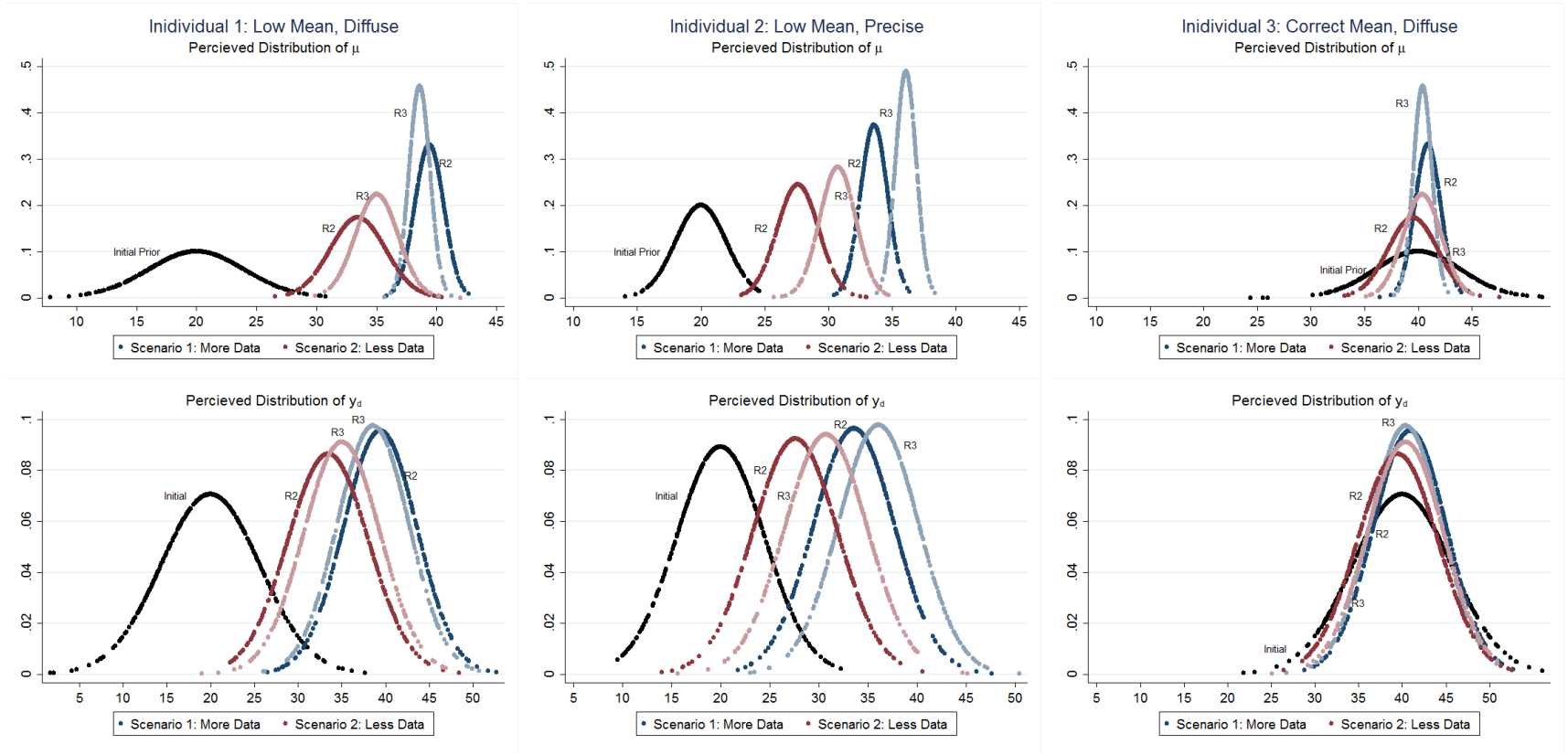


Figure A1

Model Simulations: Information and Bayesian Updating

Notes: Simulated data points from the distributions of $\mu_d$ and $y_d$ are presented for three separate individuals in two separate scenarios. Individual 1 had a diffuse prior with a low mean, individual 2 had a more precise prior with a low mean, and individual 3 had a diffuse prior with a correct mean. In scenario 1 the individual viewed ten data points from the true distribution of $y_d$ in each round (R2 and R3), and updates the posterior probability accordingly. In scenario 2 the individual views only 2 data points and updates the posterior. The initial prior and two additional iterations are shown.

Source: Author's calculations.

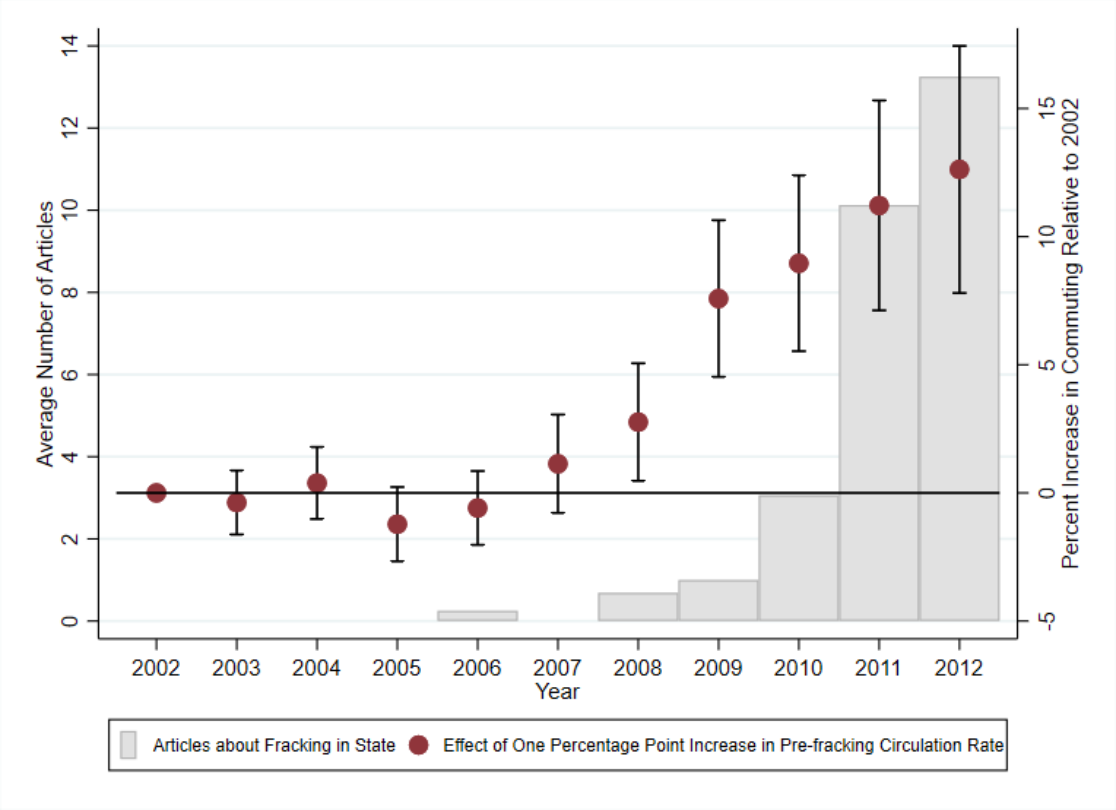


Figure A2

Trends in Commuting by Pre-fracking Circulation

Notes: For each origin, the pre-fracking circulation rate is the weighted average of the pre-fracking circulation of the *USA TODAY*, *New York Times*, and *Wall Street Journal*, where weights are the share of the total articles about fracking in each newspaper. This measure is then interacted with year indicators. Commuting data is only available starting in 2002, so 2002 is treated as the base year. The inverse hyperbolic sine of the number of cross-county commuting jobs is then regressed on this set of interactions along with origin-destination pair effects and destination-by-year fixed effects, as in the main specification, to trace out the effect of a one percentage point increase in the pre-fracking circulation rate on migration, as a percent. The marginal effect of one unit of a one percentage point increase is converted to percentage points and plotted for each year on the right axis, to look at trends by pre-fracking circulation. Standard errors from the regressions are corrected for clustering at the origin DMA level. For reference, the average number of articles about fracking in each state is also plotted for each year in bars on the left axis.

Source: Author's calculations using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and commuting flows from the LODES.

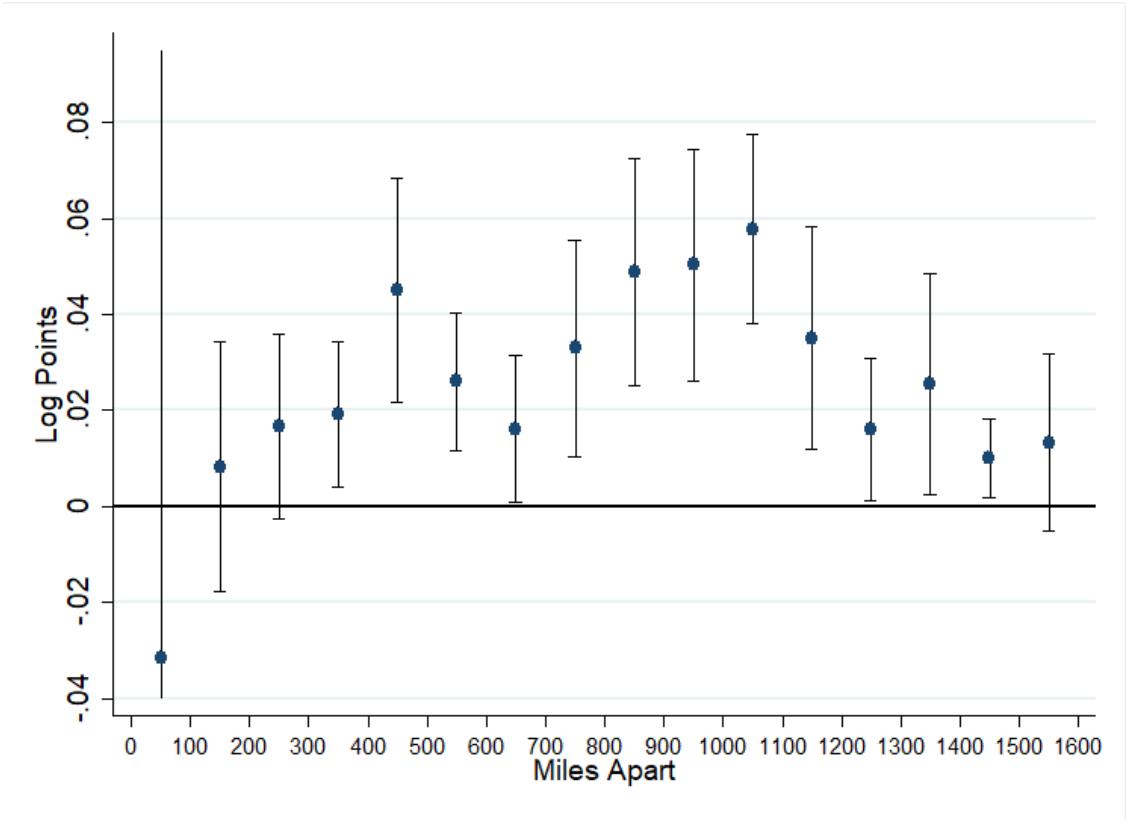


Figure A3

Marginal Impact of Newspaper Exposure on Migration by Origin to Destination Distance

Notes: Coefficients and confidence intervals plotted for the marginal effect of newspaper exposure on migration flows from equation (10), estimated over one hundred mile bins. Standard errors calculated using the Delta Method.

Source: Author's calculations using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and migration flows from the IRS SOI.

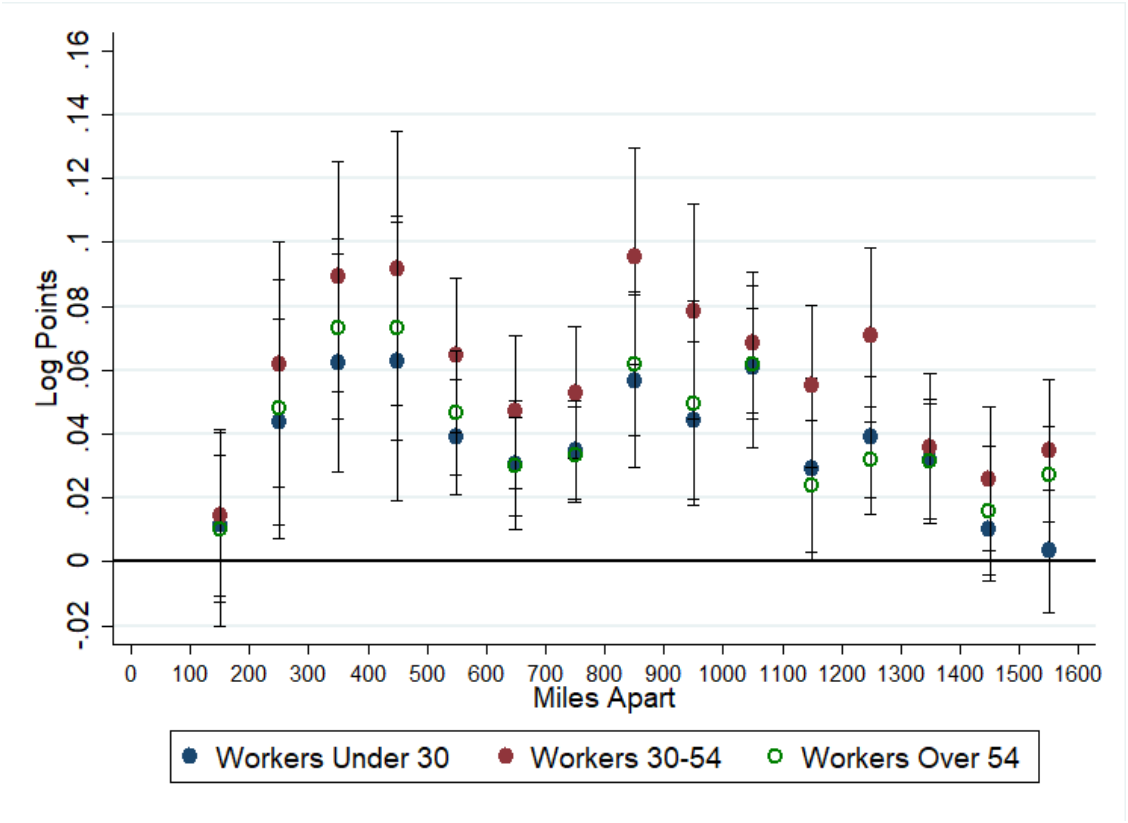


Figure A4

Marginal Impact of Newspaper Exposure on Commuting by Age and Origin to Destination Distance

Notes: Coefficients and confidence intervals plotted for the marginal effect of newspaper exposure on cross-county commute flows from equation (13), estimated over one hundred mile bins, separately by worker age. Standard errors calculated using the Delta Method. The estimated marginal impact for pairs less than one hundred miles apart for each age group are highly negative, at -0.06 (0.04), -0.10 (0.06), and -0.16 (0.06), respectively.

Source: Author's calculations using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and commute flows from the LEHD Origin-Destination Employment Statistics (LODES).

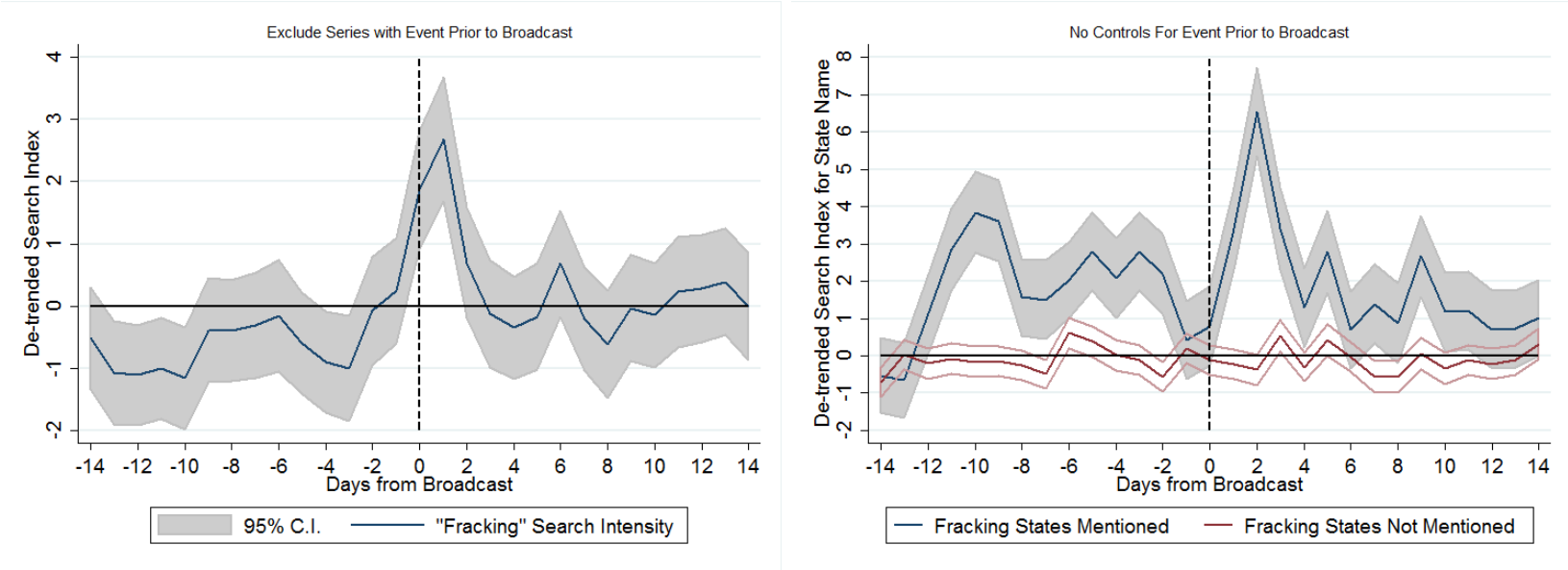


Figure A5

Google Search Interest: Exclude Event Controls

Notes: Plot depicts the same average daily search index for "fracking" and specific states mentioned as in Figures VI and VII, but does not include controls for high publicity events that occurred during the search window and were either related to fracking, or a specific state. Standard errors are clustered at the search level.

Source: Author's calculations using daily search indices from Google Trends.

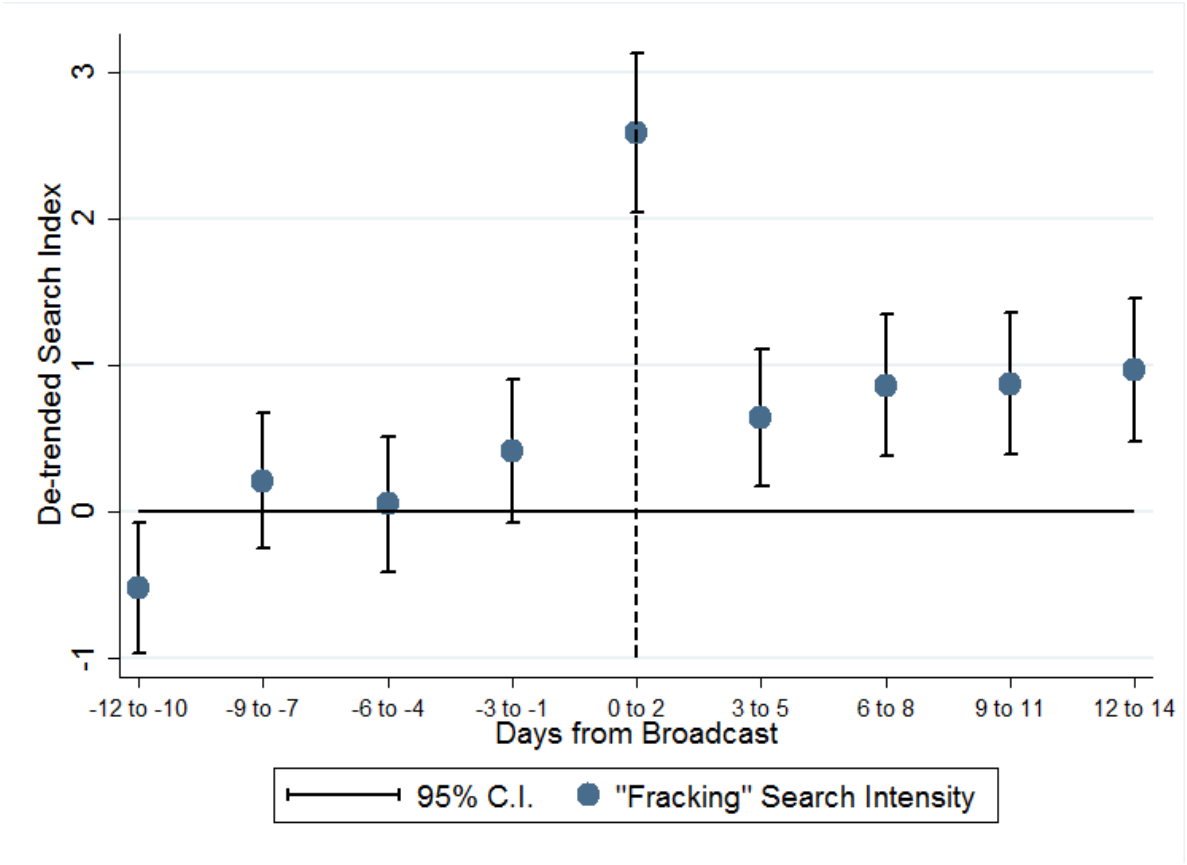


Figure A6

Google Search Interest in "Fracking" After TV News Broadcasts, 3 Day Bins

Notes: Plot depicts the average daily search index for the term "fracking" by DMA before and after 17 TV broadcast mentioning fracking or shale gas between 2006 and 2012 as recorded by the Vanderbilt Television News Archive, as in Figure VII, but groups days into 3-day bins. Search intensity is de-trended by removing day of week and search (DMA by four week publication window) specific effects. To be consistent with other analysis in the paper, one broadcast from CNN and one broadcast from Fox News are excluded. Four days prior to a news broadcast on January 28, 2012, President Barack Obama mentioned shale gas exploration due to fracking in the State of the Union Address. Four days prior to a news broadcast on January 4, 2012, there was an earthquake in Ohio that reporters linked to fracking. For both of these event I include indicator variables for the next four days. When these events are not controlled for, there is a marginally significant increase in search intensity in the days prior to the broadcast. Standard errors are clustered at the search level.

Source: Author's calculations using daily search indices from Google Trends.

# For Online Publication: Appendix B. Supplemental Data Appendix

Below I describe each of the key datasets used in my analysis, as well as important characteristics of data construction.

## DrillingInfo Oil and Gas Production Data

Well level information on drilling date, lease agreements, location, direction, and geological formation as well as other characteristics are provided through a restricted use data agreement from DrillingInfo. This data is proprietary, and obtained through an academic use agreement with DrillingInfo, available through their academic outreach initiative. DrillingInfo does not indicate if a well is a fracking well, as fracking is a means of stimulating production. To infer wells that are affected by the technological innovation associated with fracking, I use details on drilling direction and well location. Localized fracking booms occurred in part because of the combination of horizontal (directional) drilling and hydraulic fracturing. The DrillingInfo data reports whether a well is horizontally or vertically drilled. In addition, fracking was particularly impactful over shale plays, as these resources were not extractable previously. For this reason I assign non-vertical wells drilled in counties that intersect with shale plays as fracking wells. This production data is then combined with shale play boundary shapefiles provided by the Energy Information Administration to identify wells in shale plays and counties with fracking production.

## Internal Revenue Service Statistics of Income County Flows

The Internal Revenue Service (IRS) Statistics of Income (SOI) division provides annual counts of county-to-county flows. This provides the raw number of tax returns and exemptions that were filed in one county in year $t-1$ and in another county in year $t$. Each year, the IRS provides county-to-county flows of exemptions in a file with two years (e.g., 2002to2003). This represents exemptions that were in one county when filing in 2002 and in another county when filing in 2003. As most people file in the beginning of the year before April, I assign this flow to the year 2002.

Using exemptions to approximate people in a household, I can identify origin-destination county level flows. For privacy purposes, the IRS suppresses county pairs that have fewer than ten returns move in each year. As such, county pairs that have small, positive flows will be recorded as zero. This potentially introduces measurement error. For this reason, I also consider lower bound specifications where all county to county flows of zero are replaced with nine. This operates under the assumption that all flows had at least nine households move, which is likely an extreme over-estimate. In 2013 the suppression threshold increased to 20 households. This led to considerably more suppression. For this reason I limit my analysis to 2012.

Unfortunately, the IRS county to county flows only provide aggregate numbers, and do not break up the migration levels by demographic characteristics (gender, marital status, education). As such, I am unable to use the IRS measures to look at differences across demographics. The only measure provided is the total adjusted gross income for all of the moved- returns. This is the adjusted gross income in the first year, but only the average for all movers in the county pair is provided.

The IRS data does not capture every move from one county to another. Low income individuals and households are not required to file a tax return, and thus might be under represented in the data. It is likely that individuals that move to fracking areas will earn well beyond the filing threshold after moving, but they might not have been required to file in the previous year. If there are individuals that did not file in the first year, but moved in response to fracking and filed in the second year, my estimates would be attenuated.

Households that file for extensions past September will also not be included in the data, which might exclude very high income households with complicated returns.

The IRS data does not capture every move from one county to another. Low income individuals and households are not required to file a tax return, and thus might be under represented in the data. It is likely that individuals that move to fracking areas will earn well beyond the filing threshold after moving, but they might not have been required to file in the previous year. If there are individuals that did not file in the first year, but moved in response to fracking and filed in the second year, my estimates would be attenuated. Households that file for extensions past September will also not be included in the data, which might exclude very high income households with complicated returns.

### Alliance for Audited Media Newspaper Circulation Data

Newspaper circulation rates between 2005 and 2008 were obtained through a temporary academic membership at the Alliance for Audited Media (AAM). These circulation rates are provided in PDFs, which I scraped to collect county level estimates. In some cases the scrap was unable to identify the circulation rate, so hand corrections were made.

The AAM conducts regular (annual or biannual) audits of newspapers and collects circulation rates, along with other measures such as prices. This circulation rates includes the number of copies sold on the audit date and the number of copies as a percent of households for each county with over 25 copies. Counties with fewer than 25 copies sold are assigned a zero value. For most newspapers, circulation rates are reported at the county level. However, for the *New York Times* and *Wall Street Journal* these rates were provided at the DMA-level. For my county level analysis I assign each county the DMA-level, which reduces the variation and adds measurement error. However, as seen in Table 3.5, DMA-level estimates provide similar conclusions. A small subset of local newspapers that reported about fracking do not have AAM audits. For these newspapers, which often only distributed to one or two counties, statistics about local circulation was compiled from online searches. For local newspapers that were not audited annually, the intermittent values were imputed through linear interpolation. The three national newspapers report circulation every year.

### TV and Cable Factbook TV Circulation Data

TV circulation data is taken from the Television Cable Factbook for 2008 and 2016. The Factbook contains information on local TV stations as well as DMA-level circulation as reported by Nielsen's. TV circulation is reported at the DMA level for each TV station and includes viewership from both cable and non-cable households. This data is available at the station-level and not specific to news programming. The circulation rate is constructed by dividing total weekly viewership by the total number of households in the DMA. I use average weekly circulation rates throughout my analysis. For each station the "own" DMA and "other" DMA circulation is reported. Because it is not specified what "other" DMA is included, I only include circulation in "own" DMA. This is likely to attenuate the estimated effects. However, for many stations viewership outside the DMA is very low or non-existent. The 2016 circulation rates were obtained through a temporary online membership which provided only the current 2016 circulation rates.

For this reason, I also hired an undergraduate RA to collect circulation rates from the 2008 Factbook. Between 2007 and 2009, TV stations were transitioning from analog to digitally transmitted broadcasts on a market-by-market basis. When a market transitioned, viewers were required to obtain digital reception equipment, and it is unclear how this affected viewership and if 2008 viewership is representative of later years. For this reason I include estimates using both the 2008 and 2016 measures.

### LexisNexis Newspaper Content Data

Newspaper content is collected through LexisNexis by searching on key terms, "frack*", "fracing", and "hydraulic fractur*". I then take the universe of articles, remove non-US sources (e.g., Daily Mail in the UK), and remove articles that only reference things like "Frick and Frack", unrelated acronyms, or last names. I then parse the entire text of these articles for each of the 16 state names (both capitalized and lower cased). References to states in the title of newspapers or place of publication are excluded, (ex: articles published in Colorado are not included as citing Colorado unless there is a reference in the body of the text). I then parse the entire text of the articles for positive and negative terms: "new job", "creat + job", "low + unemploy", "hire", "hiring", "boom", "growth", "earthquak", "environment", "health", "contaminat", "danger", and "pollut". Positive articles are articles that reference at least two positive terms and more positive terms than negative. Negative terms are the opposite. There are "neutral" articles that refer to fewer references that are not included when looking at news content, but have been included in previous specifications. When positive, neutral, and negative news are all included the patterns are similar but less precise.

### Vanderbilt Television News Archive TV News Content

TV news content was pulled from the Vanderbilt Televisions News Archive (VTNA) and includes broadcasts that mention "fracing", "frack", or "shale". The VTNA database contains TV news recordings and transcript abstracts for nightly news broadcasts from the three major news networks (ABC, CBS, and NBC) and the cable news channels CNN and Fox News. The database only includes one hour of programming each day for both cable news outlets. Because the available content of cable news is limited, and circulation rates are only available for the TV networks, I restrict the sample to TV broadcasts from the three major news networks. I parse the transcript abstracts for search terms such as "fracking" and "shale" as well as which state is being discussed. These clips are short often ranging from one to five minutes in length.

### LEHD Origin Destination Employment Statistics Commute Data

The LEHD Origin Destination Employment Statistics (LODES) contains the number of workers for each residence/work place Census block pair. This data is available for all years since 2002, and also provides statistics by broad age (under 30, 30-54, over 54), monthly earnings (under $1,250, $1,250-3,333, over $3,333), and industry (goods, trade/transportation, other) groups. For each Census block I identify the corresponding county, and then aggregate up commute flows to the county to county level. For privacy, some noise is introduced at the Census block level, which likely remains at the county level, although to a lesser extent.