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

Riley Wilson

November 18, 2020

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, concentrated among young, unmarried, less-educated men. Commuting also increased, sentiment of the news matters, and TV news has an impact. Google searches for "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, 435 Crabtree 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, the Chicago Labor Dynamism Conference, and SOLE for useful comments. I gratefully acknowledge funding through the Economic Club of Washington Vernon E. Jordan Jr. doctoral fellowship.

I Introduction

Migration is often seen as a way for people to avoid weak labor markets, and encounter better economic opportunities (Shultz, 1961; Sjastaad, 1962). However, empirically, people are unlikely to move away when labor markets do poorly, even when labor markets with better employment prospects exist (Monras, 2015). As the previous literature notes, 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 can introduce uncertainty and increase the risk associated with migration, information provision can increase the perceived return to moving and change migration behavior. However, the link between information and migration is frequently overlooked in empirical work. 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? I am able to evaluate this relationship by exploiting the unique setting created by the fracking boom.

Over the last 10 to 15 years, the combination of horizontal drilling and hydraulic fracturing has led to localized "fracking booms" and sudden, large increases in both local employment and earnings. These booms have created large, persistent effects across industries (Feyrer et al., 2017; Maniloff & Mastromonaco, 2014), resulting in perceived net benefits (Bartik et al., 2017). This gold rush-like flurry of economic activity has led to numerous newspaper articles and television news broadcasts touting the local economic impacts of fracking or debating its adverse side effects. As such, fracking has created plausibly exogenous labor market improvements that get talked about in the news. The novelty of fracking also introduced new words and vocabulary, making it easy to parse newspaper and TV news to see which sources are talking about fracking, which places they are talking about, what they are saying, and when they are saying it. Combining content about specific fracking destinations with origin-level measures of news circulation, allows me to isolate one particular source of information transmission and estimate its causal impact on migration.

To understand the identifying variation I exploit, consider the following example. In 2012 the *USA TODAY* published six articles about fracking in Pennsylvania. My strategy tests 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 was distributed, relative to counties with historically lower circulation. In essence, this specification holds fixed 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 generalize this in a regression framework by combining all of the news articles about fracking from major national newspapers, with proprietary pre-fracking county-level circulation data. I also include origin by destination fixed effects and destination by year fixed effects. Destination by year fixed effects control for characteristics of the destination that are changing over time and make this a comparison of migration flows from different origins, with different levels of circulation exposure, to the same destination. Origin by destination fixed effects controls for time invariant pair-specific characteristic (such as distance) that might affect migration behavior, but also control for the fact that people in origins with higher circulation might be more educated, more wealthy, or more mobile on average.

I find that exposure to national newspaper news about a particular destination state increases the flow of migrants to fracking counties in that state. During my analysis period the average impact of exposure to fracking news is a 2.4 percent increase in annual origin-destination specific migration flows. Exposure to news about fracking also increases cross county commute flows by 6.6 percent on average, translating into an average annual increase of approximately one to two additional migrants and four additional commuters to the fracking destination from each origin. Although this response is small, it is economically significant given the scope of the "treatment" and the aggregate effects at the destination. In 2012 alone, the national news about local fracking booms increased migration flows to fracking counties by 4.2 percent on average, and increased commute flows by 11.7 percent. As information comes from other source as well, this likely represents a lower bound of the overall effect of information on migration. The migration responses are largest among men, young workers (under 34), the unmarried, and workers with some college but no degree.

One concern with this 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 more income growth over time. If this additional growth in income affects migration decisions, the estimates would be biased. However this does not appear to drive the results. Areas with high and low historic circulation follow similar trends in migration to fracking areas in the pre-period, and only diverge once they are "treated" with the news. The estimates are also insensitive to controlling for time-varying origin level characteristics, like average earnings or unemployment. Furthermore, because there are 16 states involved in fracking, I can include origin by year fixed effects and account for any observable or unobservable characteristics of the origin that are changing over time and affect migration. This controls for the possibility that counties with high circulation might be becoming more educated, wealthy, or dynamic over time, in ways that might affect migration decisions. The estimates are unchanged when including the origin by year fixed effects and 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. News about fracking in a particular state increases migration to that state but not other fracking states, suggesting the news conveys a location signal. Similarly, the effect of news exposure 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. “Positive” labor market news articles that discussing things like jobs, booms, or growth, have a larger positive effect on migration than “negative” environmental news articles discussing contamination, pollution or earthquakes. The effect of negative environmental news is still positive, suggesting even this news might provide information about where fracking is occurring. Positive and negative news affect cross-county commuting similarly, consistent with non-resident workers mostly experience the benefits of fracking while not incurring many of the costs (e.g., potential home water contamination). Using an analogous strategy that looks simultaneously at both newspaper exposure and TV news exposure, I find responses to both, suggesting other information sources matter as well. To understand how this information is used, I show that 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. There are similar spikes in fracking related tweets.

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 all else equal, 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. As this strategy has only focused on a few sources of information, the overall impact of information provision on migration behavior is potentially much larger.

II Information in Migration Decisions

A large literature explores the migration response to local labor market conditions and documents how this response varies by demographics, educational attainment, and geography (Bound & Holzer, 2000; Wozniak, 2010; Molloy, Smith, & Wozniak, 2011). Although heterogeneous preferences or costs might contribute to these differences (Notowidigdo, 2013;

Ganong & Shaog, 2017), there is credible evidence that liquidity constraints, credit constraints, and other market frictions impact the migration decision (Kling, et al., 2007; Bryan, et al., 2014). One potential friction is a lack of information.

The theoretical work has long recognized that information will affect migration decisions, but the empirical work has largely been limited to focusing on the role of networks or linguistic and cultural enclaves.[1] There are a few exceptions. The Moving to Opportunity (MTO) experiment and related work suggest that providing guidelines and information about local neighborhood poverty levels along with housing vouchers and assistance induced households to move to more affluent neighborhoods (Kling et al., 2007; Bergman et al., 2019). Although the MTO did not improve economic outcomes for treated adults, it did have positive long-run impacts on young children (Chetty, et al., 2015). Exploiting the Vietnam draft, Malamud and Wozniak (2012) show that college attendance increased migration rates, plausibly by increasing information about other areas through peer exposure.[2] McCauley (2019) shows that a publicized rating system of social service offices in the UK induces welfare migration. Kaplan and Schulhofer-Wohl (2017) propose a structural framework where information helps people learn about amenities in a different location. A similar information updating process can be applied to people's expectations about labor market opportunities to see how labor market information might impact migration decisions.[3]

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

[2]There is also work in the developing context suggesting that provision of labor market information in Bangladesh only impacts migration when combined with a conditional cash transfer (Bryan et al., 2014) and that access to more TV stations in Indonesia reduced the likelihood of moving (plausibly by correcting overly optimistic expectations about the returns to migration (Farre & Fasani, 2013). It is difficult to generalize these results to the United States. 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).

[3]An early related literature explores how things like unemployment risk (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, potential

In the canonical migration choice model (Sjaastad, 1962), an individual will move if the lifetime utility derived from moving to destination d minus the fixed costs of moving exceeds the utility of staying at the original location (o). However, individuals likely face incomplete information about the return to moving to destination d. This lack of information could impact the individual's willingness to move (see Appendix C for a complete conceptual model). Receiving positive information about d could lead an individual to update her beliefs about the potential return to moving there, increasing her willingness to move there.

For example, individuals exposed to numerous newspaper articles or TV news broadcasts touting the local economic benefits of fracking in Texas might adjust their beliefs about average wages or employment prospects in a Texas fracking county. Getting more information about a destination d could lead individuals to update their beliefs about the return to moving to d. Even news about negative aspects of fracking (e.g., water contamination risk) can provide information about where fracking is occurring and lead people to update their beliefs.[4] This news does not necessarily need to be correct, as long as the individual believes it contains truthful information. If information in the news does affect the migration decision, we would expect migration to increase when people's exposure to the news increases. It is also possible that each additional piece of information will have a smaller impact on migration as people become more confident in their prior.

In fact, there is evidence that exposure to news about fracking during this time increased the likelihood a person is aware of fracking and approves of it. The Pew Research Center's March 2012 Political Survey asked respondents if they have heard "a lot", "a little" or "nothing at all" about fracking. They then asked those who had heard anything about fracking if they approve of fracking. Using a measure of news exposure (described below) and respondents' state of residence, I find that state-level exposure to news about fracking

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

[4]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.

is associated with increases in the probability of hearing about fracking and increases in the probability of approving of fracking conditional on hearing about it (see Appendix Table A1). The impacts on awareness are observed across most demographic groups, while the impacts on approval are concentrated among men, younger individuals, and individuals with some college but no degree, groups that were likely to migrate to fracking (Wilson, 2020). This would suggest that news about fracking has the potential to shape people's beliefs, which could impact migration behavior. I will test these patterns empirically to determine how exposure to information in the news about fracking impacts migration to fracking areas.

III Setting and 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. An additional one million dollars of production value increases county-level total wages by approximately $80,000 and increase earnings in the commuting zone by approximately $114,000 (Feyrer et al., 2017). Fracking also created more jobs; an additional one million dollars of production value increased county-level employment rates by 0.85 percentage points. This is due to increases in mining (0.29), transportation (0.24), construction (0.12), and the government (0.10). Not only does fracking create jobs directly in oil and gas extraction, there are large, positive spillovers on employment in transportation and construction. There is also evidence of increased earnings in education, health, and other services. This is relevant as it is not necessary for potential migrants to work directly in oil and gas extraction. Kearney and Wilson (2018), Wilson (2020), and Cascio and Narayan (2019) show that the increases in earnings and employment are largest among men without a college degree (some college or less). Local fracking booms are credited with creating as many as 640,000 new jobs (Feyrer et al., 2017).

States in all 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 migration flows, news content, and news circulation. In this section I briefly describe each data source with a full description in the online data appendix (Appendix B).

First, I use well-level production data from DrillingInfo to identify fracking counties in the US. A fracking county is defined as any county with positive oil or gas extraction from a non-vertical well in a drilling formation that corresponds to a shale play. There are fracking counties 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 1. I examine migration pattern to fracking counties in these states.

Migration Data. Migration data come 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 one year and in a different county the next year. They then report the number of returns (households) and tax exemptions (individuals) that move for each county pair.[5] County pairs are censored for privacy purposes when there are fewer than 10 returns that move. In 2013, the censoring threshold increased from 10 to 20 returns, which would have suppressed 47.5 percent of the county-to-county

[5]In 2011, the IRS extending the data collect period from September to the end of the year. As such, households that file later (typically wealthier households) begin to be picked up in the data. The IRS have evaluated this change, and report that it had small impacts on state-level net migration rates (Pierce, 2015). If I exclude the states that experienced the largest change as reported by the IRS, the results are unchanged. See the data appendix for more detail.

flows between 2000 and 2012. For this reason I restrict my analysis to migration between 2000 and 2012.[6] I then aggregate up flows from each origin county to fracking counties in each fracking state. Because the IRS only provides a total count, I will use 2005 to 2012 American Community Survey microdata obtained from IPUMS to explore heterogeneous effects by demographic characteristics (Ruggles et al., 2015). This allows me to see if exposure to news about fracking affects migration differently by gender, race, education, and marital status.

Newspaper Circulation Data. Proprietary newspaper readership data is obtained from the Alliance for Audited Media (AAM). AAM conducts regular newspaper circulation audits of 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. I scraped historic circulation rates from 2005 through 2008 from pdfs.

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 collect all US based articles since 1999 that include the search terms "frack∼", "fracing", or "hydraulic fractur∼" anywhere in the text. I then parse each article to exclude spurious keyword references such as "frick and frack" or people's last names. Most of my analysis is restricted to three national news sources: *USA TODAY*, the *New York Times*, and the *Wall Street Journal*. In depth news coverage of fracking began in 2009 and dramatically increased 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, meaning

[6] To the extent possible I extend this through 2015. The impacts follow the same pattern and are equally significant, but are smaller, due to increased suppression. The patterns also hold if I only use data from 2009-2012 or 2011-2012 and omit the pre-treatment years (Appendix Table A8).

most of the variation in content comes from the last two years of the sample. Next, I parse the entire text of each of these articles to determine which of the 16 fracking states listed above each article discusses.[7] I report the number of articles by state in Appendix Table A2.

TV Viewership Data. TV viewership data is calculated from the 2008 Television and Cable Factbook using Nielsen viewership data. 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. This might have induced some viewers to substitute to other outlets (i.e., cable), meaning viewership rates in 2008 might be less correlated with viewership rates at the time of the broadcasts for markets that transitioned after 2008.[8] 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 is not program specific. The viewership rate is constructed by dividing total weekly viewership by the total number of households in the DMA.

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). VTNA only provides content for one hour of programming for the cable news outlets CNN and Fox News. Because cable news has limited content available and does not have reported viewership rates I restrict the sample to 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.[9] Between 1999 and 2012 there is far less coverage of fracking on the nightly

[7] 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.

[8] 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.

[9] These search terms differ from those used in the newspaper analysis. When I restrict my search to the term "fracking", only 12 news broadcasts during my sample window are found. Expanding to the other search terms used in the newspaper analysis ("frack∼", "fracing", and "hydraulic fractur∼") does not add additional broadcasts. However, given that only abbreviated transcripts

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. I also explore impacts on workers who live in one county but work in another using the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES). 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 provides the number of jobs for each home and work census block pair which I aggregate up to the county level to measure the number of workers who live in one county but work in another. This data is available beginning in 2002, and also provides statistics by broad age groups (under 30, 30-54, over 54), monthly earnings in the job you are commuting for (under \$1,250, \$1,250-3,333, over \$3,333), and industry of the job you are commuting for (goods, trade/transportation, other). This allows me to explore heterogeneous commute responses across different groups.

County Characteristics Data. County level economic measures 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. Other county level characteristics are obtained from the 2000 Census and ACS through the American Factfinder.

are provided, it is possible I am missing fracking related broadcasts. For this reason I expand my search terms to include "shale". This adds five more unique broadcasts. The term "shale" could reference other topics, but since there are so few broadcasts to check in the TV analysis I am able to verify that they are references to fracking in a way that is not feasible with the newspaper analysis where there are thousands of articles.

IV Empirical Strategy

Consider the hypothetical relationship between news exposure and Y_{oSt}, a measure of migration flows from origin county o to fracking counties in destination state S in year t:

$$Y_{oSt} = f(news\ exposure_{oSt}) + \phi_{oS} + \psi_{St} + \lambda_{ot} + \nu_{oSt}. \tag{1}$$

Migration flows are potentially impacted by exposure to the news, but also time invariant origin/destination specific characteristics, such as distance or industry ties; time-varying destination specific (ψ_{St}) and origin specific (λ_{ot}) characteristics, such as local labor market performance; and idiosyncratic origin-by-destination pair specific shocks (ν_{oSt}). This is potentially problematic for causal identification, as *news exposure*$_{oSt}$ might be correlated with circumstances at either the origin or destination.

These concerns are overcome in the following thought experiment: suppose we could randomly assign county-level exposure (i.e., circulation rates) to a newspaper that publishes news about the labor market impacts of fracking in Texas (or any other fracking state). By random assignment, news exposure will be uncorrelated with unobserved time-varying origin characteristics (λ_{ot}) and the origin/destination idiosyncratic term (ν_{oSt}). By comparing migration flows to Texas from counties with high and low circulation rates, everything about fracking counties in the destination state is held constant, allowing the effect of news exposure on migration to be isolated. If news about fracking in several states is being published, each of the state experiments could be stacked and estimated as follows

$$Y_{oSt} = f(news\ exposure_{oSt}) + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \tag{2}$$

Origin/destination pair fixed effects (ϕ_{oS}) control for time invariant characteristics of the pair that affect migration, like distance. Destination state-by-year fixed effect (ψ_{St}) 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 destination-level changes in fracking production, labor market characteristics, and amenities which might directly affect migration behavior and lead to higher news exposure.

In reality, exposure to the news is not randomly assigned. 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. While the content decisions of local newspapers have been shown to respond to local consumer preferences, this is less true for large national newspapers, such as *USA TODAY*, the *New York Times*, and the *Wall Street Journal*, which do not have well-defined geographic markets and operate differently (Gentzkow & Shapiro, 2010).[10] Counties across the country are exposed to the same national news, regardless of how their preferences deviate from the national trend.[11] Exposure to this news will vary based on local readership and circulation rates. Since fracking began quite suddenly, it is possible to isolate pre-existing variation in exposure that is not endogenous to preferences toward fracking. Oil and gas production from fracking only began in earnest in 2008, with little national news attention prior to 2009.

To isolate plausibly exogenous variation in news exposure, I will focus on exposure to news from national newspapers as follows

$$newspaper\ exposure_{oSt} = \sum_{n \in N} \Big(\text{total articles in } n \text{ about fracking in } S\Big)_t * Pre09\ circ.\ rate_{on}. \quad (3)$$

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

[11]Although national newspapers might report more about destinations that see large changes in labor markets or migration (nationwide trends), the destination by year fixed 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 A7 I show that the migration response is still significant if the New York City DMA is included.

Newspaper exposure$_{oSt}$ is the weighted sum of national news articles that mention fracking in destination state S in year t, where each newspaper is weighted by its fixed pre-2009 circulation rate (ranging from zero to one) in origin county o.[12] N is the set of national newspapers: *USA TODAY*, *New York Times*, and the *Wall Street Journal*. *Newspaper exposure*$_{oSt}$ is increasing in the number of articles about a particular destination, and increases by relatively more for counties that had high pre-fracking readership of the publishing newspaper.[13]

As in the thought experiment, the effect of national news exposure is identified by variation across origin counties in pre-fracking circulation rates. This 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 λ_{ot}. There are no strong geographic correlations in pre-2009 circulation of the *USA TODAY* and there is significant variation even among neighboring counties (Figure 2). Counties with low and high circulation of the *USA TODAY* also appear similar in 2000 on average (see Table 1), although low circulation counties had slightly lower employment, lower median income, higher poverty, and an older population. These level differences are not inherently problematic, as they will be controlled for by the origin/destination 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 1 suggest that migration trends in low and high circulation counties are parallel between 2000 and 2010. Pre-2009 circulation rates do not predict changes in migration between 2000 and 2010, but do have predictive power for some local demographic trends. However, these differences are quite small: the predicted differences associated with an increase in readership from the 25th to the 75th percentile of *USA TODAY* circulation are never

[12]This is average circulation between 2005 and 2008.

[13]This strategy is similar to previous work using variation in circulation exposure to explore the impact of media and news on other outcomes. 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 (2009) 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.

more than 0.51 percentage points. *New York Times* and *Wall Street Journal* pre-2009 circulation rates predict similarly small changes (Appendix Table A3).[14] Even though these differences are small and unlikely to matter, I can also include origin county by year fixed effects to control for any origin-level characteristics like these that might be changing over time.

My baseline estimation corresponds to the thought experiment as follows

$$Y_{oSt} = \beta_1 newspaper\ exposure_{oSt} + \beta_2 newspaper\ exposure^2_{oSt} + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \tag{4}$$

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 approximate the percent effect of news exposure. The state is used as the level of destination because few news articles reference specific counties by name, while state is frequently mentioned, meaning this is the level of variation. I only capture migration to fracking counties in the destination state, not the entire state, meaning I can examine flows from non-fracking counties to fracking counties in the same state. Origin counties with any fracking are excluded from the sample, as information in the news might affect the decisions of people originally living in fracking counties differently.[15]

Origin county by destination state fixed effects control for time-invariant pair specific characteristics, and destination state by year fixed effects control for changing characteristics of the fracking destinations. I include news exposure quadratically to capture decreasing marginal returns to information, although the relationship is robust to news exposure entered linearly as well as different functional forms (see Appendix Table A6). 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.[16]

[14] 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.

[15] Including these counties does not significantly impact the results (see Appendix Table A7).

[16] If I instead weight by the origin county population in 2000, the impact from equation (4) is about 1.5 times as large and significant, but less precisely estimated. The loss in precision appears

I begin with the specification in equation (4) 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 variation. First, I include time varying origin county labor market controls, including the employment to population ratio, unemployment rate, and average earnings (in 2010$) to capture observable changes in the origin labor market. Second, I include origin county by year fixed effects which account for both observed and unobserved components of λ_{ot}. This is possible because I observe migration flows to 16 different fracking states from each origin county/year pair. Origin by year fixed effects control for changing characteristics of the origin county that affect migration flows. For example, if counties with higher circulation rates, and thus higher newspaper exposure, are changing over time in ways that affect migration behavior (e.g., becoming younger, more educated, or more wealth), origin county by year fixed effects absorb these changes and exploit variation in news exposure across potential destinations from the same origin. This makes a within origin county comparison, to see if destination states that had more news exposure also experienced larger increases in migration flows. In this specification, any remaining confounding omitted variables must be origin/destination pair specific and vary over time (contained in ν_{oSt}). For this reason I next include a similarly constructed measure of local news exposure.[17] 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.[18] I also present event study evidence and use alternative strategies to verify the relationship is not driven by unobserved origin/destination specific changes over time.

to be driven by the very largest counties. If I weight by the natural log of the origin county population in 2000 or weight by the origin county population in 2000, but exclude the top ten percent of counties by population, the estimate are similar in magnitude and precisely estimated.

[17]This measure includes all domestic newspapers available through LexisNexis with available circulation data. 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).

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

As *newspaper exposure*$_{oSt}$ is a weighted sum, it is not immediate how to interpret the coefficients. If pre-2009 circulation rate in equation (3) is one (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. 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. This level of circulation is comparable to a county with high readership of *USA TODAY*.[19] 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 Results

V.A Event Study: Pre-trends and Treatment Effects by Circulation

First I present event study graphical evidence of the impact of national newspaper exposure on migration to verify that origin counties that will eventually be highly exposed to news do not have differential trends, relative to origins that are less exposed. Consistent with my main strategy, I focus on differences in initial circulation rates of national newspapers that will eventually report on fracking.[20] This tests to see if origin/destination specific news exposure is correlated with other unobserved characteristics that evolve over time and affect migration (ν_{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

[19] *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. Even though average circulation is low, there is substantial variation, which is exploited by this identification strategy. The coefficient of variation is 0.90 for *USA TODAY*, 1.07 for the *New York Times*, and 0.52 for the *Wall Street Journal*.

[20] 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.

this measure with year indicators between 2001 and 2012 (omitting 2000 as the reference year), and 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:

$$Y_{oSt} = \sum_{\tau=2001}^{2012} \theta_\tau Circulation_{oS} * 1\{t = \tau\} + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \tag{5}$$

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 3. 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.

Before 2008, only one 2006 *New York Times* article mentioned a specific destination state. Starting in 2008 there are small increases in the number of articles about fracking with a large jump in 2011 and 2012. Before 2010, migration fluctuates around zero, with only one statistically significant, negative estimate in 2003. Starting in 2006 there is a slight, 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.[21] 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.[22]

[21]This pattern could also arise if origin destination pairs with high circulation followed an upward trend that was suppressed during the Great Recession (2007-2010), only to rebound in 2011. The pattern is essentially unchanged if I control for this by including origin county unemployment rates, employment to population ratios, and average earnings.

[22]As seen in Appendix Figure A1, commuting responds similarly, although the increase is larger (8-12 percent) and begins earlier in 2009.

V.B Impact of Newspaper Exposure on Migration

Regression results from equation (4) are reported in Column (1) of Table 2. 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 $(0.025 * 1 - 0.001 * 1)$.[23] Average news exposure is also approximately one, suggesting the mean effect of news exposure on migration was 2.4 percent as well. Average news exposure was 1.8 in 2012, suggesting news about fracking increased migration flows to fracking counties by 4.2 percent on average in 2012.

I next adjust the baseline specification as outlined above to determine if changing characteristics of the origin bias the estimates. Controlling for annual origin county-level labor market measures in Column (2) does not change the coefficients. Including origin by year fixed effects in Column (3) absorbs the labor market measures included in Column (2) as well as any other unobserved characteristics of the origin that are changing over time and affect 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.[24] For the remainder of the paper, I 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.

Although these estimated impacts are small, they are both statistically and economically significant. They imply that news about fracking increased migration flows to fracking counties by 2.4 percent on average. From the levels specification, exposure to news about fracking in a particular

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

[24]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.

state led to 1.4-1.7 additional migrants from each origin on average. Another way to think about the effect size is to quantify how many people need to be exposed to the information in the news in order for one person to move. For reference the average county population in 2000 was 85,359 and the average household size was 2.59 people. If 5 percent of households receive the newspaper this would suggest that when 11,054 ($0.05 * 2.59 * 85,359$) people saw the article about fracking an additional 1.44-1.67 people would move, or only 0.015 percent. In other words, it takes about 10,000 interactions with the news about fracking for one additional person to move. In many cases, the news exposure measure is capturing more articles distributed to fewer readers. But even at a one percent circulation rate, we would still expect 2,211 ($0.01 * 2.59 * 85,359$) people to be exposed to multiple articles with 1.44-1.67 people moving, or only 0.076 percent. These ratios ignore any spillover effects of news passing through social networks, which would lead to even more people being exposed to the news. For a given origin this effect is small, but when aggregated up for a given destination the effect is large. This would suggest that providing information about potentially lucrative labor market opportunities elsewhere can increase migration to those destinations.

V.C Heterogeneous Responses Across Demographic Groups

The IRS data only provides the number of migrants and does not provide demographic characteristics. Exploring heterogeneous effects could shed light on who responds to the news and provide further evidence of the information mechanism. The annual American Community Survey (ACS) asks residents where they lived in the previous year, allowing me to construct origin/destination migration rates for demographic subgroups from the microdata. Unfortunately, the geographic data is only available starting in 2005 and the smallest geographic unit is the Public Use Micro Area (PUMA). PUMAs are geographic areas defined by population that are large enough to preserve privacy. Furthermore, migration geographic data is only available at the Migration PUMA (MIGPUMA) level, which are often even larger and contain one or more counties. There are several aspects of the data that are likely to make it more difficult to detect an effect. First, the ACS is a one percent sample of households so there is likely to be measurement error in the constructed migration rates which will reduce precision. Second, there is less geographic variation (and less variation in circulation rates) than is available at the county level leading to less precision.

Estimating an equation analogous to equation (4), I explore effect heterogeneity by gender, race, education, and marital status and plot the total effect of one unit of newspaper exposure in Figure 4. For the full population I estimate a significant 1.2 percent increase in migration to the state being mentioned in the news.[25] This effect is only half as large as the estimate using IRS data and less precise, as we would expect given the concerns about power discussed above. The impact for the full population is mostly driven by men, where I estimate a 0.95 percent increase in migration. Other groups that see significant increases in migration are 18-34 year olds, some college (at the ten percent level) and the unmarried. These are the same groups that largely drove the total migration response to fracking (Wilson, 2020). The impacts for high school graduates and dropouts (which also were drawn to fracking areas) are insignificant, but these groups are also less likely to be exposed to these news sources, which might explain the insignificant response.

Subgroups we would expect to be highly responsive, such as men with some college, unmarried men, young men (18-34), unmarried men with some college, young men with some college, and young unmarried men see significant effects ranging between 0.74 and 1.1 percent. This is highly consistent with both the data and anecdotal evidence about who moved to fracking booms. In other words, we see precisely the people we would expect to respond moving after the news.

V.D Impact of Newspaper Exposure on Cross-County Commuting

Individuals can commute to avoid the monetary, psychic, and amenity costs associated with moving. Many people took advantage of the earnings gains associated with fracking by commuting rather than migrating (Wilson, 2020). Exposure to information might also affect commuting. In Table 3 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. 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.

When looking across the three pre-defined age groups, the response to one additional news

[25] Estimates are similar if I include the local labor market controls.

article for 30 to 54 year olds is 5.2 percent and statistically larger than the response of younger workers (3.1 percent) and older workers (3.6 percent).[26] This pattern is consistent with age-specific patterns in newspaper readership and commuting which offset each other. Newspaper readership (i.e., exposure) increases with age (Pew Research, 2013b), while geographic mobility falls with age (Molloy et al., 2011).[27] In Appendix Table A4, I also report differences by the earnings and broad industry of the job they are commuting to. Consistent with people commuting to high paying fracking jobs, workers commuting to jobs that pay over $3,333 a month ($40,000 a year) are the most responsive. Consistent with previous work finding employment impacts across industries, commuting responds across industries, suggesting that news exposure not only induced people to commute for oil and gas extraction, but for other jobs affected by the labor market shock.

V.E Robustness

Estimates are robust to functional form and specification decisions. In Appendix Table A5, I re-estimate a variant of equation (7) including and excluding various fixed effects. The coefficients for migration are insensitive to the inclusion of destination by year and origin by year fixed effects, suggesting unobserved trends at the origin or destination do not play an important role. The coefficients for migration are essentially unchanged if I also add origin state by destination state by year fixed effects, to compare migration flows from origin counties in the same state to a given fracking destination and account for origin/destination pair specific trends. The coefficients for commuting are more sensitive to the inclusion of various fixed effects, but are still significant even under the most conservative specifications. The effects of newspaper exposure on both migration and commuting are robust to including newspaper exposure linearly, quadratically, as a cubic, or as the inverse hyperbolic sine (see Appendix Table A6); sample and year restrictions (Appendix

[26]The percentage effect is larger for all workers than for any of the subgroups in part because the pooled specification constrains the controls and fixed effects to be the same for each group. When run in levels, the effect for all workers is the sum of effects for each subgroup, as expected.

[27]In Figure 4 35-44 year olds exhibit a decline in migration. This is plausible as migration and commuting are different decisions and potentially even substitutes.

Tables A7 and A8);[28] or accounting for censoring in the IRS migration data (Appendix Table A9).[29] I also explore the role of distance, by estimating equation (4) for origin county by destination state pairs in one hundred mile bins and plot the total marginal effect of news exposure for each distance in Appendix Figures A2 and A3. The effect climbs to about 6 percent between 400 and 1,000 miles and then gradually falls, consistent with information provision having no effect on migration to nearby opportunities that people might be aware of, but a large impact on migration to distant potential opportunities.

The inverse hyperbolic sine specification is appropriate if information in the news has relative effects on the number of migrants. Relative effects would be present if treatment effects are heterogeneous, if information interacts with other individual or local characteristics, or if information had the largest impact on populations at the margin of moving, with higher baseline migration rates.[30] Information could instead have constant absolute effects on migration rates. In Appendix Table A11, I show that news exposure significantly increases the number of migrants per person as well. One additional article in a newspaper with a 5 percent circulation rate leads to approximately 1 more migrant for each 200,000 people at the origin. These estimates imply similar exposure to migration ratios, suggesting that 10,000 interactions with the news induced 0.8-1.2 additional people to move. I also report estimates using the inverse hyperbolic sine of the migration rate. These estimates are more precise, suggesting the data is better fit by relative effects specifications. These estimates suggest that newspaper exposure increased migration rates by 1.9 percent at the mean and provide further evidence that information in the news increased migration.[31]

[28]Estimates are insensitive to including fracking origins, excluding non-fracking origins in fracking states, or limiting the sample to counties with positive newspaper circulation to focus on intensive margin variation in exposure.

[29]Effects are similar if zero-inflated Poisson or Tobit models are used to explicitly model the selection into non-zero migration due to censoring (Appendix Table A10).

[30]As Bellemare & Wichman (2020) suggest, the inverse hyperbolic sine transformation will identify the same semi-elasticity as a log-linear specification if the mean of the outcome (un-transformed) is large. They suggest a value above 10, for a rule of thumb. The mean number of migrants is slightly below this, 7.6. However, based on their algebraic bias expressions, the semi-elasticity would be 2.37 percent rather than 2.4 (see Appendix C for a complete discussion).

[31]Heterogeneous effects by demographics are similar when using group-specific migration rates. Men, young adults, individuals with some college, the college educated, and pairwise combinations

One concern is that 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, 2020). However, readers of the *USA TODAY* look similar to typical fracking migrants. In 2007, 68 percent of *USA TODAY* readership was male, 66 percent were 25-54, 57 percent had no college degree, and 31 percent had household incomes below $50,000 (see Appendix Table A12). In Appendix Table A13, I estimate the impact of exposure to news from each of these three newspapers separately. 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*. The same patterns hold by newspaper when looking at the ACS microdata (Appendix Figure A4). This lends further support to the information mechanism as the largest responses are driven by the news source that is more accessible to the typical migrant to fracking.[32]

VI Alternative Strategy: Newspaper Market Border Comparison

Although the estimates in Table 2 are not driven by origin level characteristics, news exposure could be correlated with other, unobserved origin destination pair characteristics captured by ν_{oSt}. Counties with higher circulation of national news, might for unobserved reasons be more likely to move 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 test this, I employ an alternative strategy that exploits variation in news exposure among neighboring counties. Using all domestic newspapers that had circulation data and at least one article about fracking between 1999 and 2012, I construct geographic markets for each newspaper that capture the set of counties in the newspaper's distribution network. Distribution costs inhibit broad distribution of local newspapers and these markets are small groups of adjacent counties around a central hub.[33] I identify counties on the border of the

of these groups report significantly higher migration rates in response to news exposure.

[32]One could scale up the migration effects by the group specific probability of reading the *USA TODAY*, like a first stage, to see which groups are more responsive. Given potential information spillovers across groups it is not clear that this exclusion restriction would hold.

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

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 either side of the border. This is done in a stacked regression as follows

$$
\begin{aligned}
Y_{oSt} = \gamma_1 Articles_{nSt} * InMarket_{on} + \gamma_2 Articles^2_{nSt} * InMarket_{on} \\
+\gamma_3 InMarket_{on} + X'_{ot}\Gamma + \phi_{oS} + \psi_{nSt} + \varepsilon_{oSt}.
\end{aligned} \tag{6}
$$

The outcome is the same as above. *Articles* is the number of articles in newspaper n in year t about fracking in state S (in units of ten). *InMarket* is an indicator that equals one if the origin county o is in the market of newspaper n. To be specific n uniquely identifies each newspaper and the corresponding market border. Counties that do not receive newspaper n but are on the other side of the market border will also be assigned to border n as control counties. Time-varying origin controls and origin-destination pair fixed effects are included. Newspaper-by-destination state-by-year fixed effects are also included, making this a comparison of flows to the same destination among counties along the same market border. Because I am making a pairwise cross-border comparison, I include every county pair along a newspaper market border, meaning a county-year observation may appear multiple times if it borders several counties across the 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 compared to other local counties, similar preferences and propensities among these neighboring counties captured in ν_{oSt} will be differenced out.[34] This will identify the causal effect of news coverage as long as propensities to migrate to fracking during booms are similar within the county pair.

Relative to no articles, ten news articles significantly increased migration by 5.6 percent in counties that received the newspaper, relative to their neighbors and commuting increased by 4.9

[34]As seen in Appendix Table A14, counties on either side of these market boundaries appear quite similar in 2000 and experienced similar trends on average between 2000 and 2010. Between 2000 and 2010, counties just outside the market see a slightly larger drops in employment, smaller increases in poverty, less growth in the minority population, and more population aging when controlling for newspaper border fixed effects. However, these differences are small, for example, the difference in the employment decline is only 0.2 percent of the mean.

percent (see Table 5). In this sample, average circulation among in-market border counties was 5.5 percent, slightly higher than the benchmark 5 percent, making it easy to compare the magnitude of this effect to the previous estimates. 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.[35]

VII Additional Explorations

VII.A News about Fracking in Another State

News about fracking could provide general information about the labor market impacts of fracking or specific information about which labor markets are impacted. To evaluate the relative importance of these channels, I estimate how migration flows to fracking counties in a particular destination state respond to news about fracking in a different state. For example, observing that migration to North Dakota is less responsive to news about fracking in Pennsylvania than to news about fracking in North Dakota would suggest the location signal is important.

In practice, I randomly assign all observations indexed by S the fracking news exposure of one of the other 15 fracking states S'. 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 equation (4), but replace $\textit{News Exposure}_{oSt}$ with the randomly assigned $\textit{News Exposure}_{oS't}$, and calculate the marginal impact of a one unit increase in $\textit{News Exposure}_{oS't}$ (i.e., one news article in a county with a five percent circulation rate). I repeat this 200 times and plot the histogram of potential impacts in Panel A of Figure 5, with the estimated effect using actual news exposure from Table 2 indicated. This is not a placebo test, as general information can plausibly affect the outcome.

The effect from Table 2 is larger than all but 5 of the repetitions (2.5 percent), suggesting the

[35] The point estimates are similar if I exclude national newspapers or restrict the sample to only include one newspaper market border per county to ensure that counties only appear once. 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.

location signal is significant. 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. 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.

If I repeat the exercise but include origin by year fixed effects, the effects are centered around zero and all smaller than the effects from Table 2 (see Panel B of Figure 5.) This specification tests 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. These results are consistent with the news conveying a location signal rather than general information; however, I cannot rule out that there are simply differences across destinations in the content of the news. Because there is no average effect, it is unlikely the previous results are driven by things correlated with news exposure and pre-fracking circulation rates in general.[36]

VII.B Migration Dynamics: Chain and Short-term Migration

Understanding the dynamic nature of the migration response and the incidence of chain migration or short-term migration can shed light on the size and welfare implications of these effects (Dustmann & Gorlach, 2016a; Dustmann & Gorlach, 2016b). Chain migration is prevalent in the international migration literature and news exposure could induce this behavior if migrants inform acquaintances

[36] This is related to the advertising literature (Garthwaite, 2014), suggesting information could either lead to more migration overall (expansion) or shift people from other destinations (share stealing). Including total exposure to news about any of the 16 destination states has no effect on migration, but a small effect on commuting (see Appendix Table A15). Including the news exposure for the state that received the highest exposure within an origin and year has a positive effect, but does not reduce flows to other fracking destinations, suggesting news exposure led to expansion, rather than shifting. Information in the news also does not simply induce people to move to the closest fracking area and does not appear to shift migrants from non-fracking to fracking areas.

at their origin about the situation. This might be particularly relevant if there are cross-person spillovers as people convey news to acquaintances. To test for chain migration, I regress migration flows on news exposure as well as lags of news exposure to see if news last year or two years ago impacts migration this year (see Appendix Table A16). An impact of lagged news exposure would capture both delayed moves and people who move after a friend or acquaintance is induced to move by the news (chain migration). When I regress migration on this year's news, last year's news, and news from two years ago, the impact of current news is not significantly different from the estimate in Table 2. The coefficient on last year's news is much smaller (0.7 percent) and marginally significant, with no significant impact of news from two years ago. This would suggest that current news has an impact on migration and there is some margin for a small, delayed moving response or chain migration. If this lagged response is due to chain migration, we would expect the effect to become insignificant if we control for last year's migration. However, when I control for last year's migration, the effects of news exposure are unchanged, suggesting the effect of lagged news exposure is likely coming from delayed migration rather than chain migration.[37]

People that respond to the news might also move only temporarily before moving back. Because the IRS migration data only captures flows and does not follow individuals, I cannot directly examine the duration of migrants' moves. However, by looking at reverse migration dynamically we can understand if short-term stints and return migration are common. From the IRS data I construct the reverse migration rates from fracking counties in state S to county o. I then estimate how news exposure in origin county o about fracking state S affect migration from S to o in the future (see Appendix Table A17). Exposure to news about fracking does not lead to increased reverse migration in that same year. However, news exposure in the previous one or two years is associated with a significant increase in reverse migration. The coefficients are smaller than the direct effect on migration, but they are similar in magnitude suggesting many of the people who were induced to move to fracking areas by the news eventually move back. This is consistent with

[37]It should be noted that this lagged dependent model with origin/destination pair fixed effects is only consistent under strict assumptions. Using the ACS microdata I also estimate the effect of news exposure on the probability of joining an existing household of relatives or non-relatives to capture chain migration. There is no significant effect on joining an existing household and the estimates are close to zero (Appendix Figure A5).

the prior work documenting both elevated inflows and outflows in fracking areas (Wilson, 2020).

VII.C Positive Labor Market versus Negative Environmental News

While some news references positive characteristics of fracking such as jobs, booms, or growth, other news discusses negative aspects such as pollution, health, dangers, and earthquakes. Individuals might respond differently to positive labor market news and negative environmental news. Given the importance of the location signal, even negative environmental news could impact people with pre-conceived beliefs about fracking who did not know where it was occurring. I parse each article for keywords such as "growth", "boom", "contaminat∼", and "earthquak∼" to determine the positive and negative content of each article. These statistics are reported in Appendix Table A2. I then classify an article as positive if it has at least two positive mentions and has more positive mentions than negative. Negative articles are similarly defined. I then estimate the separate effect of positive and negative newspaper exposure on migration and commuting in Table 5.[38] Relative to no newspaper exposure in a county with a five percent circulation rate, one positive article significantly increased migration by 4.0 percent. The effect of negative news is half as large, but still positive and significant, suggesting even negative news provides information that induces migration.

In contrast to migration, positive labor market and negative environmental newspaper exposure affect commuting similarly. Unlike migrants, long distance commuters do not bear some of the costs associated with fracking such as potential home water contamination or noise on residential streets. For commuters, the location signal from negative environmental news might be just as effective as a location signal from positive labor market news.

VII.D The Role of Origin County Labor Market Conditions

Recently, there has been 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 et al., 2016).[39] I next test if the impact of news exposure varies by labor market

[38]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.

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

strength at the origin to understand if providing information is particularly impactful in weak economic areas. To do this, I modify my main specification as follows

$$
\begin{aligned}
Y_{oSt} = {} & \beta_1 newspaper\ exposure_{oSt} + \beta_2 newspaper\ exposure^2_{oSt} \\
& + \beta_3 newspaper\ exposure_{oSt} * emp/pop_{ot-1} + \beta_4 newspaper\ exposure_{oSt} * emp/pop^2_{ot-1} \\
& + \beta_5 newspaper\ exposure^2_{oSt} * emp/pop_{ot-1} + \beta_6 newspaper\ exposure^2_{oSt} * emp/pop^2_{ot-1} \\
& + \beta_7 emp/pop_{ot-1} + \beta_8 emp/pop^2_{ot-1} + X'_{ot}\Gamma + \phi_{oS} + \psi_{St} + \varepsilon_{oSt}. \qquad (7)
\end{aligned}
$$

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 A18). The effects on both migration and commuting are plotted in percentage points in Figure 6 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.[40]

The information is most impactful for people in counties with weak labor markets, where the expected gains to moving are largest. Actual exposure levels are similar for weak and strong labor markets, suggesting only a small part of the heterogeneous impacts can be explained by differential exposure. This has several potential policy implications. 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 information about

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

potential labor market opportunities in other parts of the country to all counties could 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.

VII.E Impact of TV News Exposure

Over time, newspapers have become a less important source of news, while television remains an important source for 69-74 percent of adults, and the internet has become increasingly important (Pew Research, 2013b). 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 Vanderbilt Television News Archive (VTNA) for the three major TV news networks (ABC, CBS, and NBC) I construct a measure of TV news exposure similar to the measure of national newspaper news

$$TV\ news\ exposure_{oSt} = \sum_{c \in C} \left(\text{broadcasts on } c \text{ about fracking in } S\right)_t * pre09\ view\ rate_{oc}. \qquad (8)$$

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. Between 2007 and 2009, TV stations transitioned from analog to digitally transmitted broadcasts on a market-by-market basis, requiring viewers to obtain digital reception equipment. This might have induced some viewers to substitute to other outlets (i.e., cable), making 2008 viewership rates less correlated with viewership rates at the time of the broadcasts and introducing measurement error. I also run specifications using ratings from the latest 2016 Factbook, after all markets were updated.[41] Nielsen ratings are only available at the

[41] Using 2016 ratings potentially introduces endogeneity if viewership is responding to migration and commute behavior. Circulation rates are highly persistent, suggesting this bias is likely small.

DMA-level, which is a mutually exclusive set of contiguous counties that represent a media market. As such, I aggregate up migration flows, newspaper exposure, and labor market measures from the county to the DMA-level.[42] 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.[43]

DMA-level estimates are presented in Table 6. 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 is twice as large as the county-level estimate, but not statistically different. Using 2008 viewership rates, the impact of TV news exposure on migration is insignificant. The coefficients are similar when I include both newspaper and TV news exposure in Column (3).[44] In Columns (4) and (5) I repeat the analysis, but use 2016 viewership rates to construct TV news exposure to reduce measurement error. When only TV news exposure is considered one additional broadcast in a market with 50 percent viewership leads to a 8.7 percent (0.111-0.024) increase in migration. When both newspaper and TV news are included, the effect of TV news is a similar magnitude, but no longer precise (the p-value on the first order effect is 0.11), suggesting TV news exposure might affect migration. Both newspaper and TV news have separate, significant effects on commuting.[45]

There are fewer TV news broadcasts about fracking (only 17 total) distributed to a larger group of people than newspaper news. If the accumulation of information signals is a driving mechanism (as the conceptual framework would suggest) then a small group of people getting multiple pieces of information might have a larger effect than many people getting fewer pieces of information. This is consistent with a significant impact of newspaper exposure but a much weaker impact

[42]There are only 203 DMA, which reduces the variation in the explanatory variable because circulation rates are now calculated over larger areas.

[43]This measure does not capture 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.

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

[45]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 estimates on commuting.

of TV news exposure. If the decision to commute is less costly than the decision to move, we might expect limited information (such as that provided in TV broadcasts) to be more likely to affect commuting, which is once again consistent with what we observe. The nature of information signals are different across newspapers and TV news. The TV news broadcasts are all short, only 1-5 minutes and provide less information than a newspaper article. If the quality of the information signal is lower, we might also be less likely to observe a response. This analysis would suggest that both the intensity and penetration of content influence the magnitude of the effect.

VIII Online Activity, a Potential Mechanism

Exposure to news about fracking in a particular state increases migration and commuting to fracking counties in that state. This is potentially driven by information provision, but cannot be directly tested in the data. However, using Google Trends data I can quantify how interest in fracking and the states mentioned changes after TV news broadcasts.[46] For a specified search term (e.g., "fracking"), Google Trends provides 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. As such, I can use this data to examine changes in search behavior in a given area 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.[47] I look at search intensity for the term "fracking" and for the name of any states that are mentioned in the broadcast. Only 14 broadcasts mention a state. I estimate the following event study regression

[46]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.

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

$$search\ index_{opt} = \sum_{\tau=-14}^{14} \delta_\tau * \mathbf{1}\{t \text{ is } \tau \text{ days from broadcast}\}_{op} + X_t'\Gamma + \phi_{op} + DOW_t + \varepsilon_{opt} \quad (9)$$

where $search\ index_{opt}$ is the search intensity on date t in DMA o relative to the search period p. The search period is the 30 days surrounding each broadcast and op uniquely identifies each DMA/period pair over which the relative search index is measured. The δ_τ coefficients trace out daily search intensity relative to the omitted day ($\tau = -15$). I include DMA by search period fixed effects 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 throughout the week.[48]

These effects are plotted in Figure 7 for "fracking" and the state names. Search intensity for "fracking" spikes on the day of the broadcast and remains elevated for the next two days before falling back to the previous levels. Search index is a relative measure, and cannot be used to back out how many additional searches were made. If I combine days into 3 day bins, for statistical power, I estimate a similar spike, followed by a smaller statistically significant persistent, increase, suggesting search interest remained elevated for some time (see Appendix Figure A7).

Search interest in the names of states mentioned in the broadcast also spikes one day after the broadcast and remains elevated for the next five days. Search intensity for the same set of fracking states that are not mentioned in the news does not change after the broadcast. Although not direct evidence that news coverage motivates people to move, the Google Trends analysis suggests news coverage induces people to seek more information about the potential fracking destination.

Twitter content shows similar impacts. There is a spike in the average number of tweets, retweets, and likes of tweets that include "fracking" in the days following TV broadcasts (see Figure 8).[49] These tweet often include links or mention specific Twitter users' handles. There are also spikes in positive terms like "jobs", "money", and "opportunity" (see Appendix Figures A8-A10).

[48]Several broadcasts were in close proximity to high publicity events connected to either fracking or the states mentioned in the reports. When looking at searches for indicator to control for the time around these high publicity events, listed in Appendix Table A19. The pattern is similar if I do not control for these events (Appendix Figure A6).

[49]Because I am using historic data I cannot use the Twitter API to access user location. This analysis only includes one series for each broadcast event and examines the interrupted time-series.

IX Conclusion

Migration is a way for individuals to improve labor market opportunities, but people tend not to move away from poor labor markets. In this paper I evaluate the role of information in the decision to move to labor market opportunities. The current literature speaks very little to the effect of labor market information on migration behavior. I exploit information about local fracking booms disseminated through the national news to estimate the effect of news on migration.

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. Cross-county commute flows also increase by 6.6 percent. Fracking news about a specific destination increases migration to that state, but not other fracking states, suggesting the location signal is important. There is evidence that news exposure induces short-term migration, with many people returning after several years. Migration flows are more responsive to exposure to positive labor market news than negative environmental 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). Other sources of news, such as TV news, also affects commuting and potentially migration. As further evidence that news coverage increases interest in these fracking destinations, Google search interest in both the term "fracking" and the names of states mentioned significantly increases after news broadcasts.

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 when 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. All else equal, 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.

Bellemare, Marc and Casey Wichman. 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation," *Oxford Bulletin of Economics and Statistics*, 82(1):50-60.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer. 2019. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice," NBER Working Paper No. 26164, National Bureau of Economic Research, Cambridge, MA.

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).

Cascio, Elizabeth and Ayushi Narayan. 2020. "Who Needs a Fracking Education? The Educational Response to Low-Skill Biased Technological Change," *ILR Review* accepted.

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.

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).

Dustmann, Christian and Joseph-Simon Gorlach. 2016. "The Economics of Temporary Migrations," *Journal of Economic Literature*, 54(1): 98-136.

Dustmann, Christian and Joseph-Simon Gorlach. 2016. "Estimating Immigrant Earnings Profiles when Migrations are Temporary," *Labour Economics*, 41: 1-8.

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.

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.

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

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.

McCauley, Jeremy. 2019. "The Role of Information in Explaining the Lack of Welfare-Induced Migration," *Working Paper*, University of Bristol.

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/.

Pierce Kevin. 2015. "SOI Migration Data: A New Approach, Methodological Improvements for SOI's United States Population Migration Data, Calendar Years 2011-2012," The Internal Revenue Service.

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. 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. 2020. "Moving to Economic Opportunity: the Migration Response to the Fracking Boom," *the Journal of Human Resources*, accepted.

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 1

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. Standard errors are corrected for clustering at the origin DMA level. p<0.01 ***, p<0.05 **, p<0.1 *.

Table 2

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 3

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. Observations are not limited to origin counties in the same local commuting zone and includes long distance commuters. 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 4

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. I include every county pair along a newspaper market border, meaning a county-year observation may appear multiple times if it borders several counties across the market border. 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 5

Positive Labor Market vs. Negative Environmental News: Impact of Newspaper Exposure on Migration and Commuting

	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 6

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)	(9)	(10)
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	0.111*	0.100		0.125**	0.115*	0.144**	0.134*
		(0.072)	(0.072)	(0.065)	(0.064)		(0.062)	(0.061)	(0.069)	(0.069)
TV News Exposure$^2_{oSt}$		-0.010	-0.011	-0.024*	-0.021		-0.019	-0.019	-0.021	-0.018
		(0.017)	(0.017)	(0.014)	(0.013)		(0.014)	(0.014)	(0.014)	(0.014)
2008 TV Viewership Rates		X	X				X	X		
2016 TV Viewership Rates				X	X				X	X
Dependent Mean (in Levels)	60.2	60.2	60.2	60.2	60.2	152.6	152.6	152.6	152.6	152.6
Observations	32,864	32,864	32,864	32,864	32,864	27,808	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),(5), (9), and (10) 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. $p<0.01$ ***, $p<0.05$ **, $p<0.1$ *.

Figures

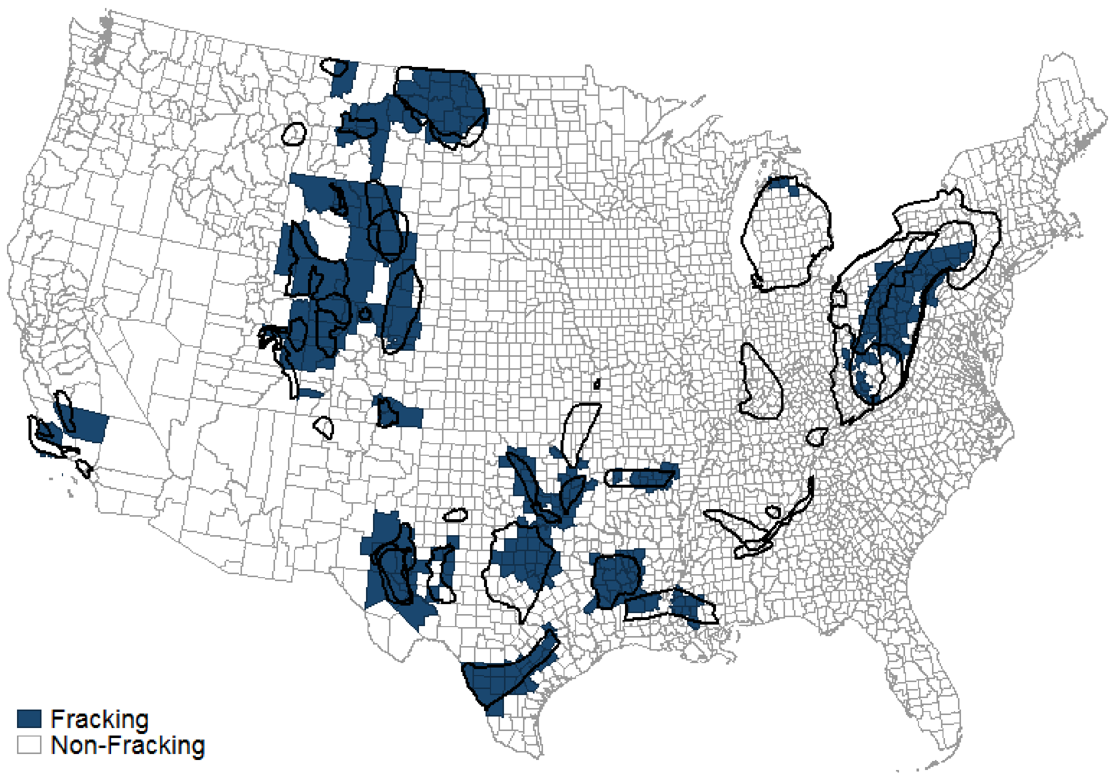

Figure 1

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.

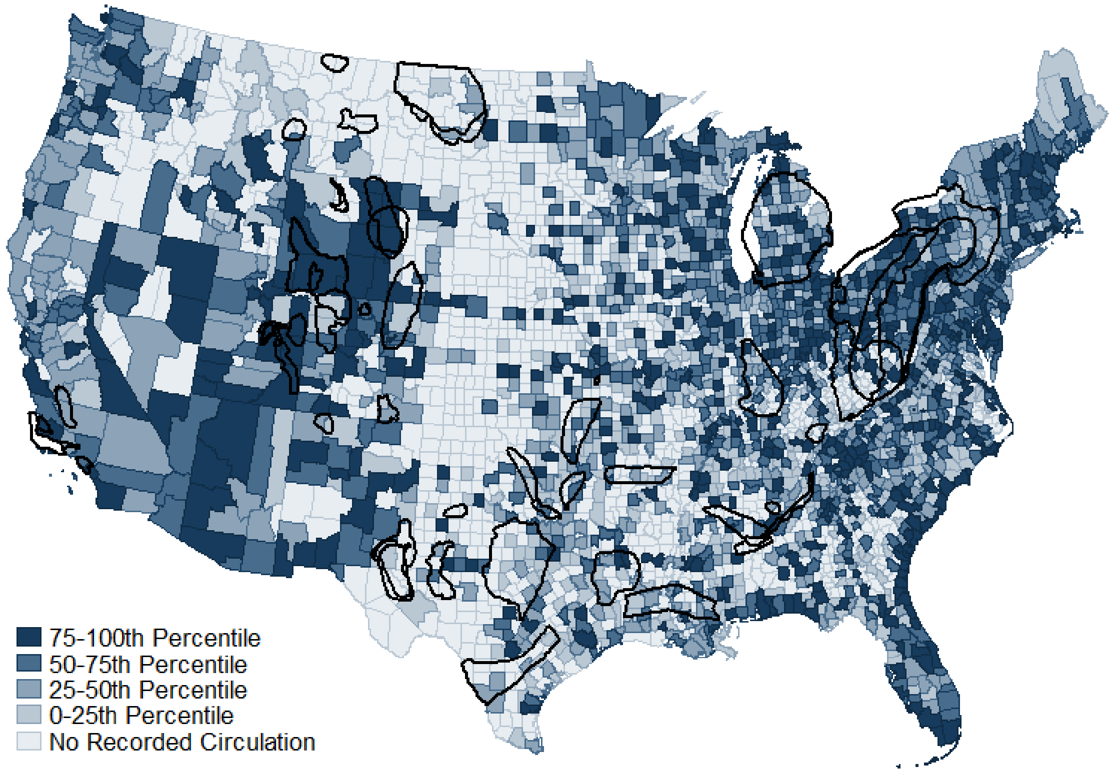

Figure 2

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.

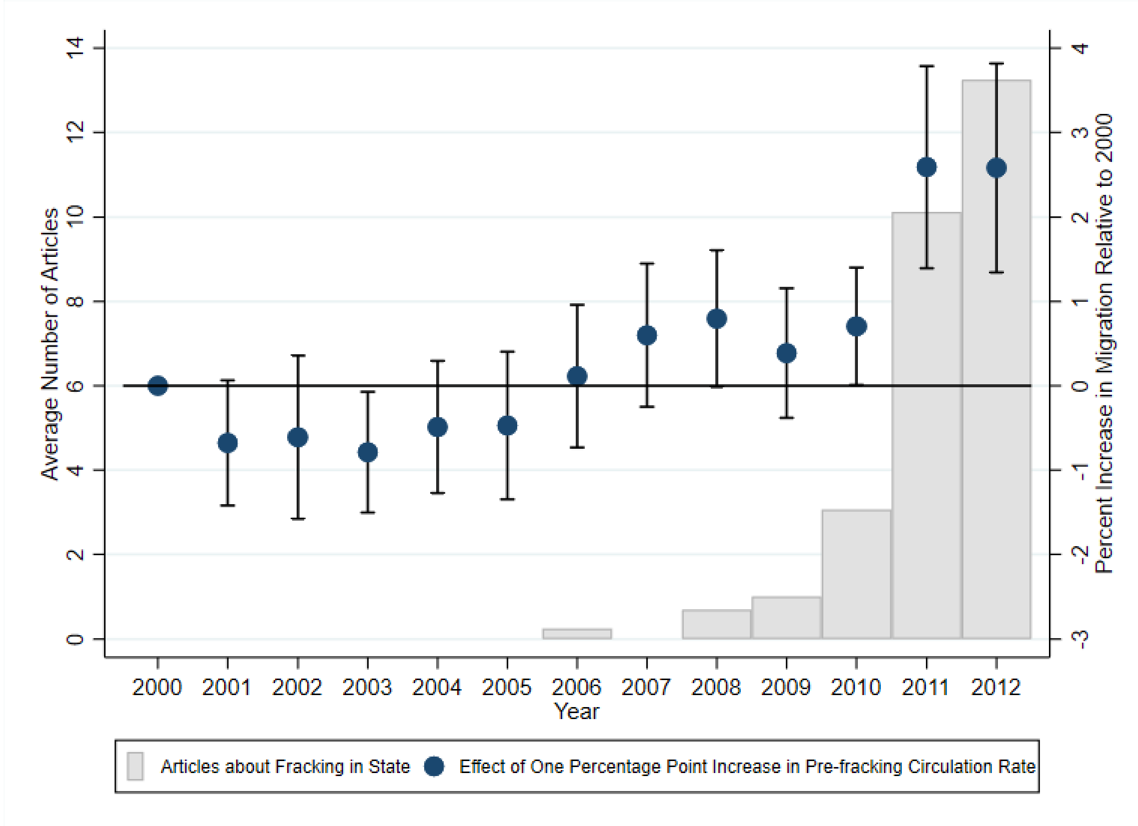

Figure 3

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 are corrected for clustering at the origin DMA level, with 95 percent confidence intervals provided. 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.

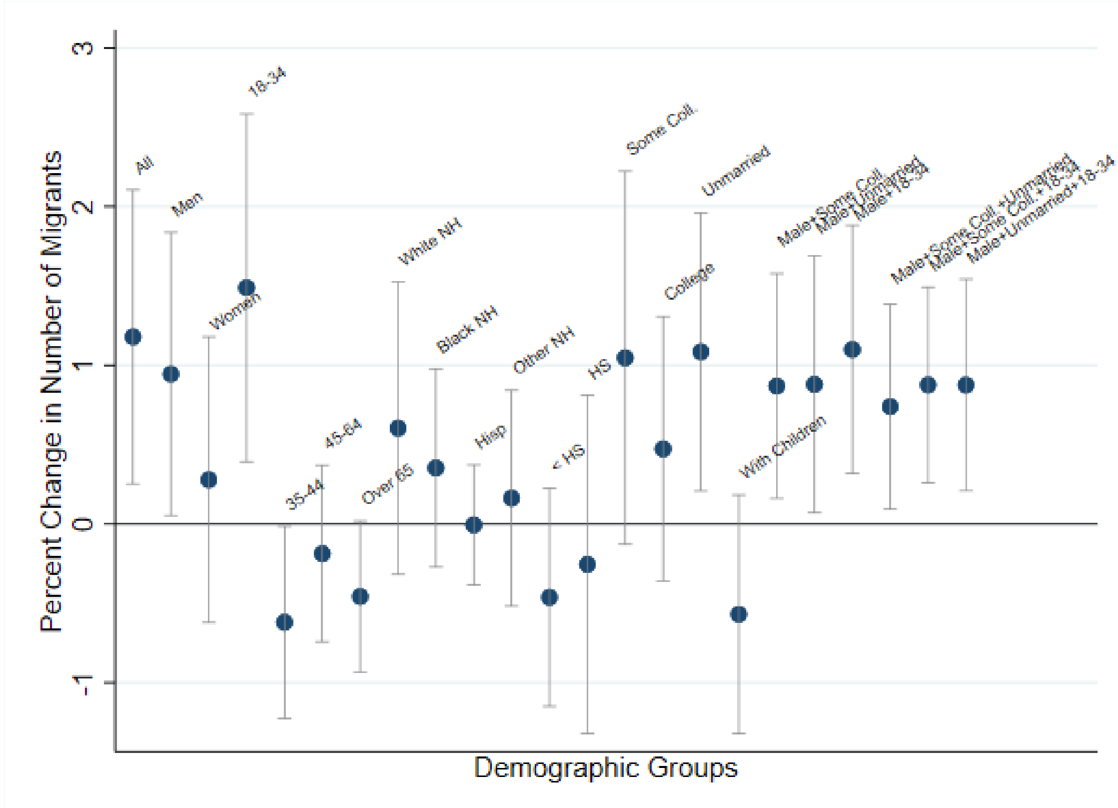

Figure 4

Migration Impact of Newspaper Exposure by Demographic Group

Notes: The total impact of a one unit increase in Newspaper Exposure is plotted with 95 percent confidence intervals. The level of observation is the origin MIGPUMA by destination state by year from 2005 to 2012 for the given group. Origin MIGPUMAs with any fracking production 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. 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 origin state.

Source: Author's calculation using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and microdata from the 2005-2012 American Community Survey.

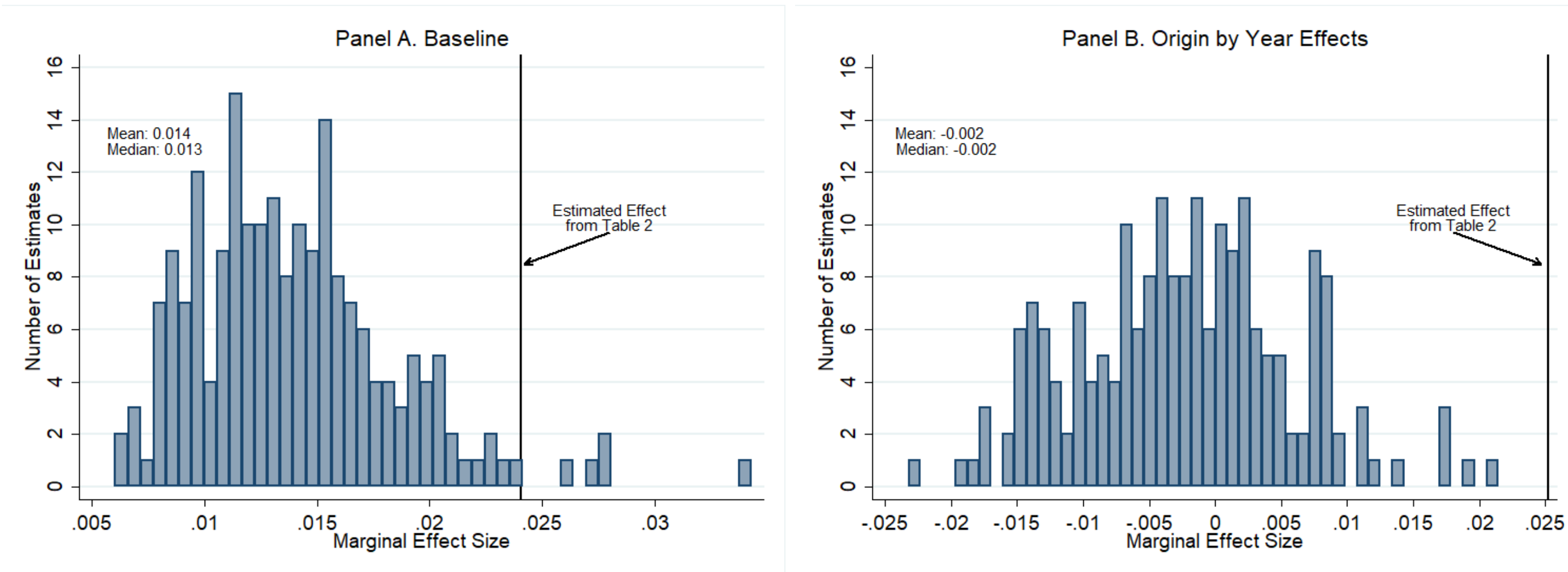

Figure 5

Location Signals: 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 (4). 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.

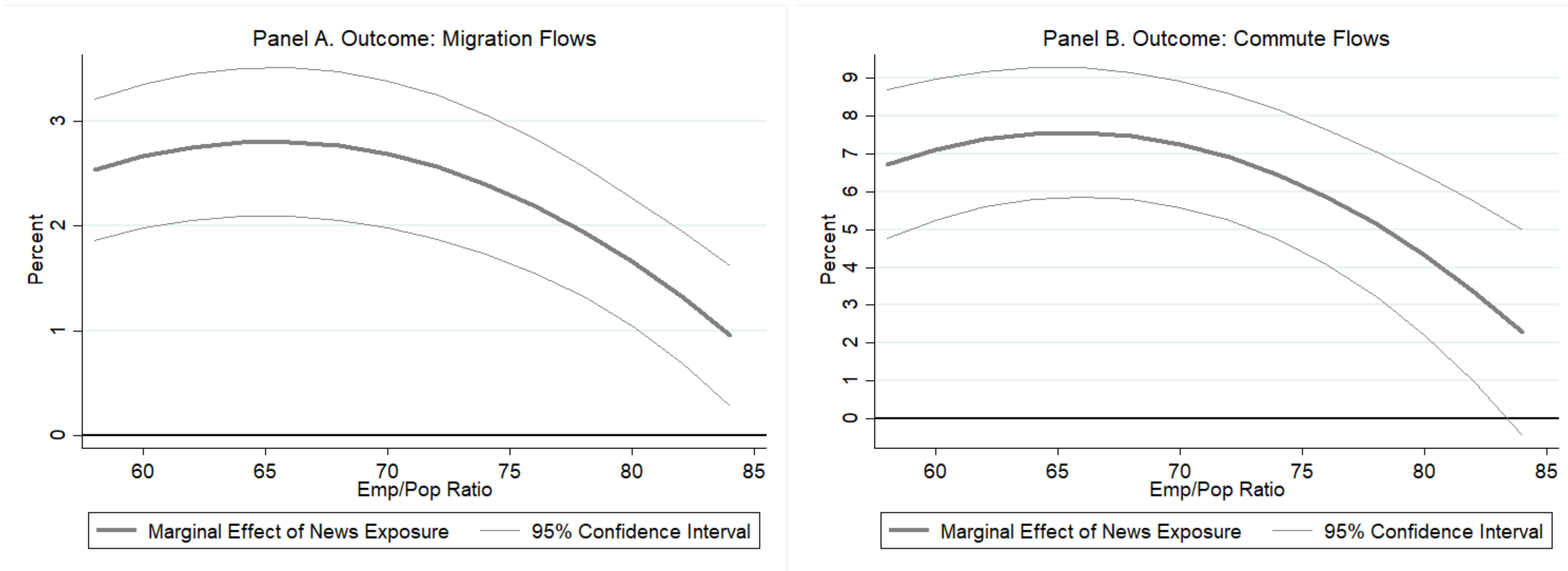

Figure 6

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.

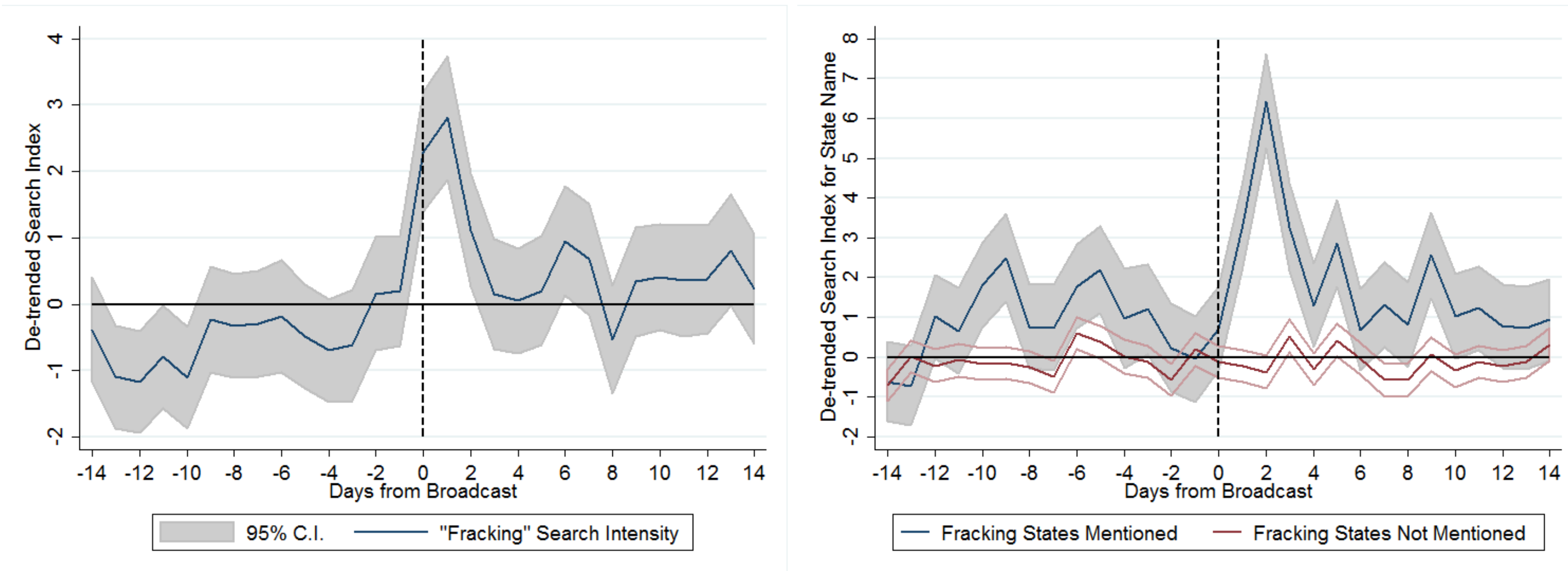

Figure 7

Google Search Interest in “Fracking” and Names of Fracking States Mentioned in 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. 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 A6). 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.

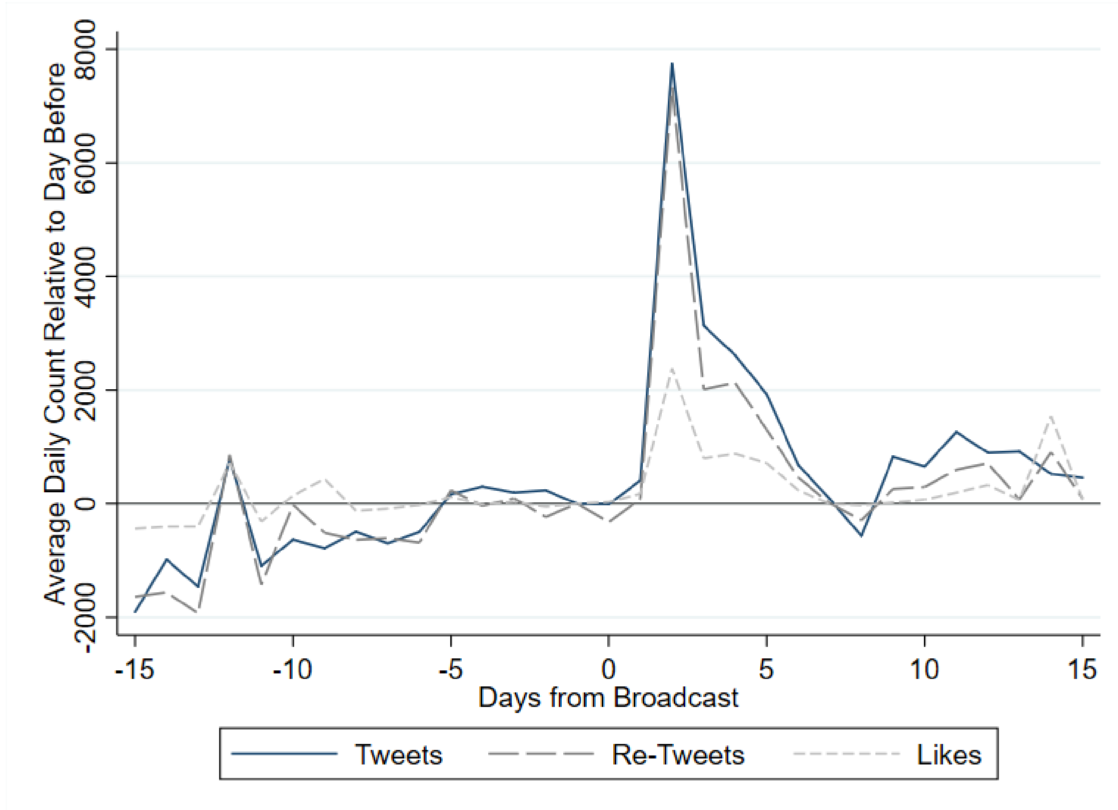

Figure 8

Average Tweets, Re-tweets, and Likes of Tweets about Fracking around a TV News Broadcast about Fracking

Notes: All tweets with the word "fracking" within 15 days of the post-2009 TV news broadcast events are included. The daily number of total tweets, re-tweets, and likes are averaged over all of the broadcast events. The Twitter-user's location is not attainable, and I am unable to exploit geographic variation in exposure to news about fracking.

Source: Author's calculations using TV news content from VTNA and Twitter tweet content.

For Online Publication: Appendix A. Additional Tables and Figures

Table A1

How Exposure to News about Fracking Impacts Awareness and Approval of Fracking

	All (1)	Male (2)	Female (3)	Under Median Age (4)	Over Median Age (5)	HS or Less (6)	Some College (7)	College Degree (8)
	Outcome: Heard About Fracking							
Newspaper Exposure	0.01***	0.01*	0.01***	-0.00	0.01***	0.01**	-0.00	0.01**
In 2011 (State-level)	(0.003)	(0.005)	(0.003)	(0.007)	(0.004)	(0.005)	(0.006)	(0.005)
Constant	0.62***	0.65***	0.62***	0.56***	0.69***	0.49***	0.64***	0.79***
	(0.017)	(0.032)	(0.021)	(0.022)	(0.025)	(0.026)	(0.037)	(0.020)
Observations	1,492	502	939	582	892	465	437	586
	Outcome: Heard A lot About Fracking							
Newspaper Exposure	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.01**	0.03***
In 2011 (State-level)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.006)	(0.006)
Constant	0.24***	0.25***	0.25***	0.20***	0.28***	0.11***	0.23***	0.41***
	(0.021)	(0.033)	(0.023)	(0.023)	(0.026)	(0.017)	(0.033)	(0.033)
Observations	1,492	502	939	582	892	465	437	586
	Outcome: Approve of Fracking							
Newspaper Exposure	0.01***	0.03***	0.00	0.02**	0.00	0.00	0.02***	0.01
In 2011 (State-level)	(0.003)	(0.006)	(0.003)	(0.007)	(0.004)	(0.006)	(0.005)	(0.010)
Constant	0.51***	0.47***	0.52***	0.49***	0.53***	0.56***	0.53***	0.44***
	(0.022)	(0.035)	(0.023)	(0.038)	(0.032)	(0.042)	(0.034)	(0.034)
Observations	1,030	357	652	346	673	252	300	477

Notes: Observations at the individual level from the Pew Research Center March 2012 Political Survey. Individuals were asked if they had heard “a lot”, “a little”, or “nothing at all” about fracking. Individuals who reported hearing “a lot” or “a little” abot fracking were then asked if they favor or oppose fracking. As such I am unable to estimate how hearing about fracking affects people’s approval of fracking. I can only estimate the impact of exposure to the news on approval, conditional on hearing about fracking. Newspaper exposure constructed from LexisNexis newspaper content and AAM newspaper circulation. The median age in the sample is 48. Observations are weighted using the weights provided by the Pew Research Center. Standard errors are corrected for clustering at the state-level. $p<0.01$ ***, $p<0.05$ **, $p<0.1$ *.

Table A2

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 A3

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	Above Median	Below Median	Above Median	
	(1)	(2)	(3)	(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. Standard errors are corrected for clustering at the origin DMA level. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A4

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		
	$\leq$\$1,250	\$1,250–\$3,333	$\geq$\$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 A5

Sensitivity to Various Fixed Effects

	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)	(9)	(10)
National Newspaper Exposure$_{oSt}$	0.024***	0.025***	0.026***	0.026***	0.027***	0.112***	0.068***	0.060***	0.040***	0.024***
	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.007)	(0.009)	(0.006)	(0.007)	(0.006)
National Newspaper Exposure$^2_{oSt}$	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.004***	-0.002***	-0.002***	-0.001***	-0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0002)
Origin by Destination F.E.	X	X	X	X	X	X	X	X	X	X
Destination by Year F.E.		X		X	X		X		X	X
Origin by Year F.E.			X	X	X			X	X	X
Origin State by Destination by Year					X					X
Dependent Mean (in Levels)	7.6	7.6	7.6	7.6	7.6	31.4	31.4	31.4	31.4	31.4
Observations	590,224	590,224	590,224	590,224	590,224	499,440	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. In all specifications origin/destination pair fixed effects are included to control for time-invariant differences across pairs. Columns (5) and (10) include origin state by destination by year fixed effects, thus exploiting variation across origin counties in the same state to the same destination state. 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

Sensitivity 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	Quadratic	Cubic	IHS	Linear	Quadratic	Cubic	IHS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
National Newspaper $\textit{Exposure}_{oSt}$	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 $\textit{Exposure}^2_{oSt}$		-0.001***	-0.0002			-0.002***	-0.006***	
		(0.0001)	(0.001)			(0.0004)	(0.001)	
National Newspaper $\textit{Exposure}^3_{oSt}$			0.00001				0.0001***	
			(0.00001)				(0.00002)	
Inverse Hyperbolic Sine of				0.046***				0.155***
National Newspaper $\textit{Exposure}_{oSt}$				(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* $\textit{Exposure}_{oSt}$ 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 A7

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	Include Fracking Origins	Exclude Non-Fracking Origins in Fracking States	Exclude Zero Circulation	Exclude Top One Percent of Exposure	Include NYC DMA	Include Fracking Origins	Exclude Non-Fracking Origins Fracking States	Exclude Zero Circulation	Exclude Top One Percent of Exposure
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
National Newspaper $\textit{Exposure}_{oSt}$	0.013***	0.025***	0.026***	0.024***	0.31***	0.045***	0.067***	0.066***	0.049***	0.123***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.006)	(0.006)	(0.009)	(0.009)	(0.011)	(0.013)
National Newspaper $\textit{Exposure}^2_{oSt}$	-0.0001**	-0.001***	-0.001***	-0.001***	-0.0003	-0.0004***	-0.002***	-0.002***	-0.001**	-0.007***
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.001)	(0.0001)	(0.0003)	(0.0004)	(0.0004)	(0.001)
Dependent Mean (in Levels)	7.7	12.2	2.1	11.1	6.1	31.2	62.2	1.3	43.5	26.9
Observations	596,256	639,840	398,944	392,832	505,664	504,544	541,504	337,568	332,416	427,840

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* $\textit{Exposure}_{oSt}$ 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 A8

Impact of Destination State Specific National Newspaper Exposure on Migration, Restrict Sample to Treatment Years

	Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$				Number of Migrants$_{oSt}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Sample: 2009-2012				
National Newspaper Exposure$_{oSt}$	0.018***	0.018***	0.015***	0.015***	1.416***	1.441***	1.314***	1.315***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.372)	(0.379)	(0.427)	(0.422)
National Newspaper Exposure$^2_{oSt}$	-0.001***	-0.001***	-0.001***	-0.001***	-0.034**	-0.035**	-0.027*	-0.031**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.020)	(0.020)	(0.022)	(0.021)
Mean Number of Migrants	8.2	8.2	8.2	8.2	8.2	8.2	8.2	8.2
Observations	181,696	181,696	181,696	181,696	181,696	181,696	181,696	181,696
				Sample: 2011-2012				
National Newspaper Exposure$_{oSt}$	0.014***	0.014***	0.013***	0.013***	0.358***	0.365***	0.219*	0.216*
	(0.004)	(0.004)	(0.005)	(0.005)	(0.128)	(0.129)	(0.126)	(0.129)
National Newspaper Exposure$^2_{oSt}$	-0.001***	-0.001***	-0.001***	-0.001***	-0.016***	-0.017***	-0.013**	-0.015**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.006)	(0.006)	(0.006)	(0.006)
Mean Number of Migrants	9.1	9.1	9.1	9.1	9.1	9.1	9.1	9.1
Observations	90,848	90,848	90,848	90,848	90,848	90,848	90,848	90,848

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. Each column replicates the corresponding estimate from Table 2, but restricts the years in the sample. 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

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

	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$_{oSt}$ Positive Flows in All Years
	As Reported	Lower Bound: Replace 0 with 9	As Reported	Lower Bound: Replace 0 with 9		
	(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. Estimates are similar if I instead impute missing migration flows as the average or maximum non-censored origin/destination flows rather than zero. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A10

Impact of Destination State Specific National Newspaper Exposure, Explicitly Modeling for Zero-values from Censoring

	Number of Migrants$_{oSt}$			Inverse Hyperbolic Sine of Number of Migrants$_{oSt}$	
	OLS	ZIP	Tobit	OLS	Tobit
	(1)	(2)	(3)	(4)	(5)
National Newspaper Exposure$_{oSt}$	6.22***	1.60**	9.93***	0.074***	0.107***
	(2.10)	(0.70)	(3.27)	(0.018)	(0.028)
National Newspaper Exposure$^2_{oSt}$	-0.22***	-0.04*	-0.36***	-0.003***	-0.004***
	(0.08)	(0.02)	(0.13)	(0.001)	(0.001)
Mean Number of Migrants	138.5	138.5	138.5	138.5	138.5
Observations	32,534	32,534	32,534	32,534	32,534

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. The outcome in columns (1)-(3) is the number of migrants. The outcome in column (4) is the inverse hyperbolic sine of the number of migrants. Origin counties with any fracking production or in the New York City designated market area are excluded. Relative to the baseline analysis sample, this sample is restricted to origin/destination pairs with at least one non-zero migrant flow between 2000 and 2012, as this is the dimension of the selection. Effects are larger for this subsample. For the zero-inflated Poisson estimation in column (2), the adult population (15-64) and the adult population squared are included in the logit first stage selection estimation. For the zero-inflated Poisson (ZIP) estimates, average marginal effects are reported. 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. Controls for the origin county unemployment rate, employment to population ratio, and average annual earnings (2010$) are 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 A11

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

	Inverse Hyperbolic Sine of the Migration Rate per 100,000$_{oSt}$				Migration Rate per 100,000$_{oSt}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
National Newspaper Exposure$_{oSt}$	0.020***	0.020***	0.023***	0.022***	0.400*	0.438*	0.663**	0.622**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.222)	(0.235)	(0.260)	(0.290)
National Newspaper Exposure$^2_{oSt}$	-0.001***	-0.001***	-0.001***	-0.001***	-0.018	-0.019*	-0.024**	-0.025**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.011)	(0.011)	(0.012)	(0.011)
Local Newspaper Exposure$_{oSt}$				0.013**				-0.629
				(0.005)				(2.617)
Local Newspaper Exposure$^2_{oSt}$				-0.0001**				0.063
				(0.00005)				(0.056)
Origin Labor Market Controls		X				X		
Origin by Year Effects			X	X			X	X
Origin/Destination Local News				X				X
Mean Migrants per 100,000	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0
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. The outcome is the number of migrants divided by the origin county population in 2000, in hundreds of thousands of people. 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 A12

Demographic Characteristics of *USA TODAY* Readership in 2007

	Readership Characteristics (1)	Fraction Adult Population (2)
Total Readers	3,882,000	0.017
Male	0.68	0.024
Female	0.32	0.011
18-49	0.63	0.018
25-54	0.66	0.20
35-54	0.50	0.022
40-59	0.48	0.022
No College	0.29	0.009
Some College (<4-year)	0.28	0.022
4-year Degree or More	0.43	0.029
Professional/Managerial	0.34	0.084
Top/Middle Manager	0.25	–
Employed	0.80	0.022
HH Income <50K	0.31	0.013
HH Income 50-75K	0.17	0.005
HH Income 75-100K	0.14	0.018
HH Income ≥100K	0.38	0.007
No Children	0.55	0.015
Any Children	0.45	0.020
Homeowners	0.74	0.018
Renters	0.26	0.016

Notes: Readership characteristics based on the MRI Fall 2007 Report (http://usatoday30.usatoday.com/media_kit/usatoday/au_general_demographics.htm). Age groups are as defined in the 2007 Report and are not mutually exclusive groups. Readers' education, income, presence of children, and homeownership status are grouped into mutually exclusive groups based on the reported groups from the 2007 report. The population measures in Column (2) are constructed using the ACS 2007 microdata for the full US population 18 and older. Top and middle managers are not identified in the ACS.

Table A13

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 A14

County Characteristics by Local Newspaper Circulation Among Counties at the Market Border

	Characteristic in 2000		Change from 2000 to 2010		Within Border Change from 2000 to 2010
	In-Market (1)	Out of Market (2)	In-Market (3)	Out of Market (4)	(5)
Migrants to fracking areas (Percent of Population)	0.22	0.24	-0.018	-0.021	0.00
Employment to Population (16+)	57.92	57.77	-0.80	-0.96	0.161*
Unemployment Rate	3.44	3.40	1.07	1.13	-0.06
Median Household Income	37,695	37,599	9,358	9,562	-185
Percent in Poverty	12.87	12.86	1.46	1.31	0.15*
Percent White	84.07	85.17	-2.08	-1.77	-0.31**
Percent Black	8.42	7.62	0.27	0.23	0.06
Percent Hispanic	6.75	6.90	2.44	2.28	0.20***
Percent Other Race	7.51	7.21	1.81	1.54	0.26***
Percent Population 20-34	18.78	18.08	-0.48	-0.70	0.20***
Percent Population 35-64	38.74	39.40	1.66	1.69	-0.04
Percent Population Over 64	14.54	14.44	0.95	1.30	-0.33***
Percent Households Renting	27.66	25.85	1.63	1.59	0.05

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. Local newspaper circulation data from the Alliance for Audited Media. Columns (1) and (2) report the means for counties inside versus outside the local newspapers' distribution markets. Columns (3) and (4) reports the average change from 2000 to 2010 for these groups. Column (5) reports the differences between Columns (3) and (4) when controlling for newspaper border, to compare counties in the same newspaper border. Standard errors corrected fro clustering at the origin DMA level. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A15

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 A16

Chain Migration: Impact of Lagged Newspaper Exposure on Migration

	Inverse Hyperbolic Sine of the Number of Migrants$_{oSt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
National Newspaper Exposure$_{oSt}$	0.025***	0.022***	0.022***	0.024***	0.022***	0.022***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)
National Newspaper Exposure$^2_{oSt}$	-0.001***	-0.001***	-0.001***	-0.001**	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
National Newspaper Exposure$_{oSt-1}$		0.007*	0.008**		0.007*	0.008**
		(0.004)	(0.004)		(0.004)	(0.004)
National Newspaper Exposure$^2_{oSt-1}$		-0.000	-0.000		-0.000	-0.000
		(0.000)	(0.000)		(0.000)	(0.000)
National Newspaper Exposure$_{oSt-2}$			-0.003			-0.004
			(0.006)			(0.006)
National Newspaper Exposure$^2_{oSt-2}$			0.001			0.001
			(0.001)			(0.001)
IHS of Number of Migrants$_{oSt-1}$				0.042***	0.042***	0.032**
				(0.012)	(0.012)	(0.012)
Dependent Mean	7.6	7.7	7.7	7.7	7.7	7.7
Observations	590,224	544,832	499,440	544,832	544,832	499,440

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 for the years specified. Origin counties with any fracking production or in the New York City designated market area are excluded. The variable National Newspaper Exposure 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 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 A17

Reverse Migration: Impact of National Newspaper Exposure at Origin on Future Migration from Destination

	Inverse Hyperbolic Sine of the Number of Migrants from S to o			
	(1)	(2)	(3)	(4)
National Newspaper Exposure$_{oSt}$	0.005	0.004	0.003	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
National Newspaper Exposure$^2_{oSt}$	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
National Newspaper Exposure$_{oSt-1}$		0.008*	0.006	0.007
		(0.005)	(0.005)	(0.005)
National Newspaper Exposure$^2_{oSt-1}$		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
National Newspaper Exposure$_{oSt-2}$			0.016**	0.017**
			(0.007)	(0.007)
National Newspaper Exposure$^2_{oSt-2}$			-0.003**	-0.003**
			(0.001)	(0.001)
National Newspaper Exposure$_{oSt-3}$				-0.012
				(0.026)
National Newspaper Exposure$^2_{oSt-3}$				-0.012
				(0.017)
Dependent Mean	7.4	7.4	7.4	7.5
Observations	590,224	544,832	499,440	454,016

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 for the years specified. Origin counties with any fracking production or in the New York City designated market area are excluded. The outcome is the number of migrants from fracking counties in state S to origin county o. National Newspaper Exposure is exposure at the origin about the destination. This specification will capture how news exposure in previous years affects reverse migration flows. 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 A18

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$_{oSt}$ *	-0.0006***	-0.0017**	-0.0100	-0.1350
Emp/Pop$_{ot-1}$	(0.0002)	(0.0006)	(0.0221)	(0.0958)
Newspaper Exposure$_{oSt}$ *	-0.0001***	-0.0002***	-0.0040***	-0.0189***
Emp/Pop$^2_{ot-1}$	(0.000)	(0.000)	(0.0013)	(0.0066)
Newspaper Exposure$^2_{oSt}$ *	0.000	0.000	-0.0009	0.0058
Emp/Pop$_{ot-1}$	(0.000)	(0.0001)	(0.0011)	(0.0043)
Newspaper Exposure$^2_{oSt}$ *	-0.000	0.000	-0.000	0.0008**
Emp/Pop$^2_{ot-1}$	(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 A19

State-level Events Controlled for in Google Trends State Specifications

Date (1)	Event (2)	States (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	All
Jan. 24, 2012	Obama State of the Union Mention of Fracking	All
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.

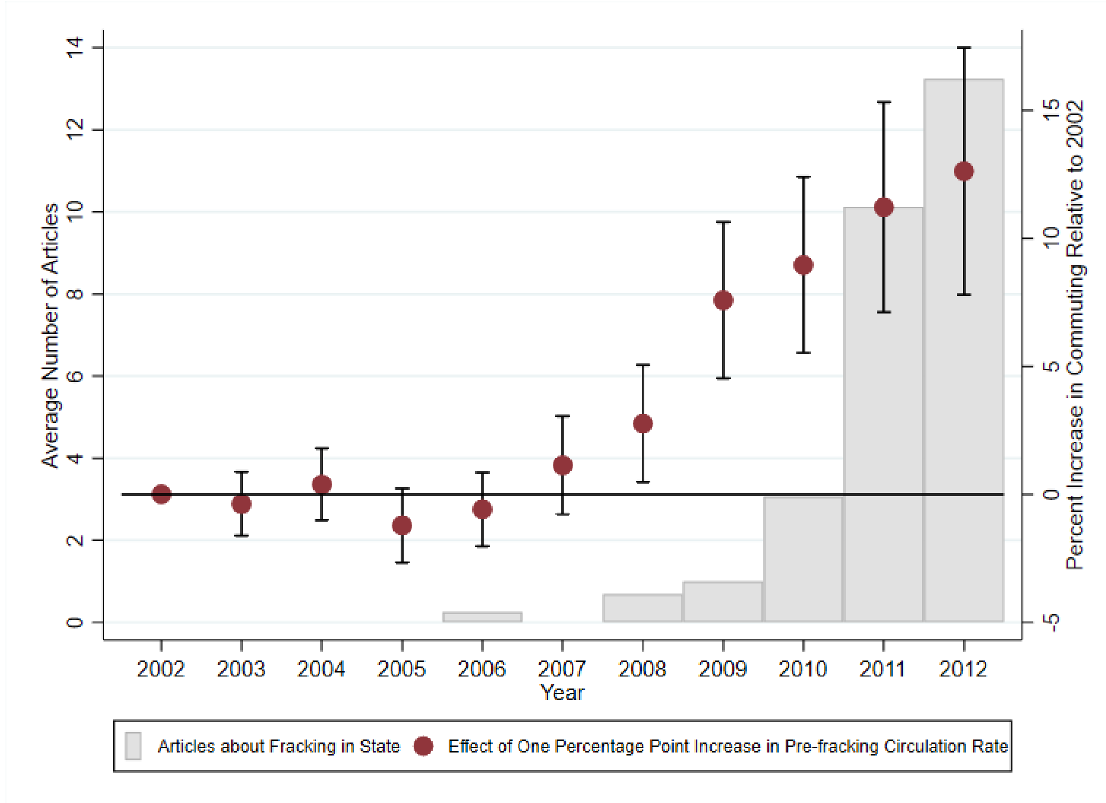

Figure A1

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.

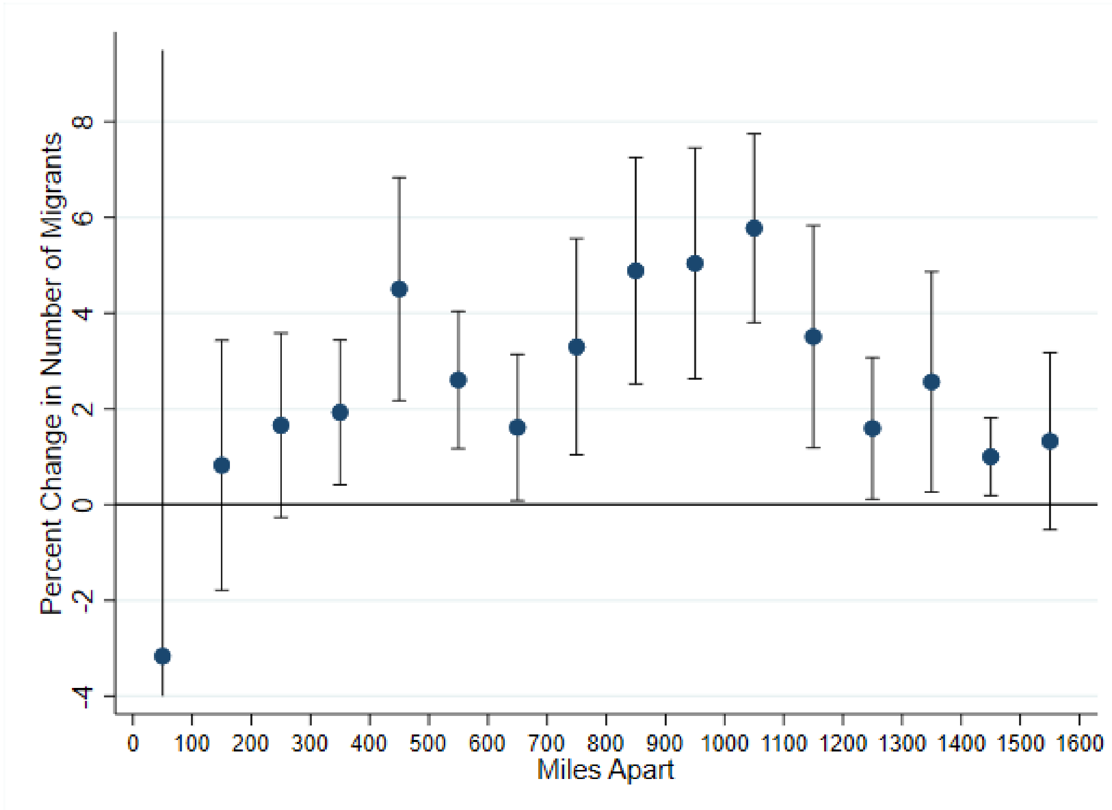

Figure A2

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

Notes: The total impact of a one unit increase in Newspaper Exposure is plotted with 95 percent confidence intervals, from equation (4), 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.

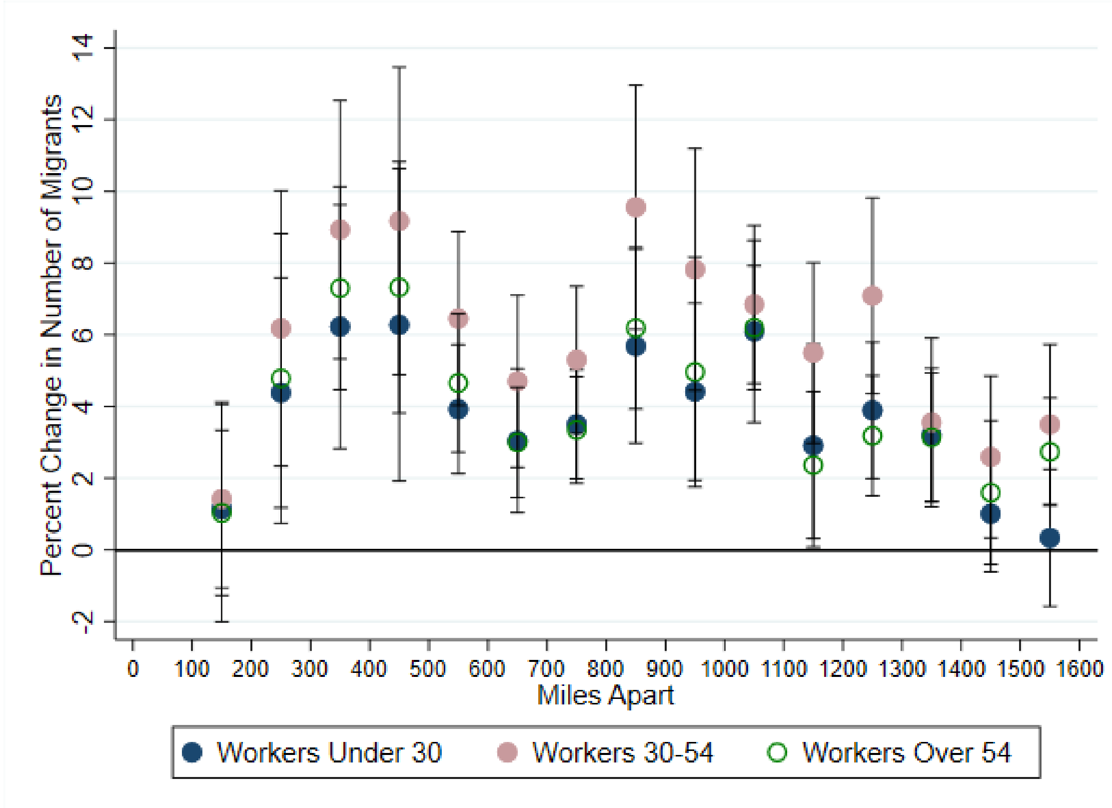

Figure A3

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

Notes: The total impact of a one unit increase in Newspaper Exposure is plotted with 95 percent confidence intervals, 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).

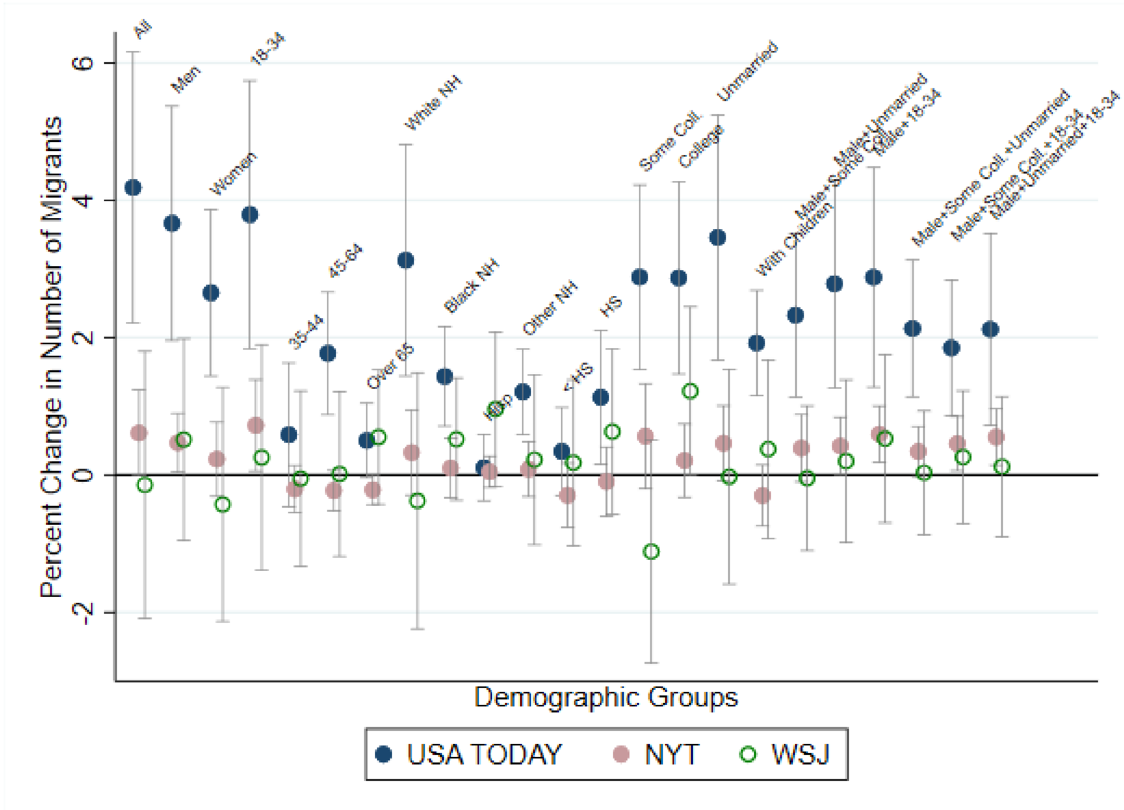

Figure A4

Heterogeneity by Newspaper: Migration Impact of Newspaper Exposure by Demographic Group

Notes: The total impact of a one unit increase in exposure for each of the three national newspapers is plotted with 95 percent confidence intervals. The level of observation is the origin MIGPUMA by destination state by year from 2005 to 2012 for the given group. Origin MIGPUMAs with any fracking production are excluded. 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. 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 origin state.

Source: Author's calculation using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and microdata from the 2005-2012 American Community Survey.

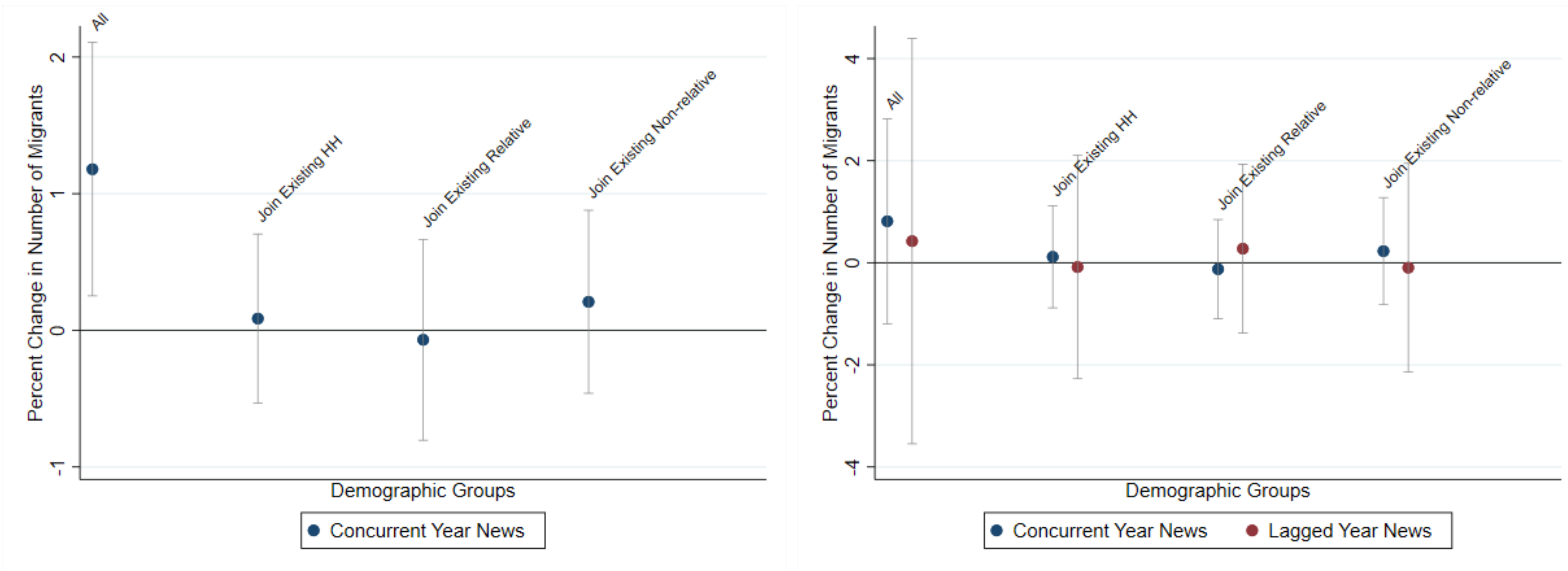

Figure A5

Household Chain Migration in ACS: Migration Impact of Newspaper Exposure on Migration into Pre-existing Households

Notes: The total impact of a one unit increase in exposure for each of the three national newspapers is plotted with 95 percent confidence intervals. The level of observation is the origin MIGPUMA by destination state by year from 2005 to 2012 for the given group. Origin MIGPUMAs with any fracking production are excluded. 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. 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 origin state.

Source: Author's calculation using circulation rates from the Alliance of Audited Media, newspaper content from LexisNexis, and microdata from the 2005-2012 American Community Survey.

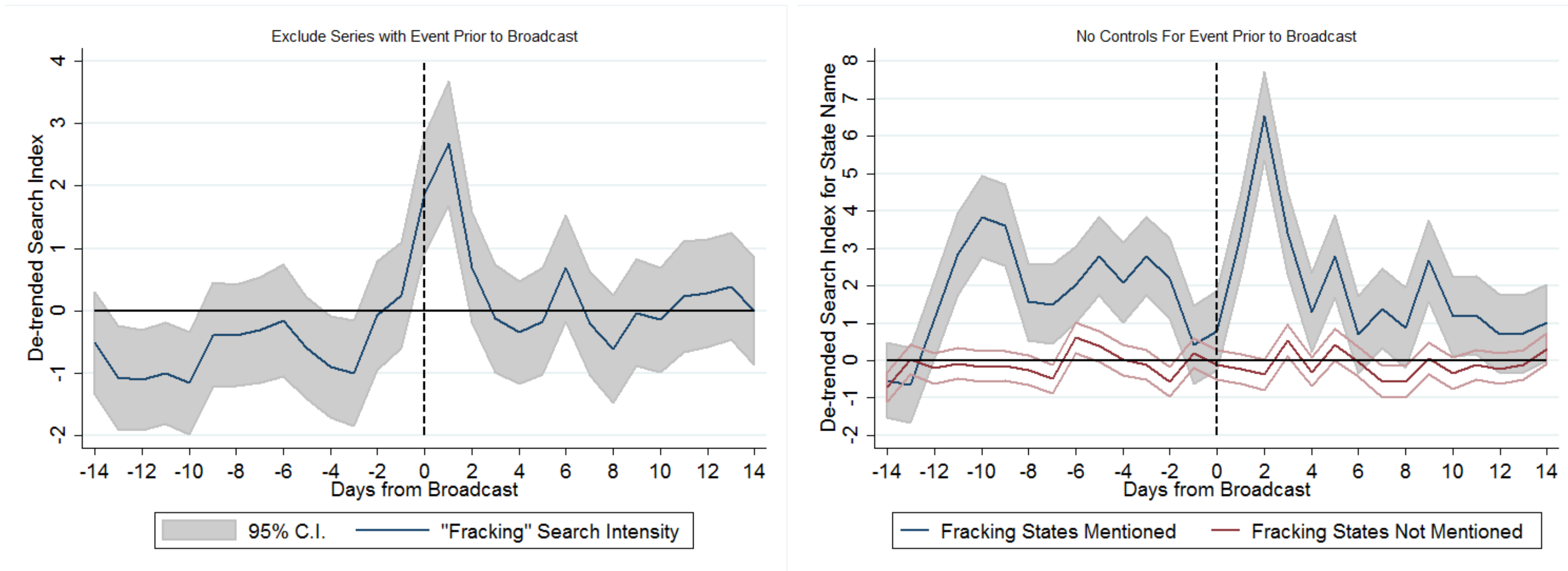

Figure A6

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.

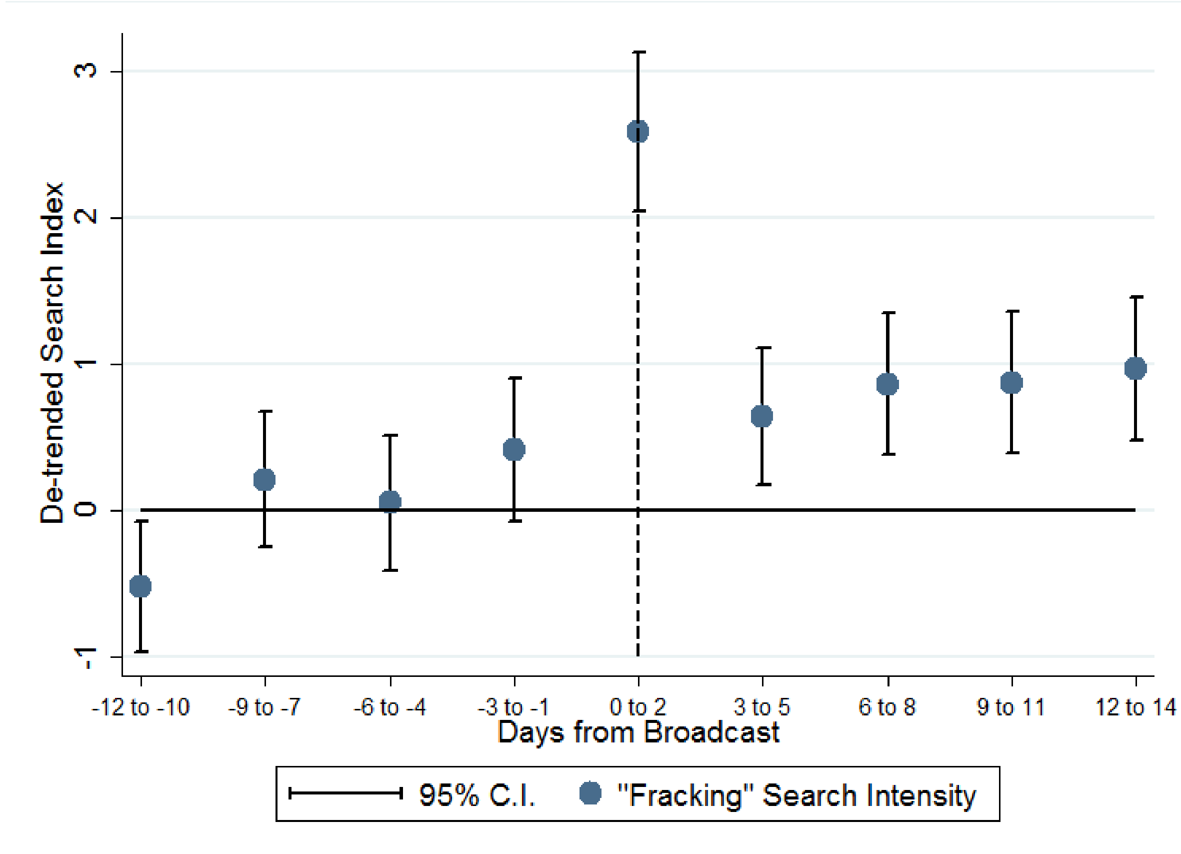

Figure A7

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 VIII, 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.

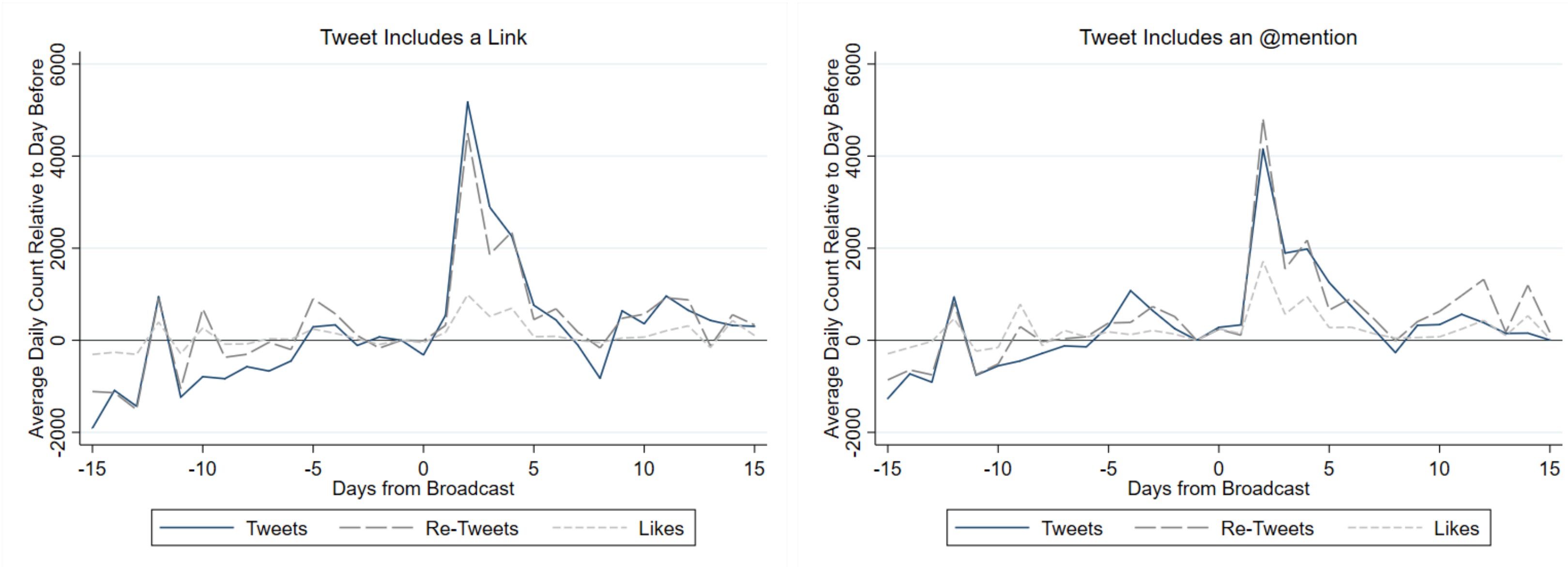

Figure A8

Average Tweets, Re-tweets, and Likes of Tweets about Fracking that Include Links or @ Mentions

Notes: All tweets with the word "fracking" and either a link or an @ mention within 15 days of the post-2009 TV news broadcast events are included. The daily number of total tweets, re-tweets, and likes are averaged over all of the broadcast events. The Twitter-user's location is not attainable, and I am unable to exploit geographic variation in exposure to news about fracking. Links might direct people to other websites with more information.

Source: Author's calculations using tweets collected from Twitter.

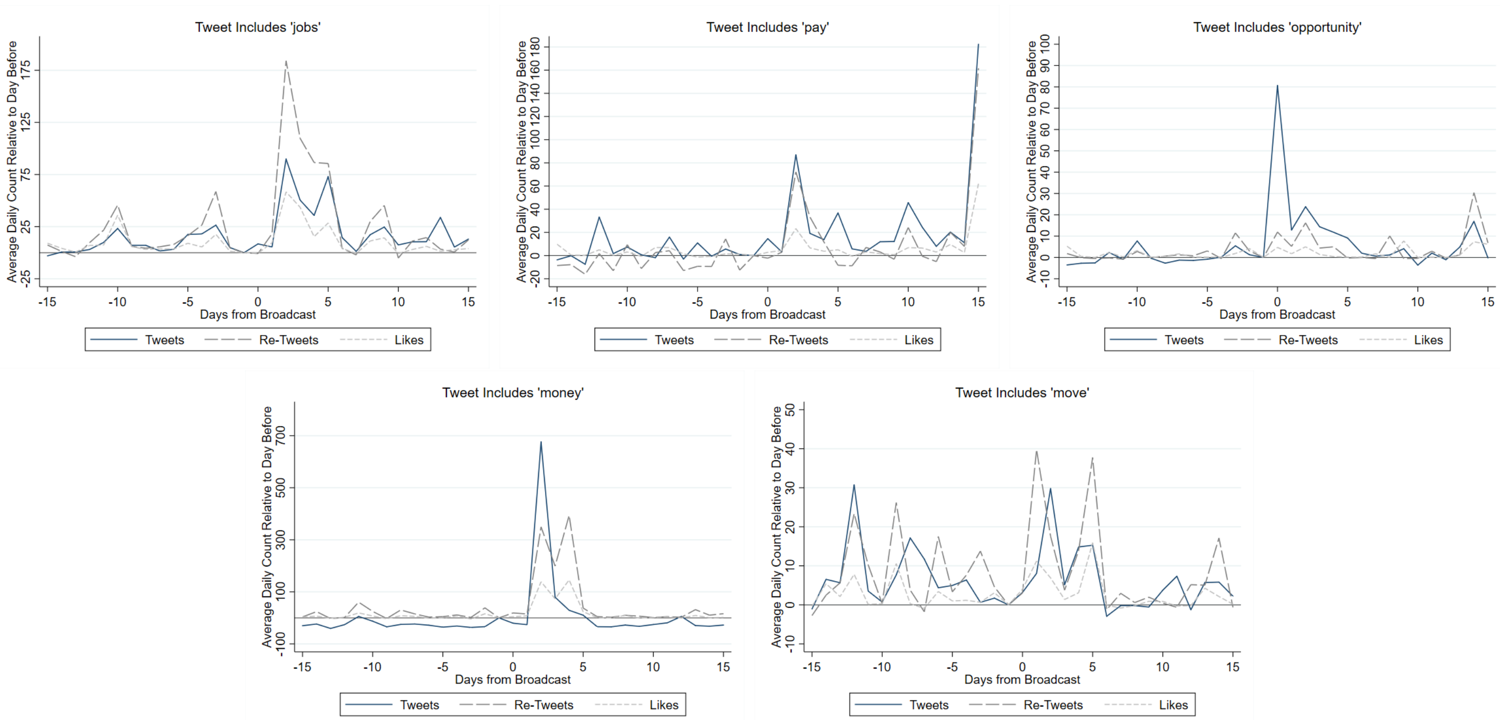

Figure A9

Average Tweets, Re-tweets, and Likes of Tweets about Fracking that Include Job Related Terms

Notes: All tweets with the word “fracking” and the listed word (i.e., “jobs”) within 15 days of the post-2009 TV news broadcast events are included. The daily number of total tweets, re-tweets, and likes are averaged over all of the broadcast events. The Twitter-user’s location is not attainable, and I am unable to exploit geographic variation in exposure to news about fracking.

Source: Author’s calculations using tweets collected from Twitter.

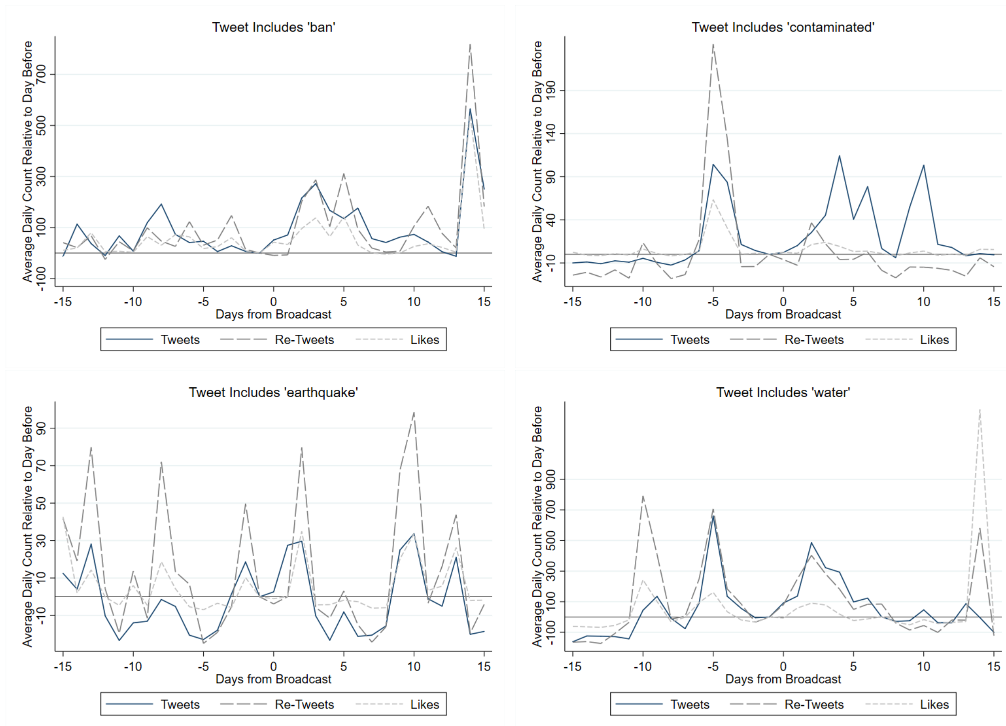

Figure A10

Average Tweets, Re-tweets, and Likes of Tweets about Fracking that Include Environmental Related Terms

Notes: All tweets with the word “fracking” and the listed word (i.e., “ban”) within 15 days of the post-2009 TV news broadcast events are included. The daily number of total tweets, re-tweets, and likes are averaged over all of the broadcast events. The Twitter-user’s location is not attainable, and I am unable to exploit geographic variation in exposure to news about fracking.

Source: Author’s calculations using tweets collected from Twitter.

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. Most filing occurs between February and April, so annual migration flows capture moves from approximately March or April from one year to the next.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).

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 2011, the IRS changed the methodology for constructing the migration flows. Prior to 2011, only returns filed before late September were included in the calculations. From 2011 on, returns filed through the end of December were included. This led to greater representation of high income households (Pierce, 2015). They also expanded the way that matches were identified to consider all heads, spouses, and dependents. Using both the new method and old method, the IRS calculated state-level net migration rates to determine how much the series was affected. They find that 44 of the states (plus DC) differed by less than 5 percent and only Wyoming varied by more than 10 percent. Geographic differences in how the methodological change affects the calculated flow is potentially problematic, but addressed in the specification. If flows into certain destinations were more affects, this will be picked up in the destination by year fixed effects. If flows out of certain origins were more affects, this will be picked up when I include the origin by year fixed effects. As seen above, this does not seem to affect the estimates. As a further check, I exclude each destination state one-by-one to see if the effect is only driven by one state. It is not. Also, the effects are essentially the same if I exclude fracking destinations where the methodological change had the biggest affect on flows: Wyoming, Louisiana, and Mississippi. In 2013 the suppression threshold increased from 10 households to 20. This led to considerably more suppression. Nearly 48 percent of the reported flows between 2000 and 2012 would have been suppressed. For this reason I limit my analysis to 2012. In the analysis sample from 2000 to 2012, 96.6 percent of flows between an origin county and fracking counties in one of the 16 destination states are reported as zero. Some of this likely due to censoring, but it is also probably that many counties do not send any migrants to a distant fracking state in any given year. Among origin/destination pairs that ever report a non-suppressed migration flow, only 39.4 percent of the annual flows are reported as zero. Even among pairs where the smallest non-zero migration flow was less than 20, only 52.5 percent of annual flows are reported as zero. These patterns would be consistent with censoring masking a lot of variation, but with most of the zero reported flows begin true zeros.

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. 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).

American Community Survey Microdata

To explore heterogeneous responses to news exposure I also exploit the 2005-2012 American Community Survey Microdata obtained from IPUMS (Ruggles et al., 2015). Each year the Census conducts the ACS which surveys approximately one percent of households each year. In addition to collecting information about household structure, demographics (age, race/ethnicity, gender, marital status), education, and employment, it asks individuals where they lived in the previous year. As such, I am able to construct origin/destination migration rates for demographic subgroups. Unfortunately, the geographic data is only available starting in 2005 and the smallest geographic unit is the Public Use Micro Area (PUMA). PUMAs are geographic areas defined by population that are large enough to preserve privacy. Furthermore, migration geographic data is only available at the Migration PUMA (MIGPUMA) level, which are often even larger. These MIGPUMA can contain one or more counties. There are several aspects of the data that are likely to reduce power, making it more difficult to detect an effect. First, because this is based off of a one percent sample of households there is likely to be measurement error in the constructed migration rates will reduce precision. Second, I have less geographic variation (and less variation in circulation rates) than is available at the county level leading to less precision.

To construct my estimates I use a county to MIGPUMA cross walk and collapse county-level measures to the MIGPUMA level. I then construct MIGPUMA origin to fracking state destination migration flows for each group by calculating the fraction of people in each group from each MIGPUMA origin that moved to a fracking MIGPUMA in a fracking state. This is analogous to the IRS SOI estimation but at the MIGPUMA rather than county level. After Hurricane Katrina, several PUMA and MIGPUMA in New Orleans, Louisiana were combined. I drop individuals living in these combined PUMA in 2005.

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.

Google Trends Data

For a given search term, Google will provide a measure of search intensity within a given geographic area over a specified time period. The geographic regions include the entire country, states, or the DMA-level. All of my analysis is conducted using DMA-level search indices. For a given term (like "fracking") Google will identify the day in the search window where the term is the most searched as a fraction of all searches. This day is assigned a value of 100. All other days will be indexed to that day. So a value of 20 would represent a search intensity of 20 percent. Because these indices are produced within the search area and time they are internally consistent but not comparable across queries. If there is not enough search for a given term over the period, no values are reported. 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.

Twitter Tweet Data

Twitter provides an API to access tweets, but only recent tweets (in the last 14 days) are recoverable. Because I am interested in historic tweets, I scrape them directly from Twitter using the GetOldTweets3 package in Python. Using this method I can only recover location if users include this with their post, and in practice this happens rarely. Instead, I consider the total tweet count for the entire country at the daily level. Using the GetOldTweets3 package, I collect all tweets that include the word "fracking" anywhere in the text. I also observe the entire text, the day and time, the number of likes and re-tweets, the sender, and whether any other Twitter users handles (@) were specifically mentioned. Tweets that are re-tweets (without additional text) are not included as separate tweets and only counted as re-tweets.

For Online Publication: Appendix C. Conceptual Model: Information and Migration

In the canonical migration choice model (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 real returns $y_d(t)$ and $y_o(t)$ for each period in each location (which account for earnings, cost of living, local amenities, and idiosyncratic fit) as well as the fixed utility cost c_{od} associated with migrating from o to d. These location specific returns can vary over time, and are discounted by β. Under complete information with no uncertainty, the decision to move from o to d (m_{od}) is as follows

$$m_{od} = \begin{cases} 1 \quad \textit{if} \quad \sum_{t=0}^{T} \beta^t u(y_d(t)) - c_{od} \geq \sum_{t=0}^{T} \beta^t u(y_o(t)) \\ 0 \qquad\qquad\qquad \textit{else} \end{cases} \tag{10}$$

However, individuals likely face incomplete information about conditions in the potential destination. As such, the individual views $y_d(t)$ as a random variable, where $y_d(t) \sim G(y;\theta)$.[50] Now, the individual will only migrate if

$$\sum_{t=0}^{T} \beta^t \big(Eu(y_d(t)) - u(y_o(t))\big) - c_{od} \geq 0 \tag{11}$$

where the E represents the expected value at time zero. Equation (2) yields a threshold moving cost (c^*_{od}) at which the individual is indifferent between staying and moving

$$c^*_{od} = \sum_{t=0}^{T} \beta^t \big(Eu(y_d(t)) - u(y_o(t))\big). \tag{12}$$

Changes in the parameters θ which govern the distribution of $y_d(t)$ will affect this cost threshold. For simplicity, consider the case where $y_d(t)$ is distributed normally with a mean (μ_d) and variance (σ^2) such that $\theta = \{\mu_d, \sigma^2\}$, and that the individual is risk averse with monotonic preferences (e.i., $u' > 0$ and $u'' < 0$). Under these assumptions

$$\frac{\partial c^*_{od}}{\partial \mu_d} = \sum_{t=0}^{T} \beta^t \frac{\partial Eu(y_d(t))}{\partial \mu_d} > 0 \text{ and } \frac{\partial c^*_{od}}{\partial \sigma^2} = \sum_{t=0}^{T} \beta^t \frac{\partial Eu(y_d(t))}{\partial \sigma^2} \leq 0. \tag{13}$$

Intuitively, as the mean increases, the individual places less weight on low values of y_d leading to higher expected utility and the individual is now willing to pay a larger moving cost. As Rothschild & Stiglitz (1970) show, 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.[51] The increase in variance leads to more risk, and the individual's moving cost threshold weakly falls, as she must be compensated by a lower cost to move to compensate for the added risk. If the individual's prior belief is that the return to migration

[50] The model implications are similar if the individual lacks information about conditions at the origin.

[51] 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 σ^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.

is low, she will be less willing to move (as seen in equation (13)). Similarly, if her prior is diffuse and the investment in migration appears more risky, she will also be less willing to move.

People might have incomplete information about the parameters that govern the distribution of $y_d(t)$ (μ_d and σ^2).[52] Receiving additional pieces of information about these parameters can lead individuals to update their beliefs about μ_d and σ^2, and potentially change migration behavior. This seems like a plausible scenario in the case of fracking where there was often little known about the places involved in fracking and there was a lot of information about fracking provided in the news. This corresponds to a framework of updating, where people recognize that they are missing information and have prior beliefs about the distribution of μ_d and σ^2 that they update upon receiving new information. This updating process will impact their perceived distribution of $y_d(t)$.

Suppose $y_d(t) = \mu_d + \varepsilon_d$, where μ_d and ε_d are independent and $\varepsilon_d \sim N(0, \sigma^2)$, but the decision-making individual has incomplete information about the mean of $y_d(t)$.[53] Given her prior observations of the world, she believes $\mu_d \sim N(\mu_{d0}, \lambda_0)$. Now suppose she is exposed to information (perhaps in the newspaper or through TV news) that gives information about $G_0 = (g_1, g_2, ..., g_{n0})$, draws from the distribution of $y_d(t)$, which are assumed to be independent. The individual can use this information to update their believes about the distribution of μ_d using a process like Bayesian Updating. According to Bayes Rule

$$P(\mu_d|G_0) = \frac{P(G_0|\mu_d)P(\mu_d)}{P(G_0)} \tag{14}$$

Because $P(G_0)$ is fixed constant, it is also true that

$$P(\mu_d|G_0) \propto P(G_0|\mu_d)P(\mu_d). \tag{15}$$

In other words, the individual's perceived distribution of μ_d, conditional on observing G_0 is proportional to the likelihood of observing G_0 given the parameter μ_d times the prior perceived distribution of μ_d. By observing both the likelihood of G_0 and the prior of μ_d, the individual can update her perception of the distribution of μ_d, which is known as the posterior. If she received new information sets $(G_0, G_1, G_2, ...)$ over time, this updating process can be repeated iteratively, allowing the individual to incorporate the new information and adjust her beliefs about the distribution of μ_d. For certain distributions, such as the normal distribution, the posterior and prior probability distributions from Bayesian Updating are in the same family of distributions, and the parameters of the posterior distribution are formulaically adjusted by the new data observed. In this case, the posterior distribution will be

$$P(\mu_d|G_0, \sigma^2) = N\big(\mu_d|\frac{\sigma^2}{n_0\lambda_0 + \sigma^2}\mu_{d0} + \frac{n_0\lambda_0}{n_0\lambda_0 + \sigma^2}\bar{x}, \frac{\sigma^2\lambda_0}{n_0\lambda_0 + \sigma^2}\big) \tag{16}$$

where $\bar{x}$ is the sample average, n_0 is the number of data points, μ_{d0} is the mean from the prior, σ^2 is the true variance of $y_d(t)$ and λ_0 is the variance from the prior probability distribution.[54] From this updating

[52]This type of incomplete information 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).

[53]In this scenario, the individual is assumed to know the variance of $y_d(t)$. A similar process arises if both the mean and variance are unknown, but now the posterior will vary across the two parameters.

[54]This is a general result in the Bayesian Updating literature (Murphy, 2007).

process, the individual can identify a new updated posterior mean and variance for the unknown distribution of parameter μ_d

$$\mu_{d1} = \frac{\sigma^2}{n_0\lambda_0 + \sigma^2}\mu_{d0} + \frac{n_0\lambda_0}{n_0\lambda_0 + \sigma^2}\bar{x} \tag{17}$$

$$\lambda_1 = \frac{\sigma^2\lambda_0}{n_0\lambda_0 + \sigma^2} \tag{18}$$

Although the individual is likely not performing these calculations to incorporate new information, using sample moments from the new information to update beliefs is reasonable at an intuitive level.[55] The posterior mean is a weighted average of the prior mean and the sample mean from the new draw of data. The weights correspond to the relative dispersion. An increase in the sample mean from the new information will lead to a larger posterior mean. The magnitude of this increase will depend on how precise or diffuse the individual's prior belief is. If the individual has a diffuse prior (λ_0 is small), she will attribute the difference between μ_{d0} and $\bar{x}$ to observational noise and not change her beliefs in response to the new information. If the individual is uncertain about her prior (λ_0 is large) she will put more weight on the new data and take differences in the sample average as a signal that her prior needs to be adjusted.

Notice that λ_1, the posterior variance of μ_d, is decreasing in n_0. In other words, as more pieces of information and data are received the individual becomes more confident in her prior. This will lead them to put less weight on future data as they become more confident in their prior. As such, the marginal impact of each additional piece of data will become smaller as more data is received, potentially leading to non-linear, decreasing returns to new information. The impact of each piece of information will depend on her confidence in her initial belief about μ_d.

As the individual updates her beliefs about μ_d this will shape her belief about $y_d(t)$, the return to moving to destination d. As μ_d changes her belief about the distribution of $y_d(t)$ can be computed by determining the likelihood of observing some return y given data points G_0 as

$$\begin{aligned} P(y|G_0) &= \int P(y|\mu_d)P(\mu_d|G_0)d\mu_d \\ &= \int N(y|m_d,\sigma^2)N(\mu_d|\mu_{d1},\lambda_1)d\mu_d \\ &= N(y|\mu_{d1},\lambda_1+\sigma^2). \end{aligned} \tag{19}$$

We can now determine how information affects the distribution of $y_d(t)$ (see Bishop (2006) for proof). The effect of additional information on the perceived mean of $y_d(t)$ will interact with prior beliefs. If the prior belief about average returns in a potential destination was lower than the news suggests, (a likely case in this scenario given that many fracking locations were rural and unknown and that much of the coverage of fracking was positive), additional data will increase μ_{d1}. As we see in equation (13) this will result in a higher moving cost threshold as the individual becomes more willing to move. The opposite case is also possible. If the individual's prior belief was that the average return in fracking locations were higher than portrayed in the news, new information will result in a lower μ_{d1} in the updating process, making migration less desirable.

[55] 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.

Obtaining more data will always reduce the dispersion of $y_d(t)$ by reducing λ_1 (confidence in prior of the mean). However, there will still be uncertainty associated with the distribution of returns (σ^2). Shifting upward her prior about the mean, holding all else equal makes moving less risky and will make migration more desirable for risk averse individuals as in equation (13).

People's responses to news about localized fracking booms might have followed a similar process. 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 occurring and change people's beliefs.[56] For example, individuals exposed to numerous newspaper articles and TV news broadcasts touting the local economic benefits of fracking in Texas might adjust their beliefs about the average or dispersion of the return 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.

As the individual incorporates new information about the parameters (μ_d or σ^2), she better understands the distribution of $y_d(t)$ and can compute the likelihood of observing the return y 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 μ_d. This in turn increases c^*_{od}, meaning she is more willing to move (see equation (13)).

Linking the Conceptual Framework to the Empirical Model

This conceptual framework is meant to highlight the potential role of labor market information, and represents a simplified version of the migration decision with only one origin and one destination. With any empirical strategy it is impossible to observe both the treated (individual receives new information) and counterfactual, untreated states of the world. For this reason, I will be comparing migration flows from different origins that faced different levels of news exposure. Thus, origin counties with less news exposure provide a counterfactual for other origins that received more news.

If the origins provide a reasonable counterfactual for each other, then this comparison would approximate the conceptual model and allow me to evaluate how additional information affects migration flows. To make the empirical analysis approximate the conceptual framework, I include destination state by year (ψ_{St}) fixed effect. These fixed effects make the regression a comparison between migration flows from different origins, that faced different levels of news exposure, to the same destination. This is similar to equation (12) where the threshold moving cost for each origin destination pair equals the difference in expected utility. The regression provides estimates of what happens to the relative level of migration as exposure to labor market news about destination d increases. This estimation is directly related to equation (13). If origin counties with more news exposure see relative larger increases in migration flows to destination d, this would consistent with the partial derivatives in equation (13), suggesting that the threshold moving cost increases as people's perception of μ_d changes as they receive more new.

For some origin destination pairs, the expected utility gaps (and thus threshold moving costs) will be bigger or smaller because of other local characteristics that impact utility, but are not explicitly modeled

[56]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.

in equations (10)-(12). For example, some origin destination pairs have persistent differences in weather and geography that impact location specific utility. These differences can result in origins providing worse counterfactuals for each other, masking any effects of news exposure. For this reason, I also exploit variation across time. With multiple periods for each origin destination pair, I can include origin by destination fixed effects which will account for persistent gaps in expected utility, making origins more similar, absent any news exposure. These fixed effects make the empirical model a comparison between relative changes in migration, but allow me to isolate variation similar to the conceptual framework.

The conceptual model would suggest that additional news will increase the individual propensity to move, which would result in an aggregate change in the migration rate. This simple framework assumes a homogeneous average return in destination d, μ_d. Under this assumption, exposure to the news would have a constant absolute effect on the migration rate from the origin. To estimate the average absolute effect of news exposure on migration in the empirical specification, we would use the migration rate, or the number of migrants over the fixed, baseline population at the origin as the outcome. The baseline population is used to avoid endogenous changes in the population and isolate variation due to migration (just as the news circulation population share is fixed in the construction of news exposure).

However, as noted above, differences in expected utility between origins and destinations will vary depending on local characteristics. It is possible these differences might have an interactive effect with news exposure, leading to heterogeneous average returns and heterogeneous treatment effects. For example, in origin destination pairs that have higher migration flows absent any labor market news, there might be more density around the migration cost threshold, meaning more people at the margin of moving. This could result in news exposure having a bigger effects in origin destination pairs that have higher pre-treatment migration flows, leading to heterogeneous treatment effects. In this setting, constant absolute effects on the migration rate seem unlikely. There could also be heterogeneous treatment effects across individual characteristics, such as age, gender, race, or marital status (consistent with Figure 4).

If we relax the assumption of homogeneous returns, the model would suggest a pattern of relative effects, rather than absolute effects. Suppose that instead of the return $y_d(t)$ having a homogeneous average return μ_d, there were heterogeneous return, $y_d(t) = f(X_i, X_{od}, News_{od}) + \varepsilon_d$. In this case, the expected return would be $E(y_d(t)) = f(X_i, X_{od}, News_{od})$. In other words, an individual's expected return depends on individual specific characteristics (X_i), origin destination pair specific characteristics (X_{od}), and the news people in o receive about conditions at d. For example, the expected return in a given location for women might be different than for men, or the expected return could depend on how similar the labor market composition (e.g., industry, skill composition) is between the origin and destination. To see how receiving positive news about opportunities in d affects the propensity to move, consider the derivative of c^*_{od} with respect to $News_{od}$.

$$\frac{\partial c^*_{od}}{\partial News_{od}} = \sum_{t=0}^{T} \beta^t \frac{\partial Eu(y_d(t))}{\partial f(X_i, X_{od}, News_{od})} \frac{\partial f(X_i, X_{od}, News_{od})}{\partial News_{od}} \tag{20}$$

If information in the news matters, we would expect positive news to have a positive impact on expected utility in destination d, or both partial derivatives to be positive. Note, however, that if for some element j of X_i or X_{od}

$$\frac{\partial^2 f(X_i, X_{od}, News_{od})}{\partial News_{od} \partial X_{ij}} \neq 0 \text{ or } \frac{\partial^2 f(X_i, X_{od}, News_{od})}{\partial News_{od} \partial X_{odj}} \neq 0 \tag{21}$$

the effect of the news will vary with an individual's or location's other characteristics and there will be heterogeneous treatment effects. With out making further assumptions about the utility function, it is unclear which elements of X_i or X_{od} would interact with information. However, if the news shifts people's priors about the returns to moving, it is likely to result in more migration for populations already at the margin of moving, which empirically would be groups with higher baseline migration rates. This would imply that the effect of the news would be larger in subgroups or origin/destination pairs with higher baseline migration rates. If this is the case, the effects might more closely approximate relative effects, rather than constant absolute effects. Estimating relative effects in the empirical specification can be done by considering the natural log of the migration rate $ln(\frac{Migrants_{odt}}{Population_o})$. This is equal to $ln(Migrants_{odt}) - ln(Population_o)$. If the baseline population is used, $ln(Population_o)$ is a constant, so if we were to take a derivative with respect to an independent variable, $News\ Exposure_{odt}$

$$\frac{\partial ln(\frac{Migrants_{odt}}{Population_o})}{\partial Migrants_{odt}} \frac{\partial Migrants_{odt}}{\partial News\ Exposure_{odt}} = \frac{\partial ln(Migrants_{odt})}{\partial Migrants_{odt}} \frac{\partial Migrants_{odt}}{\partial News\ Exposure_{odt}} \tag{22}$$

Examining the natural log of the number of migrants will be the same as examining the natural log of the migration rate, with a fixed population.

The inverse hyperbolic sine transformation approximates the natural log transformation but is defined at 0. It is as follows

$$IHS(Migrants_{odt}) = ln(Migrants_{odt} + \sqrt{Migrants_{odt}^2 + 1}) \tag{23}$$

Note that the derivative of the inverse hyperbolic sine transformation of the number of migrants, with respect to an independent variable, $News\ Exposure_{odt}$ is

$$\frac{\partial IHS(Migrants_{odt})}{\partial News\ Exposure_{odt}} = \frac{1}{\sqrt{Migrants_{odt}^2 + 1}} \frac{\partial Migrants_{odt}}{\partial News\ Exposure_{odt}} \tag{24}$$

As $Migrants_{odt}$ becomes larger, the first term converges to the natural log derivative and

$$\frac{\partial IHS(Migrants_{odt})}{\partial News\ Exposure_{odt}} \approx \frac{\partial ln(Migrants_{odt})}{\partial News\ Exposure_{odt}} \tag{25}$$

A similar relationship holds if we examine the inverse hyperbolic sine of the migration rate, if a fixed, baseline population $Population_o$ is used

$$\frac{\partial IHS(\frac{Migrants_{odt}}{Population_o})}{\partial News\ Exposure_{odt}} = \frac{1}{\sqrt{Migrants_{odt}^2 + Population_o^2}} \frac{\partial Migrants_{odt}}{\partial News\ Exposure_{odt}} \tag{26}$$

If $Population_o$ is small and fixed, as $Migrants_{odt}$ becomes larger

$$\frac{\partial IHS(\frac{Migrants_{odt}}{Population_o})}{\partial News\ Exposure_{odt}} \approx \frac{\partial IHS(Migrants_{odt})}{\partial News\ Exposure_{odt}} \approx \frac{\partial ln(Migrants_{odt})}{\partial News\ Exposure_{odt}} = \frac{\partial ln(\frac{Migrants_{odt}}{Population_o})}{\partial News\ Exposure_{odt}}. \tag{27}$$

Each of these transformations will yield approximately the same effect. When considering the migration rate, I measure the number of migrants per 100,000 people. Thus, population is measured in hundreds of thousands of people. In 2000, the average county population was 89,539 (or 0.895 when measured in hundreds of thousands of people). The median county population was 24,527 (0.245). Less than 17 percent of counties had a population over 100,000 and less than 9 percent had a population over 200,000. By dividing by the population in hundreds of thousands, population is small, leading to a closer approximation in equation (27).

In the empirical analysis, I use the number of migrants or the inverse hyperbolic sine of the number of migrants as my preferred outcome. Bellemare & Wichman (2020) have derived the algebraic elasticities and semi-elasticities with the inverse hyperbolic sine. For large values of the un-transformed dependent variable (e.g, the number of migrants), the arcsinh-linear specification directly yields the semi-elasticity, just as the log-linear specification would. Their work suggests two potential concerns that might arise in the empirical specification at hand.

First, the estimated semi-elasticity will be biased if the mean of the un-transformed dependent variable is small. Using simulations, they suggest a mean of the un-transformed dependent variable greater than 10 will result in little to no bias as a rule of thumb (Bellemare & Wichman, 2020). For my full sample the mean annual number of migrants is 7.6, suggesting my semi-elasticity estimates are slightly biased. This bias is multiplicative, with the magnitude $\frac{\sqrt{y^2+1}}{y}$, where y is the mean of the un-transformed dependent variable (the number of migrants). This would suggest that my migration semi-elasticity estimates are biased upward by 0.9 percent ($\frac{\sqrt{7.6^2+1}}{7.6} = 1.009$). In other words, rather than a 2.4 percent effect from Table 1, the semi-elasticity is closer to $\frac{2.4}{1.009} = 2.37$ percent. As such, any biased due to small outcome values is likely to be small.[57] The mean number commuters in my analysis sample is 31.4. Based on Bellemare & Wichman's (2020) results, this should have minimal impact on the semi-elasticity estimates.

Second, Bellemare & Wichman (2020) caution that when dealing with data with many zero-valued observations it might be best to explicitly model the data generating process, using a Tobit or zero-inflated Poisson model. As a rule of thumb, they suggest modeling this selection explicitly if over one third of the observations are zero-valued. However, as they note this cutoff is entirely arbitrary. In my analysis, there is a large number of origin/destination flows that are zero. Some of these are true zeros, but it is likely that many are censored flows to protect privacy. Following Bellemare & Wichman's (2020) suggestion, I explicitly model the selection process using both zero-inflated Poisson and Tobit models. It should be noted, that neither of these methods are designed for panel data with fixed effects. When I estimate Tobit and zero-inflated Poisson models, I find significant increases in the number of migrants that are consistent with the baseline OLS estimates (Appendix Table A10).

Regardless of whether exposure to labor market news has absolute or relative effects on migration rates and flows, the number of migrants and the inverse hyperbolic sine of the number of migrants provide a weaker test of whether or not the news increases migration. Rather than impose functional form assumptions on the migration decision, I also show that the pattern is significant when examining the migration rate to capture constant absolute effects or inverse hyperbolic sine of the migration rate (Appendix Table A11) and is robust to various empirical issues that arise using the inverse hyperbolic sine (Appendix Tables A7 and A9).

Simulated Effects of News Information on Updating

To visualize how information can change migration decisions when individuals lack information a simple simulation is shown in Appendix Figure C1. Two scenarios are presented for three types of people. Individual 1 has a diffuse prior over the expected return (μ_d) and incorrectly believes μ_d is lower than the true mean. Individual 2 has a precise belief that μ_d is low. Individual 3 has a diffuse prior, but correctly predicts μ_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 μ_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 μ_d and y_d. Additional information also reduces dispersion of μ_d and y_d, which

[57] In my baseline specifications, *News Exposure*$_{odt}$ is included quadratically. The quadratic in the inverse hyperbolic sine transformation-linear specification leads to the same semi-elasticity as the log-linear specification ($\hat{\beta}_1 x + 2\hat{\beta}_2 x^2$) with the same bias $\frac{\sqrt{y^2+1}}{y}$.

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.

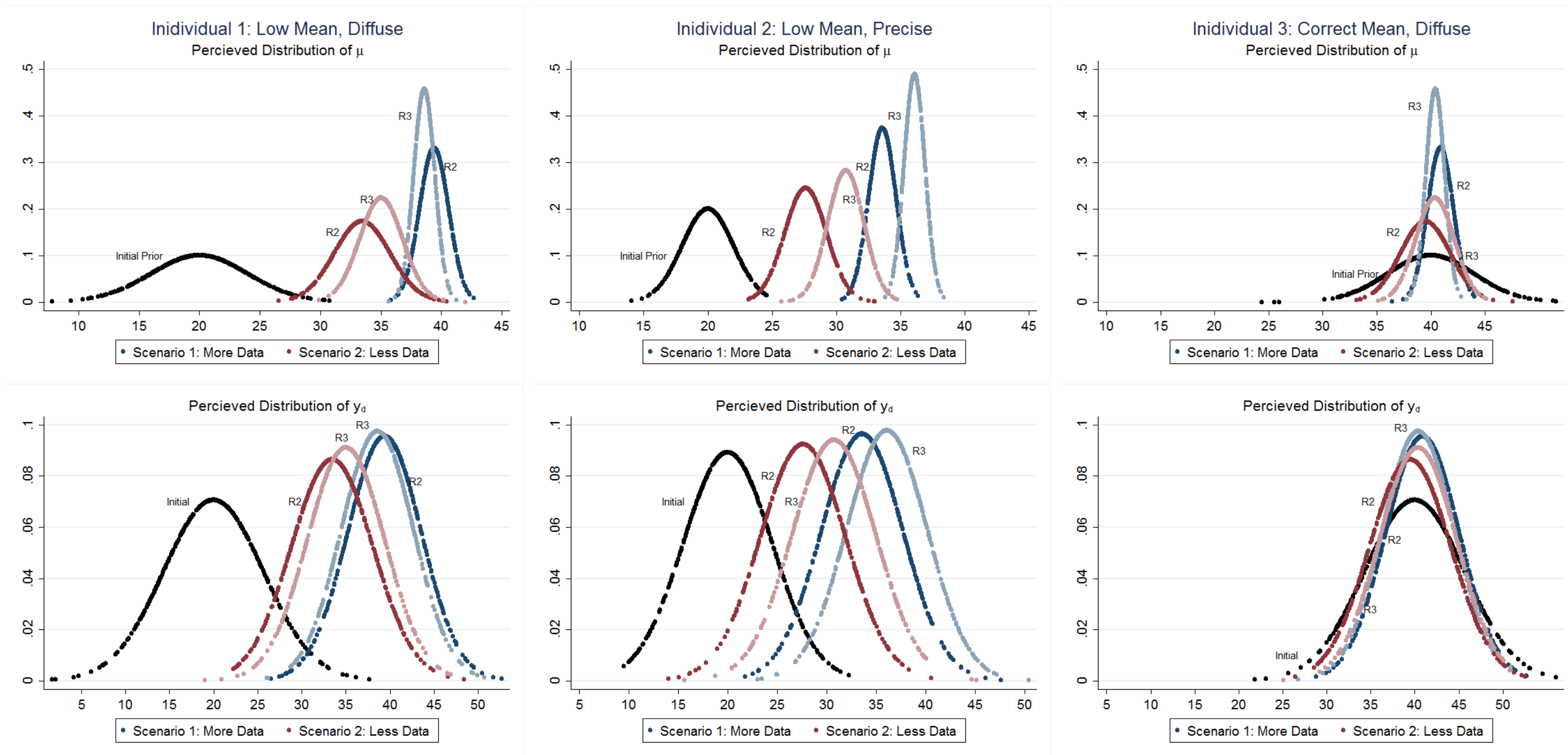

Figure C1

Model Simulations: Information and Bayesian Updating

Notes: Simulated data points from the distributions of μ_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.