Does Local Access to Finance Matter?

Evidence from U.S. Oil and Natural Gas Shale Booms*

Erik Gilje†

Job Market Paper

November 13, 2012

Abstract

I use oil and natural gas shale discoveries as a natural experiment to identify whether local access to finance matters for economic outcomes. Shale discoveries lead to large unexpected personal wealth windfalls, which result in an exogenous increase in local bank deposits and a positive local credit supply shock. Using this shock I examine whether local credit supply influences economic outcomes and how local banking market structure affects the importance of credit supply. After a credit supply shock, the number of business establishments in industries more reliant on external finance increases 4.6% relative to those less reliant on external finance. This increase is more than five times higher in counties dominated by small banks relative to all other counties. Local credit supply still matters despite the use of improved lending technology, the increased securitization of loans, and banking deregulation.

*I would especially like to thank Phil Strahan for his comments and advice. I would also like to thank Ashwini Agrawal, Allen Berger, David Chapman, Thomas Chemmanur, Jonathan Cohn, Simon Gilchrist, Evgenia Golubeva, Edith Hotchkiss, Steven Kaplan, Sari Kerr, Darren Kisgen, Elena Loutskina, Tobias Moskowitz, Ramana Nanda, Jonathan Reuter, David Robinson, Jérôme Taillard, Bent Vale, and participants at the 2012 Kauffman Entrepreneurship Mentoring Workshop, 2012 Western Finance Association Annual Meeting, 2012 Financial Intermediation Research Society Conference, 2012 European Finance Association Annual Meeting, 2012 BC/BU Green Line Meeting, and Steven C. Agee Economic Research and Policy Institute for helpful comments and suggestions. Additionally, I would like to thank Evan Anderson, Registered Professional Landman, for background and expertise on oil and gas leasing. I would like to also thank the Ewing Marion Kauffman Foundation for providing financial support for this project as part of the Kauffman Dissertation Fellowship program. All errors are my own.

†Carroll School of Management, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467. Email: gilje@bc.edu

1 Introduction

In perfect markets, entrepreneurs and firms should be able to obtain funding for all positive net present value projects. In such a world, changes in the availability of local sources of capital would have no effect on economic outcomes. However, if informational, regulatory, or other frictions interfere with financing then suboptimal economic outcomes may occur. In the United States, significant progress has been made to mitigate frictions and reduce the importance of geographic proximity for bank lending relationships (Petersen and Rajan (2002), Berger (2003)). These advances include increased use of credit score models and securitization, as well as state level banking deregulation. The contribution of this study is to examine whether local credit supply still matters for economic outcomes and how local banking market structure affects the importance of credit supply. I use exogenous variation in local credit supply to document that local credit supply still matters for economic outcomes, especially in areas dominated by small banks.

I use shale discoveries as a natural experiment to obtain exogenous variation in local credit supply. I identify shale discoveries ("booms") at the county level in seven states between 2003 and 2009 using a unique dataset of 16,731 individual shale wells. Unexpected technological breakthroughs in shale development have caused energy companies to make high payments to individual mineral owners for the right to develop shale discoveries. I find that the increase in individual mineral wealth associated with shale booms raises local bank deposits by 8.2%. More deposits enhance a bank's ability to make loans, resulting in a positive local credit supply shock.

To assess how this credit supply shock affects economic outcomes I compare the number of business establishments before a boom to after a boom in industries with different external financing needs. After a boom, the number of business establishments in industries with high dependence on external finance increases 4.6% relative to industries with low external finance needs.[1] This difference increases to 13.5% in counties where small banks have high market

[1] I have excluded all economic outcome measures directly related to oil and gas extraction, construction, real estate, and financial services, because economic outcomes for these industries potentially improve due to reasons unrelated to better local credit supply.

share relative to 2.5% in all other counties. These results suggest that local credit supply is still important for economic outcomes, particularly in areas dominated by small banks.

To formally test whether credit supply is more important in areas dominated by small banks I undertake a triple differencing strategy, by comparing how outcomes differ in boom counties dominated by small banks relative to all other counties. This specification is a comparison of 1) boom county-years vs. non-boom county-years 2) high external finance dependent industries vs. low external finance dependent industries and 3) small bank dominated counties vs. all other counties. This empirical strategy is a direct test of whether local credit supply shocks affect counties dominated by small banks differently.

Using a triple differencing strategy also addresses potential alternative interpretations of a basic differences-in-differences approach that only compares industries with different external financing needs after a boom. For example, some industries could benefit differentially from a shale discovery due to consumer demand shocks, wealth shocks, or other non-credit based shocks associated with a shale discovery. If any of these shocks are correlated with external financing needs, then a credit supply based interpretation of the results could be problematic. However, for these alternative shocks to alter the interpretation of the triple differencing specification, they would also need to be correlated with the size of a county's local banks. In fact, I find no evidence that after booms demand shocks differ across counties with different bank sizes. Specifically, retail sales, a proxy for demand, increase by similar amounts after booms in counties dominated by small banks as other counties. Additionally, there is no evidence that deposits grow faster after booms in counties dominated by small banks than in other counties, as one might expect if demand shocks affected counties differently. More broadly, the empirical design of this paper requires an alternative, non-banking based, interpretation of results to reconcile why outcomes for industries with distinct external financing needs respond differently after a shale boom, and why these different responses are larger in counties dominated by small banks.

In addition to relying on a triple differencing strategy to rule out alternative explanations, I undertake a variety of robustness and falsification tests. Using falsification tests I show that the results of this paper are not driven by pre-existing growth trends. I also conduct a separate falsification test to examine whether general non-shale growth shocks impact counties dominated by small banks differently and find no evidence that counties dominated by small banks disproportionately benefit from growth shocks relative to other counties. I demonstrate that the main results of this study are not driven by any single industry or industry exposure to economic fluctuations as proxied by industry asset beta. Additionally, I conduct robustness tests related to local banking structure and find that my main results are not driven by changes to local banking markets after a boom, different small bank size definitions, or banks that are part of holding companies.

The evidence from my study indicates that local credit supply is still important, however existing literature suggests that geographic proximity to finance matters less than in the past. For example, Petersen and Rajan (2002) state: "...technology is slowly breaking the tyranny of distance, at least in small business lending." They document that between 1973 and 1993, the distance between lenders and small firms increases, and that small firms communicate with their lenders less in person. Further research documents increases in borrower-lending distances from 1993 through 2001 (DeYoung et al. (2011)). However, Becker (2007) studies the U.S. banking system from 1970 to 2000, and using senior citizens as an instrument for local deposit supply, argues a causal link between local deposit supply and local economic outcomes.

I extend the existing literature by studying both if and where local credit supply matters after the erosion of frictions from banking deregulation, lending technology, and securitization.[2] Prior literature documents the importance of local credit supply (Peek and Rosengren (2000), Ashcraft (2005), Becker (2007)). However, there is little evidence on where local credit supply might be most important or whether local credit supply is still important after the widespread use of credit score models and securitization. Additionally, while significant research is devoted to bank size and borrower type (Strahan and Weston (1998), Berger et al. (2005), Berger et al. (2007)), and bank size and access to funds (Houston et al. (1997), Jayaratne and Morgan (2000), Kashyap and Stein (2000), Campello (2002)), far less research

[2]I follow the approach of other studies and focus on economic outcome variables, because detailed bank level loan data is typically unavailable.

examines whether local bank size influences the importance of local credit supply. Due to endogeneity concerns, the field often has challenges in cleanly identifying these questions.

Why might local credit supply be important, particularly in counties dominated by small banks? If local banks are large, capital can be redeployed geographically to fund projects. However, if local banks are small it could be more difficult for capital to be redeployed from elsewhere to be lent locally.[3] Furthermore, small banks are typically more reliant on deposit funding than large banks, and may have more challenges in obtaining external capital. Prior research also suggests that small banks may be more adept at lending to "soft" information borrowers (Stein (2002), Berger et al. (2005)). If areas with more small banks have more "soft" information borrowers, the inability of a small bank to obtain outside funding for these types of borrowers would also lead to worse economic outcomes. The ultimate set of frictions influencing outcomes could be frictions between borrowers and banks as well as frictions between banks and funding sources. This study provides new evidence that due to these frictions local credit supply is most important in areas dominated by small banks.

In Section 2 I provide an overview of the hypothesis tested in this study and the related literature. Section 3 provides detail on my identification strategy and background on my natural experiment. Section 4 discusses data and variable definitions. Section 5 discusses my results, and Section 6 concludes the paper.

2 Hypothesis Development and Related Literature

The underlying research question in this paper: "Does local access to finance matter?" is a dual hypothesis test of two sets of frictions 1) frictions between borrowers and banks 2) frictions between banks and access to funds for lending. Both sets of frictions have to be present for the observed results.

If firms could seamlessly access capital regardless of location, then neither local credit supply, local banking characteristics, nor a local bank's ability to obtain external funds

[3]Prior research discussing this issue includes Houston et al. (1997) and Jayaratne and Morgan (2000)

for lending would matter for local economic outcomes. Any local negative credit shock would be counteracted by distant lenders stepping in to fund positive net present value projects. Recent research suggests that geography and distance currently play less of a role in enhancing informational frictions between borrowers and banks due to improved use of information technology. Berger (2003) documents the rise of internet banking, electronic payment technologies, and credit scoring, while (Loutskina and Strahan (2009)) document the importance of securitization. These advances would suggest a reduced importance of local access to finance, because borrowers can more easily convey information about themselves to banks that are farther away.

Regulatory based frictions in the U.S. have also been eroded over time, reducing the importance of distance in lending relationships. Banking deregulation in U.S. states has affected output growth rates (Jayaratne and Strahan (1996)), the rate of new incorporations (Black and Strahan (2002)), the number of firms and firm-size distribution (Cetorelli and Strahan (2006)), and entrepreneurship (Kerr and Nanda (2009)). Additionally, Bertrand et al. (2007) document that banking deregulation in France leads to better allocation of bank loans to firms and more restructuring activity.

If distance does aggravate information based frictions between borrowers and lenders, then local credit supply may matter. In particular, if the cost to overcoming distance related frictions is prohibitive as could be the case with "soft" information borrowers[4], then local credit supply could be important. In this setting, the frictions that a bank faces in obtaining external funding become important for local economic outcomes. Existing literature suggests that bank size is a key characteristic along which frictions in obtaining external capital may vary. Kashyap and Stein (2000) document that monetary policy influences lending for small banks more than for large banks, while Bassett and Brady (2002) document that small banks rely more on deposit funding. Smaller banks also have fewer sources of funding outside a local area (Houston et al. (1997), Jayaratne and Morgan (2000), Campello (2002)). If small banks need to raise capital externally, while large banks can redeploy capital internally across

[4]Small banks may focus more on relationship lending based on "soft" information relative to transaction lending (Berger and Udell (2006)).

different geographic regions, then areas with more small banks may have more agency and informational issues related to obtaining external funding. These bank funding frictions may mean that areas with a higher proportion of small banks could be less likely to have access to funding beyond local deposits.

This paper is also more broadly related to other papers which use natural experiments to document the importance of access to finance for economic outcomes in different settings earlier in the United States (Peek and Rosengren (2000), Ashcraft (2005), Chava and Purnanandam (2011)) and internationally (Khwaja and Mian (2008), Iyer and Peydro (2011), Schnabl (2011), Paravisini (2008)). In other related work, Guiso et al. (2004) use Italian data to document the importance of financial development on new firm entry, competition, and growth. Recent literature has also used natural experiments in the U.S. to document the importance of local access to finance for productivity (Butler and Cornaggia (2011)) and risk-management (Cornaggia (2012)). Additionally, Plosser (2011) uses shale discoveries as an instrument for bank deposits, but focuses on bank capital allocation decisions during financial crises. My contribution differs from these papers in that: 1) I present evidence on the importance of local credit supply after banking deregulation, and wide adoption of new lending technology and securitization 2) I document that local credit supply is particularly important in areas dominated by small banks.

3 Identification Strategy: Unexpected Development of Shale

3.1 Natural Gas Shale Industry Background

The advent of natural gas shale development is one of the single biggest changes in the U.S. energy landscape in the last 20 years. According to the U.S. Energy Information Agency, in its 2011 Annual Energy Outlook, there are 827 Trillion Cubic Feet (Tcf) of technically recoverable unproved shale gas reserves in the United States, this estimate is a 72% upward revision from the previous year. 827 Tcf of natural gas is enough to fulfill all of the United

States' natural gas consumption for 36 years. On an energy equivalent basis 827 Tcf represents 20 years of total U.S. oil consumption or 42 years of U.S. motor gasoline consumption. As recently as the late 1990s, these reserves were not thought to be economically profitable to develop, and represented less than 1% of U.S. natural gas production. However, the development of the first major natural gas shale "play" in the United States, the Barnett Shale in and around Fort Worth, TX, changed industry notions on the viability of natural gas shale.

In the early 1980s Mitchell Energy drilled the first well in the Barnett Shale (Yergin (2011)). However, rather than encountering the typical, highly porous, rock of conventional formations, Mitchell encountered natural gas shale. Shale has the potential to hold vast amounts of gas, however, it is highly non-porous which causes the gas to be trapped in the rock. Over a period of 20 years Mitchell Energy experimented with different techniques, and found that by using hydraulic fracturing (commonly referred to as "fracking") it was able to break apart the rock to free natural gas. With higher natural gas prices and the combination of horizontal drilling with "fracking" in 2002, large new reserves from shale became economically profitable to produce. Continued development of drilling and hydraulic fracturing techniques have enabled even more production efficiencies, and today shale wells have an extremely low risk of being unproductive (unproductive wells are commonly referred to as "dry-holes").

The low risk of dry-holes and high production rates have led to a land grab for mineral leases which were previously passed over. Prior to initiating drilling activities a firm must first negotiate with a mineral owner to lease the right to develop minerals. Typically these contracts are comprised of a large upfront "bonus" payment, which is paid whether the well is productive or not, and a royalty percentage based on the value of the gas produced over time. Across the U.S., communities have experienced significant fast-paced mineral booms. For example, the New Orleans' *Times-Picayune (2008)* reports the rise of bonus payments in the Haynesville Shale, which increased from a few hundred dollars an acre to $10,000 to $30,000 an acre plus 25% royalty in a matter of a year. An individual who owns one square mile of land (640 acres) and leases out his minerals at $30,000/acre would receive an upfront one-time payment of $19.2 million plus a monthly payment equal to 25% of the value of all the gas produced on his lease. The media has dubbed those lucky enough to have been sitting on shale mineral leases as "shalionaires." The significant personal windfalls people have experienced in natural gas shale booms has led to increases in bank deposits in the communities that they live in. Since the first major shale boom in the Barnett (TX), additional booms have occurred in the Woodford (OK), Fayetteville (AR), Haynesville (LA + TX), Marcellus (PA + WV), Bakken (Oil ND), and Eagle Ford (TX).

3.2 Identification Strategy

The booms experienced by communities across the U.S. due to shale discoveries are exogenous to the underlying characteristics of the affected communities (health, education, demographics etc). The exogenous factors driving shale development include technological breakthroughs (horizontal drilling/hydraulic fracturing) and larger macroeconomic forces (demand for natural gas and natural gas prices). Acknowledging the unexpected nature of shale gas development John Watson, CEO of Chevron, stated in a *Wall Street Journal (2011)* interview, that the technological advances associated with "fracking" took the industry "by surprise." The development of shale discoveries is typically undertaken by large publicly traded exploration and production companies that obtain financing from financial markets outside of the local area of the discovery. The exogenous nature of a shale boom and the effect it has on local deposit supply creates an attractive setting for a natural experiment, which I use to identify the importance of local credit supply and local banking market structure.

Figure 1 depicts an example of how economic outcomes, measured as establishment levels, change over time for high external finance dependent and low external finance dependent industry groups in a boom county (Johnson County, TX), relative to deposits and drilling activity.

To track shale development I use a unique data set which has detailed information on the time and place (county-year) of drilling activity associated with shale booms. For example, in Johnson County, TX (Figure 1) the number of shale wells, which are used to develop shale natural gas and oil,[5] grew from 0 to 2,336 between 2002 and 2009 . As can be seen in Figure 1, drilling activity began in 2003, but significant activity did not occur until 2004 and 2005. After this date, bank deposits grew from 10% above 2000 levels to 64% above 2000 levels. The influence of these increased deposits can be seen in the disproportionate increase in the level of establishments with high dependence on external finance. Specifically, after the onset of the boom, the number of establishments in high external finance dependent industries grew from 7% above 2000 levels to 29%, while the number of establishments in low external finance dependent industries grew from 5% above 2000 levels to 9% above 2000 levels. This study will provide statistical evidence that the basic result shown in Figure 1 holds across all boom counties.

3.2.1 Effect of Boom on Deposits

The first step in my analysis is to quantify the deposit shock observed in Figure 1 for the entire sample. Specifically what is the impact of a shale boom on local deposit supply? In order to do this I estimate the following regression model

$$Log\ Deposit_{i,t} = \alpha + \beta_1 Log\ Pop_{i,t} + \beta_2 Boom_{i,t} + Year\ FE_t + County\ FE_i + \varepsilon_{i,t}$$

$Boom_{i,t}$ is a measure of shale activity, in my tests I use both logarithm of total shale wells, and a binary dummy boom variable to measure the shale boom. $Log\,Deposit_{i,t}$ is the logarithm of deposits summed across all branches in county i at time t. $Log\,Population_{i,t}$ is included as a control and is the logarithm of the population of county i at time t. County fixed effects are included to control for time invariant county effects and year effects are included to account for time-varying effects, these enter the specification in the form of $Year\ FE_t$ (year fixed effect) and $County\,FE_i$ (county fixed effect). The key variable of interest in this specification

[5] I use horizontal wells as my key measure of shale development activity. Horizontal drilling is a component of the key technological breakthrough that enables the production of shale resources to be economically profitable. Nearly all horizontal wells in the U.S. are drilled to develop shale or other unconventional oil and gas resources.

is the coefficient β_2, which indicates the change in $Log\,Deposit_{i,t}$ attributable to the $Boom_{i,t}$ variable.

A primary concern in my empirical setting may be whether counties with different bank size characteristics experience similar shocks. If a deposit shock were correlated with the underlying banking structure in a county it could suggest problems for my broader empirical tests. To test whether counties with different banking characteristics are affected differently by the deposit shock, I estimate the following regression:

$$Log\ Deposit_{i,t} = \alpha + \beta_1 Log\ Pop_{i,t} + \beta_2 Boom_{i,t} + \beta_3 Small\ Bank_{i,t}$$
$$+\beta_4 Small\ Bank_{i,t} * Boom_{i,t} + Year\ FE_t + County\ FE_i + \varepsilon_{i,t}$$

The key coefficient of interest in measuring whether counties with different bank size characteristics experience different deposit shocks is the interaction coefficient (β_4).

3.2.2 Effect of Credit Supply on Economic Outcomes: Differences-in-Differences

To identify the economic outcomes related to the local credit supply shock, I use a regression specification which distinguishes between economic outcomes for industries with high external financing needs relative to those with low external financing needs. To achieve this aim, I use a regression form of differences-in-differences, where the first difference (β_2) can be thought of as the difference in economic outcomes between boom county-years and non-boom county-years. To identify the effect of the credit component of a boom I incorporate a second difference (β_4), the difference in economic outcomes for industries with high dependence on external finance and industries with low dependence on external finance.

$$Log\ Establishments_{i,j,t} = \alpha + \beta_1 Log\ Pop_{i,t} + \beta_2 Boom_{i,t} + \beta_3 High_j + \beta_4 Boom_{i,t} * High_j$$
$$+IndustryTrends\ FE_{j,t} + CountyIndustry\ FE_{i,j} + \varepsilon_{i,j,t}$$

Where $LogEstablishment_{i,j,t}$ is the number of establishments in county i and industry group j at time t. Due to the low number of establishments in different industries at the county level, I have grouped establishments into two industry types: one industry group which has a high dependence on external finance, for which $High_j = 1$ and one industry group with low dependence on external finance $High_j = 0$.[6] Thus, for every county I have only two industry groups, which are delineated by dependence on external finance. I also include two sets of fixed effects. $IndustryTrends\ FE_{j,t}$ control for time-varying differences in industry growth, while $CountyIndustry\ FE_{i,j}$ control for county specific differences in industry make-up.[7]

This specification is a regression form of differences-in-differences, with the key variable of interest being the coefficient on the interaction term, β_4. If industries with a high dependence on external finance benefit more from shale booms, β_4 would be positive, which would indicate the importance of the credit supply component of a boom. Alternatively, if local credit supply does not influence local economic outcomes, β_4 would be zero. That is, while the boom may benefit all industries through the coefficient β_2 (overall increased demand for goods and services), there would be no evidence that the credit supply component of a boom enhances local economic outcomes.

3.2.3 Effect of Bank Size and Credit Supply on Economic Outcomes: Differences-in-Differences-in-Differences

To estimate the importance of local bank size for local credit supply I use a triple differencing specification. The first two differences are: non-boom county-years vs. boom county-years, high dependence on external finance vs. low dependence on external finance. The third difference tests whether the effect from the first two differences is bigger in areas dominated by small banks: high small bank market share vs. low small bank market share. $SmallBank_{i,t}$ is a variable representing small bank market share in county i at time t. To measure small bank market share, $Small\ Bank_{i,t}$, I use both the proportion of branches in a county which

[6] $High_j$ is not reported in the regression results because this variable is subsumed by the county-industry fixed effects, $CountyIndustry\ FE_{i,j}$

[7] I document in Appendix B that my main results are similar and statistically significant when using different fixed effects

belong to small banks as well as a dummy variable for the counties which are in the highest quartile of small bank branch market share in any given year. The interaction of $SmallBank_{i,t}$ with the other terms in the specification yields a regression form of differences-in-differences-in-differences. [8]

$$
\begin{gathered}
Log\ Establishments_{i,j,t} = \alpha + \beta_1 Log\ Pop_{i,t} + \beta_2 Boom_{i,t} + \beta_3 High_j + \beta_4 Small\ Bank_{i,t} \\
+\beta_5 Boom_{i,t} * High_j + \beta_6 Boom_{i,t} * Small\ Bank_{i,t} + \beta_7 High_j * Small\ Bank_{i,t} \\
+\beta_8 Boom_{i,t} * Small\ Bank_{i,t} * High_j + IndustryTrends\ FE_{j,t} + CountyIndustry\ FE_{i,j} + \varepsilon_{i,j,t}
\end{gathered}
$$

In this regression the key variable of interest is β_8. If industries with higher dependence on external finance benefit more from a local credit supply shock in counties dominated by small banks this coefficient would be positive.

4 Data and Variable Definition

For my panel data set I include the seven states that have experienced shale booms from 2000 through 2009. These are Arkansas, Louisiana, North Dakota, Oklahoma, Pennsylvania, Texas, and West Virginia. There are 639 counties in these states with at least one bank branch over the sample period. Each of these states have counties that have experienced shale booms, as well as counties which have not, and it is these non-boom county-years which serve as the control group in empirical tests. The data is constructed on an annual frequency and compiled from four different sources:

- Well Data (From Smith International Inc.)
- Deposit and Bank Data (From FDIC Summary of Deposits Reports)

[8] $High_j$ is not reported in the regression results because this variable is subsumed by the county-industry fixed effects, $CountyIndustry\ FE_{i,j}$

- County Level Economic Outcome Data by Industry (Census Bureau, Establishment and Employment Data)
- External Finance Dependence Measures (From Compustat)

4.1 Well Data

Well data is used to calculate the $Boom_{i,t}$ variables in the regressions. The well data is obtained from Smith International Inc. which provides detailed information on the time (year), place (county), and type (horizontal or vertical) of well drilling activity. I use horizontal wells as the key measure of shale development activity, as the majority of horizontal wells in the U.S. drilled after 2002 target shale or other unconventional formations. In order to best measure the influence of shale development activity I focus on two different measures.

- $Boom_{i,t} = Dummy_{i,t}$: A dummy variable set to 1 if county i at time t is in the top quartile of all county-years with shale well activity (total shale wells > 17) in the panel dataset. Once the variable is set to 1, all subsequent years in the panel for the county are set to 1. Based on this definition 88.1% of all shale wells are drilled in boom county-years.
- $Boom_{i,t} = Log\,Total\,Shale\,Wells_{i,t}$: The logarithm of the total number of shale wells drilled in county i from 2003 to time t.

Regressions are based on the total shale wells drilled for the year leading up through March. This corresponds to when the County Business Pattern Data are tabulated. Summary statistics on sample states, counties, and well data are presented in Table 1. Figure 2 presents a map of the intensity and location of shale development activity.

4.2 Deposit and Bank Data

Deposit and bank data are obtained from the Federal Deposit Insurance Corporation (FDIC) Summary of Deposit data, which is reported on June 30 of each year and provides bank data for all FDIC-insured institutions. I use the Summary of Deposit data as opposed to data from the Reports of Condition and Income (Call Reports) because Summary of Deposit data provides deposit data at the branch level, while Call Reports only provide data at the bank level. Additionally, Summary of Deposit data provides detailed information on the geographic location of each branch that a bank has, so I can directly observe the branches in boom counties and the banks they belong to. To obtain county level deposit data I sum deposits across all branches in a county. To calculate small bank market share in a county I calculate the proportion of branches in a county which belong to small banks. I define small banks to be banks with assets below a threshold which could cause a bank to be funding constrained. For the results in this paper I use $500 million (year 2003 dollars) as the asset threshold for small banks.[9] Prior literature (Black and Strahan (2002), Jayaratne and Morgan (2000), Strahan and Weston (1998)), has suggested that banks with assets in the $100 million to $500 million range may be funding constrained. In my empirical tests I use two measures of small bank market share. Specifically, I use dummy variables set to 1 for the counties with the highest small bank branch market share (top quartile) in each year, and 0 otherwise. Additionally, I also use the ratio of small bank branches to total branches in a county. Summary data for bank and branch variables are provided in Table 2.

[9] I document that the main results remain statistically significant when using $200 million or $1 billion in assets as the definition of a small bank. The results are also robust to basing this definition off of bank holding company assets.

4.3 County Level Economic Outcome Data by Industry

Economic outcome variable data by industry was obtained from the County Business Patterns survey, which is released annually by the Census Bureau. It is worth noting, that the survey provides data only on establishments, not firms, for example, a firm may have many establishments. The survey provides detailed data on establishments and employment in each county, by North American Industry Classification System (NAICS) code as of the week of March 12 every year. My main results are based on economic outcomes grouped at the two digit NAICS code level, which I match with corresponding Compustat two digit NAICS code external finance dependence measures. More disaggregated NAICS codes (six digit NAICS as opposed to two digit NAICS) provide fewer NAICS code matches to Compustat, which I rely on for external finance dependence measures. I exclude codes 21 (Oil and Gas Extraction), 23 (Construction), 52 (Financials), 53 (Real Estate) because they may be directly influenced by booms. I exclude 99 (Other) due to lack of comparability with Compustat firms. However, my results remain similar and statistically significant when any of these industries are included.[10]

After matching County Business Pattern data with Compustat external finance dependence measures, I aggregate all industry codes into two industry groups, one with above median dependence on external finance (high) and one with below median dependence on external finance (low). The two digit NAICS code from the County Business Patterns data is used to obtain an external finance dependence measure from Compustat, which is described in more detail in the next subsection. The objective of the matching is to have the cleanest sorting of NAICS codes into high external finance dependence and low external finance dependence bins. Details on the industries in these bins are provided in Table 3.

While the County Business Patterns Survey provides detailed data on establishment counts by industry, employment data may be suppressed, for privacy reasons, if there are too few establishments in a particular industry. Employment data suppression is a particular

[10] Using three digit NAICS code industries poses two problems 1) There are 71 industries as opposed to 14, so there are far fewer comparable Compustat firms for some industries 2) There was a change in industry categorization that occurred in 2002-2003, which creates problems when constructing a pre-boom control period for booms that occur in 2003 and 2004.

problem for counties with smaller populations, for this reason the number of observations in employment regressions is reduced. Furthermore, this suppression of employment data makes including employment in the regressions related to small bank market share problematic, as 62% of establishments in high small bank market share counties have employment reporting suppressed.

4.4 External Finance Dependence Measures

I use an external finance dependence measure similar to the measure used by Rajan and Zingales (1998). The main difference is that while they use this measure only for manufacturing firms, I use it for all industry groups similar to Becker (2007). Specifically, over the 1999 to 2008 time period for each firm in Compustat I sum the difference between capital expenditures and operating cash flow. I use the time period 1999 to 2008 because these fiscal years, which end in December for most public firms, correspond most closely to March of the following year (2000 to 2009), which is when the county business patterns survey is conducted. By summing over several years the measure is less susceptible to being driven by short term economic fluctuations. I then divide this sum by the sum of capital expenditures. Specifically, for firm n, the measure is calculated as:

$$ExtFinDependence_n = \frac{\sum_{1999}^{2008}(CapitalExpenditures_{n,t} - OperatingCashFlow_{n,t})}{\sum_{1999}^{2008} CapitalExpenditures_{n,t}}$$

I take the median of this measure to get an industry's external finance dependence. The calculation of this measure for each industry is displayed in Table 3. The underlying assumption in the Rajan and Zingales (1998) measure is that some industries, for technological reasons, have greater dependence on external financing than others. The measure is based on public firms in the United States which have among the best access to capital of any firms in the world, therefore the amount of capital used by these firms is likely the best estimate that can be obtained of an industry's true demand for external financing.

5 Results

5.1 Effect of Shale Booms on Deposit Levels

Table 4 provides regression results of log deposits on different shale boom variables. The evidence suggests a causal relationship between shale booms and bank deposits, specifically, that the individual mineral wealth generated by shale booms translates into more bank deposits. In Panel A of Table 4 columns (1) and (2) provide results on different measures of the $Boom_{i,t}$ variable. In each case, the $Boom_{i,t}$ variable is found to have both economic and statistical significance. For example, the dummy variable measure of $Boom_{i,t}$ can be interpreted as a boom increasing local deposits by 8.2%. To put this in context, the average annual growth rate in deposits across all counties from 2000 to 2009 was 4.6%, so a boom county would experience an additional increase of 8.2% (4.6% + 8.2% = 12.8% total increase), or a total increase in deposits roughly triple its average annual increase.

Further tests will focus on comparisons between counties with high small bank market share and low small bank market share. An assumption in this comparison is that both types of counties experience similar deposit shocks. To directly test this assumption I estimate interactions of county bank size characteristics interacted with the shale boom variables. Panel B reports the results of this specification. The key coefficient of interest in assessing whether counties experience different shocks based on their banking structure is the coefficient on the interaction term (β_4). This coefficient is neither economically nor statistically significant, suggesting that counties with different banking structure receive similar deposit shocks.

An additional concern may be that deposits could be rising in anticipation of a boom, or that there could be some spurious correlation in a county during part of the boom period which is causing the result in Table 4. To test the precise timing of the boom relative to deposit growth I replace the boom dummy variable used in Table 4 with dummy variables based on the position of an observation relative to a boom. So, for example, if a boom occurs in 2006 in county i, then the observation in county i in 2003 would receive a t-3 boom dummy, county i observation in 2004 would receive the t-2 boom dummy and so on. I include a set of dummies for each year relative to a boom from t-3 to t+3. Due to limited observations beyond t+3, I group any observations after t+3 with the t+3 dummy (3+). Figure 3 is a graph of the coefficients from this regression, and provides visual evidence that the deposit level does not change substantially until time 0, the first year of the boom. This serves to alleviate concerns regarding whether deposits rise in anticipation of a boom, as well as concerns about possible spurious correlations during part of the boom period.

5.2 Effect of Credit Supply Shock on Economic Outcomes

Table 5 provides the results of the effect the shale boom on log establishments for all industries industry types (high external finance dependent and low external finance dependent). The economic interpretation of the results in column (1) is that the establishment levels of all industries increases by 2.2% in a county when there is a boom. However, demand for goods and services for all establishments may increase when there is a boom, so it is not surprising to see the results in Table 5. In order to draw a more direct causal relationship between the credit supply shock associated with a shale boom and economic outcomes, it is necessary to look at the difference between outcomes for industries with a high dependence on external finance compared to those with low dependence on external finance. In Table 6 I estimate the regression specification in Table 5 on each industry group separately. The larger magnitude of coefficients in the regressions for industries more dependent on external finance (Ind = High) suggest that the industries in this industry group benefit more from a boom. Specifically, the coefficients from (1) of Panel A in Table 6 can be interpreted as a 4.5% increase in high external finance dependent establishments when a boom occurs relative to no increase in low external finance dependent establishments.[11]

There may be some concerns as to the timing of the boom and changes in local economic

[11] I document in Appendix A that for banks that have all branches in a single county, both deposits and Commercial & Industrial loans increase after a boom. Overall interest income and interest paid on deposits are unchanged after a boom. Lending driven purely by demand would be more likely to result in higher interest rates and interest income.

outcomes. If establishment levels of low external finance dependent industries and high external finance dependent industries trend differently prior to the boom, they may be poor control/treatment groups. Additionally, if external finance dependent establishments trend higher well before the boom, it would suggest a problem with my empirical design, as the deposit levels in Figure 3 do not increase until time 0. To directly assess the validity of these concerns I construct a graph similar to Figure 3, but for establishments. Specifically, for each of the industry groups I run the regressions in (1) and (2) of Table 6, but replace the $Boom_{i,t}$ variable with a set of dummy variables based on the time period of an observation relative to a boom for any given county i (similar to what is done in Figure 3). The coefficients from this regression are graphed for each industry group in Figure 4. As can be seen, from time t-3 to t-1, each industry group tracks relatively closely, then at time 0, the first year of a boom, there is a divergence in trends, which increases through t+3. This indicates that when the boom occurs, establishments in high external finance dependent industries benefit disproportionately more compared to low external finance dependent industries. The evidence presented in Figure 4 should serve to address concerns regarding the change in establishment levels relative to the precise timing of a boom. The results outlined in Panel A of Table 6 and Figure 4 can be formalized in a regression form of differences-in-differences.

Panel B of Table 6 provides a direct test of the evidence presented in Panel A of Table 6 and Figure 4 in a regression form of differences-in-differences. The coefficient of interest for assessing whether improved local credit supply plays a role in local economic outcomes is the interaction term $Boom_{i,t} * High_j$. The sign and magnitude of this term indicates whether one industry group benefits disproportionately when there is a boom. The coefficient on the interaction term is positive and statistically significant in all specifications, suggesting that local economic outcomes for industries more dependent on external finance benefit more than outcomes for industries with low dependence on external finance. The economic interpretation of the interaction coefficient in (1) of Panel B of Table 6 is that, when there is a boom, establishments in industries with high dependence on external finance increase 4.6% relative to establishments in industries with low dependence on external finance. To put this number in context, the average annual increase in high external finance dependent establishments from 2000 to 2009 is 0.9%.

5.3 Effect of Bank Size and Credit Supply on Economic Outcomes

As previously discussed, local bank size composition could play a role in the importance of improved local credit supply for economic outcomes. Specifically, counties dominated by small banks may benefit more from a credit supply shock. To test this in the differences-in-differences framework, I subdivide counties into high small bank market share and low small bank market share counties, based on whether a county is in the top quartile of small bank market share in a given year. I estimate the specification presented in Panel B of Table 6 for each of these subgroups, and present the results in Table 7.

In every specification the counties dominated by small banks have a higher coefficient for the interaction term $Boom_{i,t} * High_j$. The magnitude of differences is often quite large, with high small bank market share counties (Bank = High Small Bank Mkt Share) having coefficients more than five times the coefficients of low small bank market share counties (Bank = Low Small Bank Mkt Share), depending on the specification. In order to address concerns regarding anticipation and spurious correlations, I graph coefficients as in Figure 4, but further subdivide high external finance and low external finance industries by bank size characteristics to form four separate subgroups in Figure 5. As can be seen, all subgroups trend similarly until time 0, when the subgroup that comprises high external finance dependent industries in high small bank market share counties trends higher.

To formally test the results in Table 7 and Figure 5, I estimate a regression form of differences-in-differences-in-differences, with the results shown in Table 8. This is done by adding additional interactions with small bank market share variables. The coefficient of interest in these tests is the triple interaction term $Boom_{i,t} * High_j * Small\ Bank_{i,t}$. A positive coefficient on the triple interaction term indicates that high external finance dependent industries benefit more when there is a boom in an area with high small bank market share. Specifically, the interpretation of (1) in Table 8 is that high external finance dependent establishments increase by 10.9% relative to establishments in industries with low dependence on external finance in boom counties dominated by small banks relative to the difference between these industry groups in other boom counties. Across all specifications the coefficient on $Boom_{i,t} * High_j * Small\ Bank_{i,t}$ is positive and statistically significant, providing formal evidence that higher small bank market share counties were funding constrained prior to the boom. Specifically, if there were no frictions, additional deposits from the boom should not disproportionately benefit high external finance dependent industries in high small bank market share counties.

The results in Table 8 also address concerns regarding alternative explanations from the prior differences-in-differences tests conducted. An important concern is whether high external finance dependent industries disproportionately benefit from a boom for a reason other than the credit supply component of a boom. For example, it could be the case that high external finance dependent industries benefit more in general when there is an economic boom (high asset beta). However, this explanation would not account for the differential impact experienced in high small bank market share counties relative to low small bank market share counties. An additional concern may be that there could be more demand for goods and services for industries in the high external finance dependence industry group. However, in order for this explanation to be consistent with the results in Table 8, there would also need to be a rationale for why this demand differential is relatively higher in counties with high small bank market share. In summary, the results in Table 8 provide added robustness for the initial differences-in-differences tests, while also documenting that frictions are more problematic in counties dominated by small banks.

5.3.1 Alternative Measures of Economic Outcomes

The results presented in the main tests use logarithm of establishments as the primary economic outcome measure. Table 9 reports the results of specifications that use the logarithm of employment, establishments per capita, and logarithm of establishments per capita as economic outcome measures. The results for these alternative measures are consistent with the results reported in Table 6 and Table 8. One primary drawback of using employment data as an economic outcome measure is that in many areas it is suppressed for privacy reasons.

5.4 Robustness

5.4.1 Sensitivity of Results to Industry Classifications

A potential concern with my empirical design is whether local economic outcomes for industries more dependent on external finance improve relative to outcomes for industries less dependent on external finance for some reason other than improved local credit supply. The differences-in-differences-in-differences tests help rule out several alternative explanations, however, an additional test of this assumption is included in Table 10. Specifically, for each industry group I calculate a measure of exposure to underlying economic fluctuations, asset beta, using two different asset beta methodologies.

$$\beta_{Asset1} = \frac{\beta_{Equity}}{1 + (1 - Tax\ Rate) * \frac{Debt}{Equity}}$$

$$\beta_{Asset2} = \frac{\beta_{Equity}}{1 + \frac{Debt}{Equity}}$$

The asset betas used are industry median asset betas. If it is the case that the asset betas for each industry group are different it could be cause for concern, as this would suggest that one industry group would be more sensitive to overall fluctuations in an economy. The results in Panel A of Table 10 provide evidence that the high external finance dependent industry group does have a higher asset beta. However, when the two highest asset beta industry groups are dropped from the regressions causing both industry groups to have similar asset betas, as in Panel B of Table 10, the interaction and triple interaction coefficients from the differences-in-differences regression and differences-in-differences-in-differences regression are still positive and statistically significant. This suggests that the difference in underlying asset betas between the groups is not driving my main results. Additionally Table 10 provides evidence that the regression results presented in Table 6 and Table 8 are not being driven by any single industry group, and either its inclusion (Panel A (3) and (4)) or exclusion (Panel C (3) and (4)) in the study.

5.4.2 Falsification Tests

A potential concern in differences-in-differences tests is whether results are driven by pre-existing trends or are otherwise anticipated. To directly test whether any of the local economic outcome changes begin prior to a boom, I include dummy variables for the two years prior to the first shale development. These enter the regressions in the form of the $False\,Boom_{i,t}$ variable. As can be seen in the results in Table 11, neither the $False\,Boom_{i,t}$ variable, nor any of the interaction variables are statistically significant. This result provides direct evidence that the changes in economic outcome variables documented in this paper do not occur prior to the onset of shale development activity, and that there are no statistically significant pre-existing trends. Furthermore, because shale discoveries occur in different years in different counties (not just a single event in all counties at the same time), alternative interpretations of results would need to address changes in economic outcomes that happen to coincide with boom events in different locations at different points in time.

I conduct a second falsification test to assess whether growth shocks in general favor one industry group over another. Specifically, in Table 12 I use data from the states immediately adjacent to the seven shale states to test whether generic non-shale growth shocks or "booms" benefit one set of industries or counties dominated by small banks. $Growth\,Shock_{i,t}$ dummy variables are inserted after high growth county-years so that the number of false boom county years is approximately the same proportion of shale boom county years obtained in the main sample (5% of all county-years). The key coefficient of interest to test whether high external finance dependent industries always benefit when there is growth in an area is on the interaction term $Growth\,Shock_{i,t} * High_j$, this coefficient is not statistically significant. Additionally, the triple interaction term $Growth\,Shock_{i,t} * High_j * Small\,Bank_{i,t}$ is neither positive nor statistically significant. These results suggest that the credit component of shale booms make shale growth shocks unique from general localized growth shocks.

5.4.3 Are Demand Shocks from Shale Booms Correlated with Bank Size?

A potential concern for the validity of my empirical design is whether real shocks associated with a shale boom are larger in counties dominated by small banks relative to other counties. If this is the case, my interpretation of my empirical tests may be problematic. To test this concern, I use retail sales data from the Economic Census conducted by the U.S. Census Bureau every 5 years. For this test, I use data on retail sales to proxy for demand in an area. The specific comparison I make is based on the 2002 and 2007 Economic Census data. Using this data I can test whether retail sales increase more in counties dominated by small banks after a boom relative to other counties after a boom. The key coefficient of interest in this test, is the interaction term $Boom_{i,t} * SmallBank_{i,t}$. If this coefficient is different from 0, it would suggest that retail sales increase more in a county with a particular type of bank structure, and therefore indicate that demand shocks may be different across different counties. As can be seen in the specifications in Table 13, the coefficients on the interaction term $Boom_{i,t} * SmallBank_{i,t}$ are not statistically different from 0, suggesting that demand shocks are not correlated with bank size.

5.4.4 Bank Size, Bank Holding Companies, Post-Boom Banking Market Changes

The main results in this study categorize any bank with fewer than \$500 million in assets as a small bank. However, existing literature has sometimes used different small bank definitions. Additionally, banks that are part of larger bank holding companies may have fewer funding constraints than banks that are not (Houston et al. (1997)). Table 14 reports regression results that use different small bank asset thresholds: \$200 million, \$500 million, \$1 billion (DeYoung et al. (2004)). Coefficients on the key interaction term of interest $Boom_{i,t} * High_j * Small\ Bank_{i,t}$ are positive and statistically significant, indicating that the primary results reported in this paper are robust to alternate definitions of small bank. Additionally, categorizing banks based off of bank holding company assets does not alter the main results.

There could be some concern that a county's bank size composition endogenously changes after a shale discovery. To address this concern I estimate regressions that hold banking market structure constant as of last year prior to a boom. These results are reported in Table 15. The coefficients on the triple interaction term are similar in magnitude to the main specifications in Table 8 and remain statistically significant. While bank structure could be endogenous in a given year, it is unlikely that it is changed due to the anticipation of a boom. Therefore local bank structure is not correlated with whether a county is treated (experiences a boom) or not. Furthermore, the results from Table 15 indicate that if bank structure is changing after a boom, it does not alter the main results significantly.

6 Conclusions

The United States has one of the most developed banking systems in the world. Prior research has demonstrated that deregulation, the adoption of lending technology and securitization, have led to improved economic outcomes. However, this paper provides new evidence that, despite improvements, economically significant frictions still remain in the U.S. banking system. I use oil and gas shale discoveries to obtain exogenous variation in local credit supply to document the economic magnitudes of these frictions. When there is a positive local credit supply shock, economic outcomes for industries with more dependence on external finance improve relative to industries with less dependence on external finance, suggesting that local credit supply matters for economic outcomes.

The importance of local credit supply is linked to local bank size. Consistent with the view that either small banks are funding constrained or are in areas with more "soft" information borrowers, counties dominated by small banks experience a fivefold higher benefit from a local credit supply shock. These findings suggest that deregulation, increased use of lending technology and securitization have not fully alleviated economically important frictions, particularly in areas dominated by small banks.

References

Ashcraft, A. B., 2005. Are banks really special? New evidence from the fdic-induced failure of healthy banks. American Economic Review 95, 1712–1730.

Bassett, B., Brady, T., 2002. What drives the persistent competitiveness of small banks? Finance and Economics Discussion Series working paper 2002-28 Board of Governors of the Federal Reserve System.

Becker, B., 2007. Geographical segmentation of US capital markets. Journal of Financial Economics 85, 151–178.

Berger, A. N., 2003. The economic effects of technological progress: Evidence from the banking industry. Journal of Money, Credit and Banking 35, 141–176.

Berger, A. N., Miller, N., Petersen, M. A., Rajan, R. G., Stein, J. C., 2005. Does function follow organizational form? Evidence from the lending practices of large and small banks. Journal of Financial Economics 76, 237–269.

Berger, A. N., Rosen, R. J., Udell, G. F., 2007. Does market size structure affect competition? The case of small business lending. Journal of Banking and Finance 31, 11–33.

Berger, A. N., Udell, G. F., 2006. A more complete conceptual framework for SME finance. Journal of Banking and Finance 30, 2945–2966.

Bertrand, M., Schoar, A., Thesmar, D., 2007. Banking deregulation and industry structure: Evidence from the french banking reforms of 1985. Journal of Finance 62, 597–628.

Black, S. E., Strahan, P., 2002. Entrepreneurship and bank credit availability. Journal of Finance 57, 2807–2833.

Butler, A. W., Cornaggia, J., 2011. Does access to external finance improve productivity? Evidence from a natural experiment. Journal of Financial Economics 99, 184–203.

Campello, M., 2002. Internal capital markets in financial conglomerates: Evidence from small bank responses to monetary policy. Journal of Finance 57, 2773–2805.

Cetorelli, N., Strahan, P., 2006. Finance as a barrier to entry: Bank competition and industry structure in local U.S. markets. Journal of Finance 61, 437–461.

Chava, S., Purnanandam, A., 2011. The effect of banking crisis on bank-dependent borrowers. Journal of Financial Economics 99, 116–135.

Cornaggia, J., 2012. Does risk management matter? Evidence from the U.S. agriculture industry. Journal of Financial Economics forthcoming.

DeYoung, R., Frame, W., D.Glennon, Nigro, P., 2011. The information revolution and small business lending: The missing evidence. Journal of Financial Services Research 39, 19–33.

DeYoung, R., Hunter, W. C., Udell, G. F., 2004. The past, present, and probable future for community banks. Journal of Financial Services Research 25, 85–133.

Guiso, L., Sapienza, P., Zingales, L., 2004. Does local financial development matter? Quarterly Journal of Economics 119, 929–969.

Houston, J., James, C., Marcus, D., 1997. Capital market frictions and the role of internal capital markets in banking. Journal of Financial Economics 46, 135–164.

Iyer, R., Peydro, J., 2011. Interbank contagion at work: Evidence from a natural experiment. Review of Financial Studies 24, 1337–1377.

Jayaratne, J., Morgan, D., 2000. Capital market frictions and deposit constraints at banks. Journal of Money, Credit and Banking 32, 74–92.

Jayaratne, J., Strahan, P., 1996. The finance-growth nexus: Evidence from bank branch deregulation. Quarterly Journal of Economics 111, 639–670.

Kashyap, A., Stein, J. C., 2000. What do a million observations on banks say about the transmission of monetary policy. American Economic Review 90, 407–428.

Kerr, W. R., Nanda, R., 2009. Democratizing entry: Banking deregulations, financing constraints and entrepreneurship. Journal of Financial Economics 94, 124–149.

Khwaja, A. I., Mian, A., 2008. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. American Economic Review 98, 1413–1442.

Loutskina, E., Strahan, P. E., 2009. Securitization and the declining impact of bank finance on loan supply: Evidence from mortgage originations. Journal of Finance 64, 861–889.

Paravisini, D., 2008. Local bank financial constraints and firm access to external finance. Journal of Finance 63, 2161–2193.

Peek, J., Rosengren, E., 2000. Collateral damage: Effects of the Japanese bank crisis on real activity in the United States. American Economic Review 90, 30–45.

Petersen, M. A., Rajan, R. G., 2002. Does distance still matter? The information revolution in small business lending. Journal of Finance 57, 2533–2570.

Plosser, M., 2011. Bank heterogeneity and capital allocation: Evidence from 'fracking' shocks. Working Paper.

Rajan, R. G., Zingales, L., 1998. Financial dependence and growth. American Economic Review 88, 559–586.

Schnabl, P., 2011. The international transmission of bank liquidity shocks: Evidence from an emerging market. Journal of Finance forthcoming.

Stein, J. C., 2002. Information production and capital allocation: Decentralized versus hierarchical firms. Journal of Finance 57, 1891–1921.

Strahan, P. E., Weston, J. P., 1998. Small business lending and the changing structure of the banking industry. Journal of Banking and Finance 22, 821–845.

Times-Picayune, 2008. 'SWEET SPOT; A recent rush on natural gas drilling in northwest louisiana is turning many landowners into instant millionaires, and stoking others' hopes September 17.

Wall Street Journal, 2011. Oil without apologies April 15.

Yergin, D., 2011. The Quest: Energy, Security, and the Remaking of the Modern World. The Penguin Press.

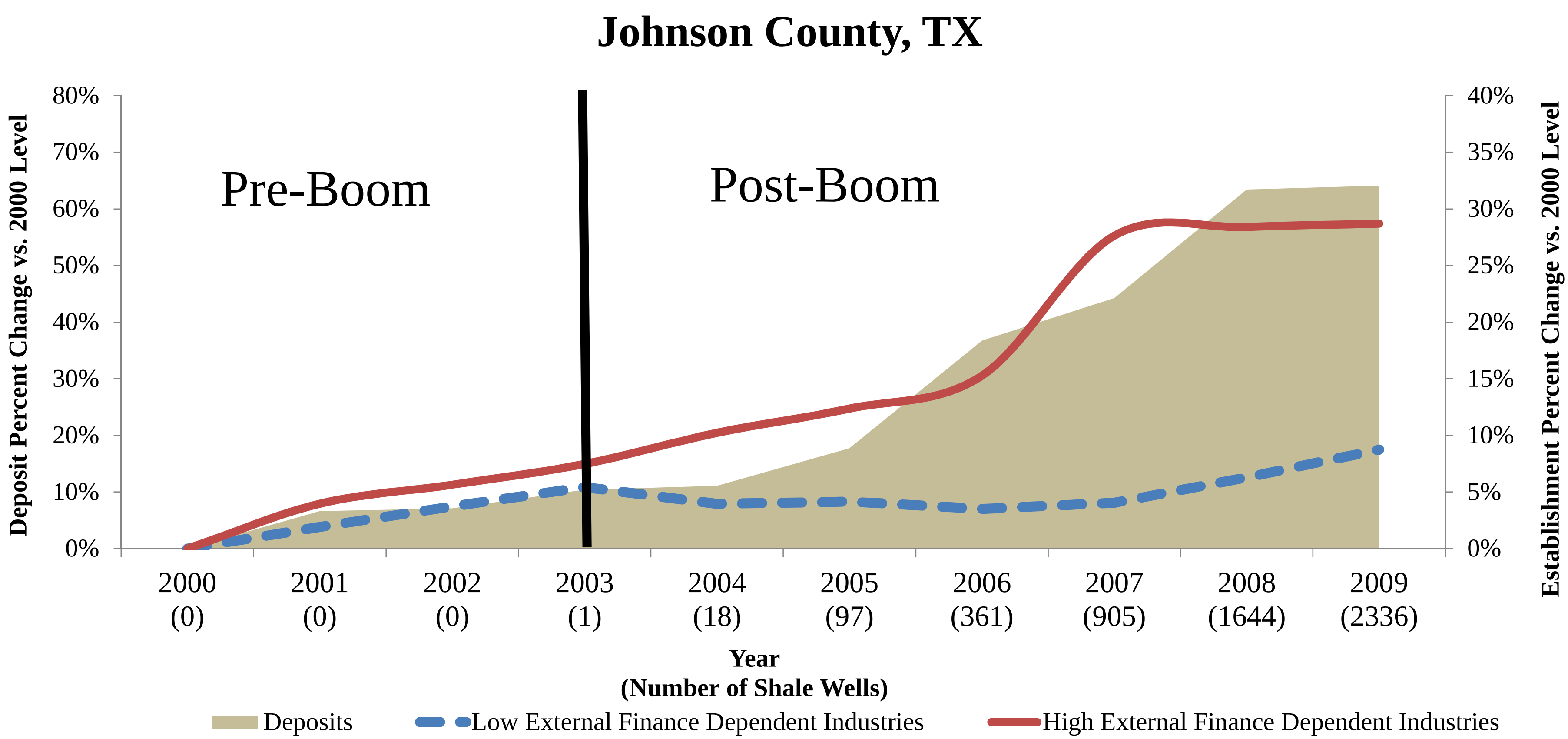

Figure 1: Boom County, Johnson County, TX: This figure plots the relative change in deposits levels and establishment levels in Johnson County, TX. Establishments are divided into two industry groups, high external finance dependent industries and low external finance dependent industries. The numbers in parenthesis under the years on the x-axis are the total number of shale wells drilled in the county by that point in time.

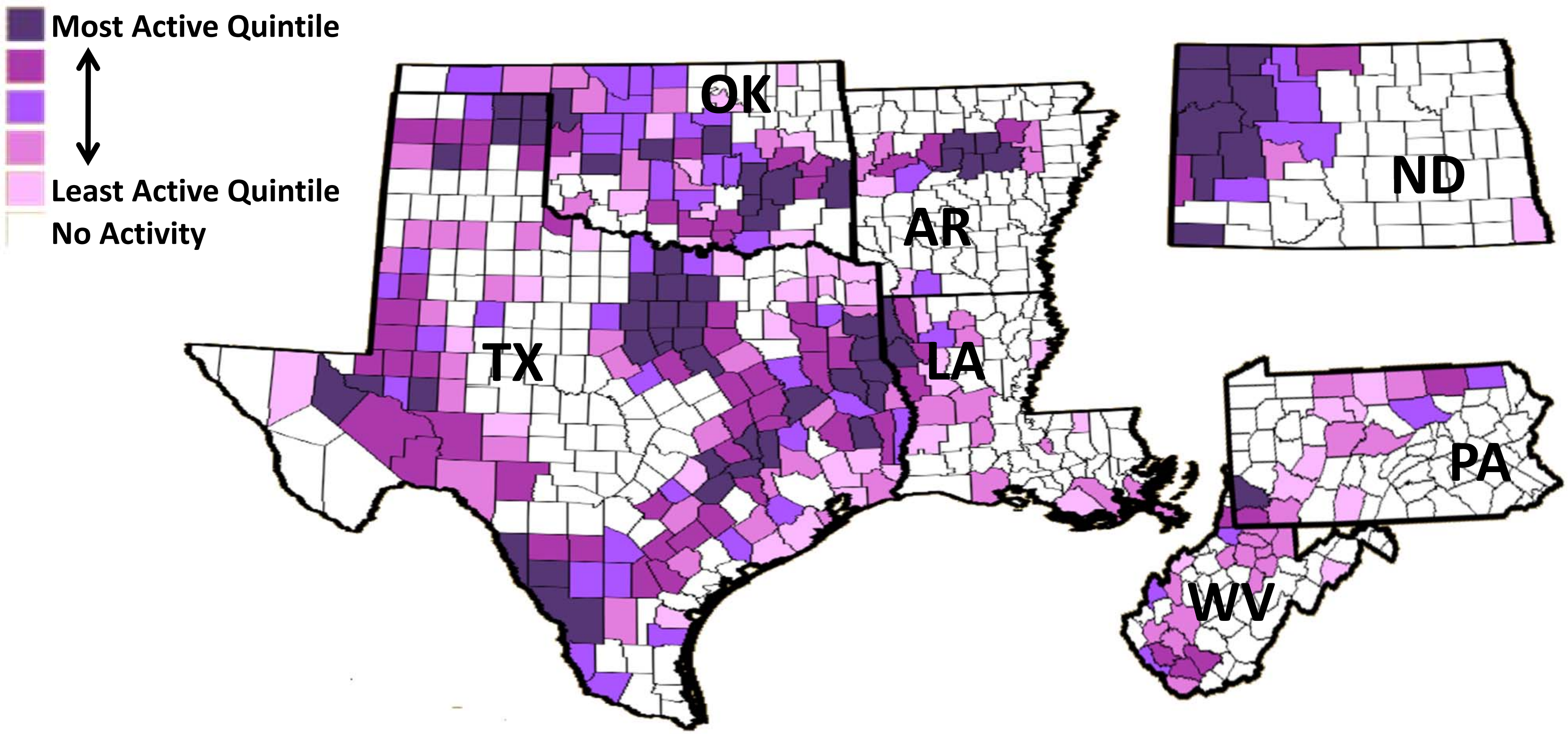

Figure 2: Location and Intensity of Shale Activity
The figure maps the counties of the 7 shale boom states included in this study: OK, TX, LA, WV, PA, ND and AR. White counties are counties with no shale development activity. The remaining counties are shaded based on intensity of activity related to the total number of shale wells drilled through 2009.

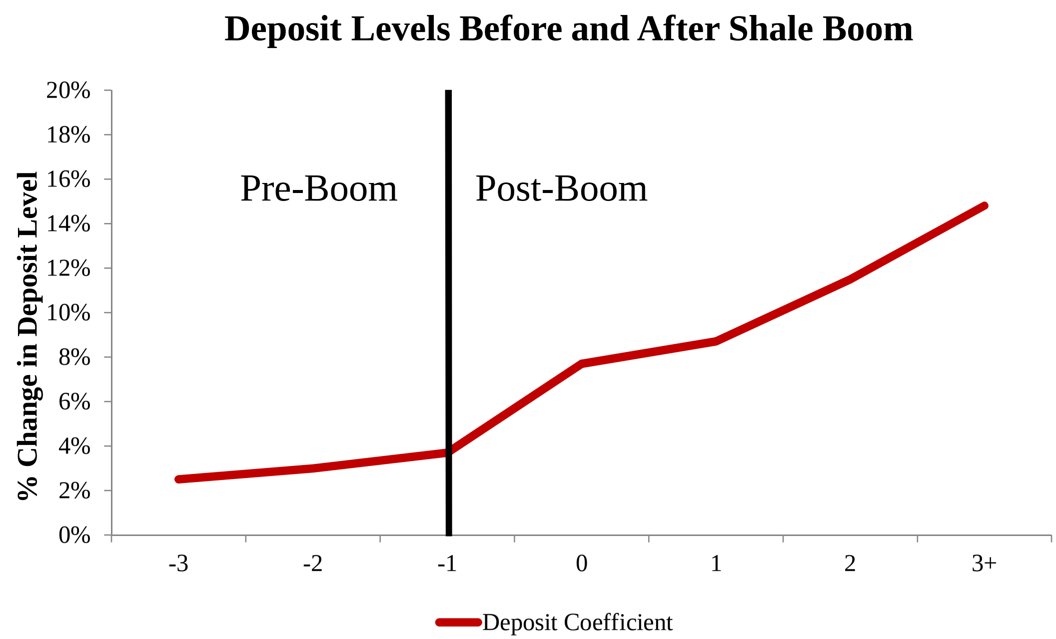

Figure 3: Deposit Levels Before and After Shale Boom

This figure plots the regression of dummy variables based on the year relative to a boom. The first year of a boom is year 0, and the definition of boom that is used is Boom Dummy (previously defined). For example, the first point is the plot of a dummy variable for time t-3 relative to the boom. Due to limited observations for times greater than t+3, all observations after time t+3 are grouped with the t+3 dummy (3+). The dependent variable is the logarithm of total deposits in the county, so the coefficients can be interpreted as the percentage change in the level of deposits at different points in time relative to the boom. The logarithm of population, year fixed effects, and county fixed effects were included in the regression as well.

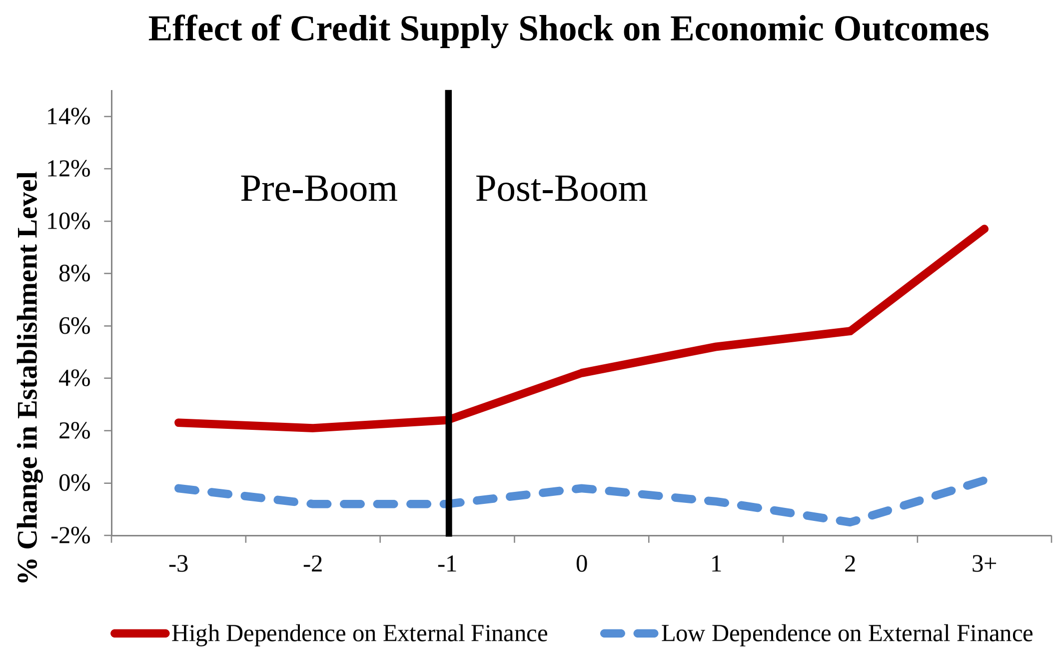

Figure 4: Establishment Levels Before and After Credit Supply Shock

This figure plots separately the regression coefficients of dummy variables of the year relative to a boom for industries with high dependence on external finance and low dependence on external finance. The first year of a boom is year 0, and the definition of boom that is used is Boom Dummy (previously defined). For example, the first point is the plot of a dummy variable for time t-3 relative to the boom. Due to limited observations for times greater than t+3, all observations after time t+3 are grouped with the t+3 dummy (3+). The dependent variable is logarithm of establishments in an industry in a county, so the coefficients can be interpreted as the percentage change in establishment levels at different points in time relative to the boom. The logarithm of population, year fixed effects, and county fixed effects were included in the regression as well.

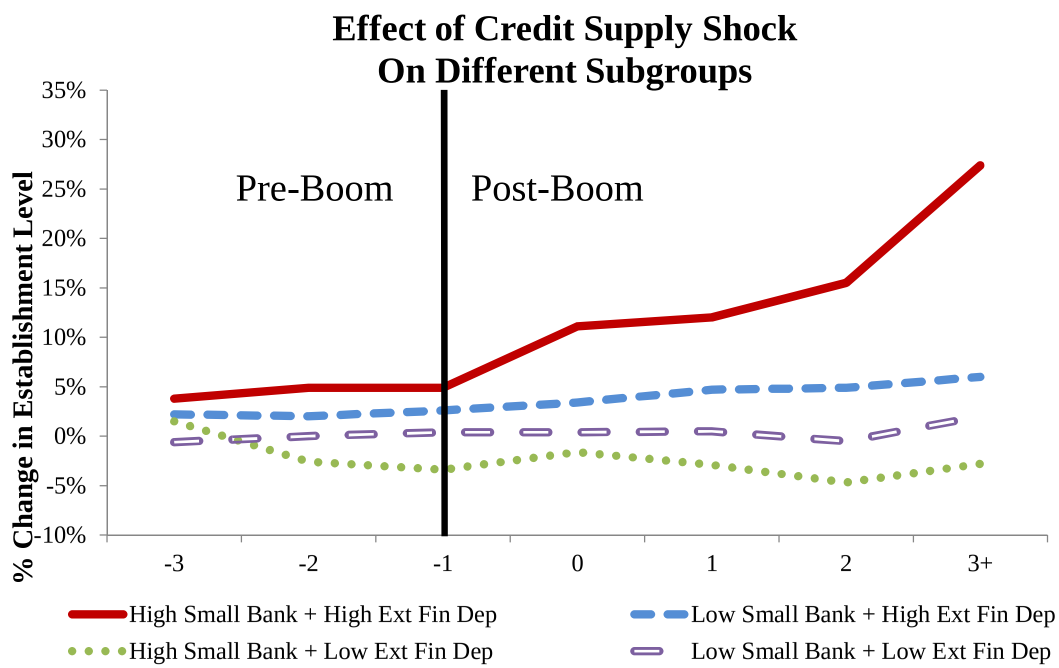

Figure 5: Effect of Credit Supply Shock on Counties with Different Bank Sizes

This figure plots separately the regression coefficients of dummy variables of the year relative to a boom for different subgroups. Specifically four different group designations are used based on whether an establishment has high or low dependence on external finance and whether it is in a county with high or low small bank market share. The first year of a boom is year 0, and the definition of boom that is used is Boom Dummy (previously defined). For example, the first point is the plot of a dummy variable for time t-3 relative to the boom. Due to limited observations for times greater than t+3, all observations after time t+3 are grouped with the t+3 dummy (3+). The dependent variable is logarithm of establishments in an industry in a county, so the coefficients can be interpreted as the percentage change in establishment levels at different points in time relative to the boom. The logarithm of population, year fixed effects, and county fixed effects were included in the regression as well.

Table 1: Summary Statistics of States and Counties With Shale Booms

This table contains summary statistics for the well data used in this study. Development of shale and other unconventional formations is done using horizontal drilling, so I use horizontal well activity as the primary method of measuring when and where booms occur. The states in the sample are states situated in the primary shale development areas: Barnett (TX), Woodford (OK), Haynesville (LA + TX), Fayetteville (AR), Marcellus (PA + WV), Eagle Ford (TX), Bakken (ND). Well data was obtained from Smith International Inc.

Panel A: States, Counties, Shale Well Activity

Number of States	7
Number of Counties	639
Number of Boom Counties	104
Total Number of Shale Wells	16,731
Time Period	2000 - 2009

Panel B: Shale Discoveries ("Booms") Over Time

Year	Discoveries	Total Discoveries as of Date
2003	6	6
2004	11	17
2005	14	31
2006	11	42
2007	22	64
2008	22	86
2009	18	104
Total		**104**

Table 2: Panel Regression Summary Statistics

This table contains summary statistics for the data used in the panel regressions. The unit of observation for economic outcome variables in the panel is at the county-year-industry (external finance industry group) level, while the unit of observation for bank deposits is at the county-year level. Data on establishments and employment are from the County Business Patterns survey. Economic outcome variables are summed across all industries into two groups based on an industry's dependence on external finance. Hence for each county-year there are two industry groups, one with high dependence on external finance and one with low dependence on external finance. Data on annual population levels are from the Census Bureau. Small Banks are categorized as banks with less than $500 million in assets, adjusted for inflation (year 2003 dollars). Bank data was compiled from the FDIC Summary of Deposit reports. Shale well information is based on well data obtained from Smith International Inc.

	Obs	Mean	Std Dev
Deposits			
Log Deposits	6,382	12.60	1.41
Deposits ($ in thousands)	6,382	1,195,253	5,228,131
Economic Outcomes			
Log Establishments	12,764	5.24	1.44
Establishments	12,764	680	2,193
Control/Explanatory Variables			
Log Total Shale Wells	12,764	0.45	1.07
Small Bank Branch Market Share	12,764	0.63	0.30
Log Population	12,764	10.08	1.40
Population	12,764	75,938	229,054

Table 3: External Finance Dependence of Industries

This table reports the industry groups used in this study. The industry groups are based on the two digit North American Industry Classification System used in the reporting of the County Business Patterns survey, which is reported annually by the Census Bureau. For each industry a measure of dependence on external finance is calculated, based on the method used by Rajan and Zingales (1998). The external dependence measure reported for each industry is the industry median dependence on external finance. The data used to calculate the external finance dependence measure is from Compustat for the period from 1999 to 2008 (the fiscal years that are closest to the March data collection of the County Business Patterns survey from 2000 to 2009). The economic outcome measures used are aggregated into two separate industry groups in each county, one with above median dependence on external finance (External Dependence Flag = 1), and one with below median dependence on external finance (External Dependence Flag = 0).

Two Digit NAICS	Two Digit NAICS Name	External Dependence Measure	External Dependence Flag
62	Health Care and Social Assistance	-0.87	0
42	Wholesale Trade	-0.79	0
11	Agriculture, Forestry, Fishing and Hunting	-0.43	0
61	Educational Services	-0.43	0
81	Other Services (except Public Administration)	-0.29	0
44	Retail Trade	-0.20	0
22	Utilities	-0.14	0
56	Administrative and Support and Waste Management and Remediation Services	-0.01	1
48	Transportation and Warehousing	-0.01	1
31	Manufacturing	0.16	1
72	Accommodation and Food Services	0.18	1
71	Arts, Entertainment, and Recreation	0.45	1
54	Professional, Scientific, and Technical Services	0.81	1
51	Information	0.95	1

Table 4: Effect of Shale Booms on Bank Deposits

This table reports the results of regressions which measure the effect of different boom variables on deposits. The dependent variable in these regressions is the log of total deposits in county i in year t. The explanatory variables are different shale boom variables, which have previously been defined. Log of county population is included as a control, as are county and year fixed effects. Panel A documents the effect of a shale boom on county bank deposits, while Panel B tests whether there is a differential effect on bank deposits based on local bank size. The Small Bank Dummy variable used in Panel B is equal to 1 if a county is in the top quartile of small bank branch market share in a given year, and 0 otherwise. The definition of small bank in these regressions is any bank with less than $500 million in assets adjusted for inflation. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

$$\textit{Log Deposits}_{i,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \textit{Year FE}_t + \textit{County FE}_i + \varepsilon_{i,t}$$

Panel A: Effect of Shale Booms on Bank Deposits

	Boom = Dummy	Boom = Log Total Shale Wells
	(1)	(2)
$\text{Boom}_{i,t}$	0.082***	0.022***
	(4.09)	(3.80)
$\text{Log Population}_{i,t}$	0.562***	0.546***
	(3.16)	(3.10)
Year FE_t	Yes	Yes
County FE_i	Yes	Yes
R^2	0.658	0.659
N	6,382	6,382

$$\textit{Log Deposits}_{i,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_3 \textit{Small Bank}_{i,t}$$
$$+ \beta_4 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t} + \text{Year } \textit{FE}_t + \textit{County FE}_i + \varepsilon_{i,t}$$

Panel B: Effect of Shale Boom on Bank Deposits: Counties With Different Bank Sizes

	Small Bank = Dummy	
	Boom = Dummy	Boom = Log Total Shale Wells
	(1)	(2)
$\text{Boom}_{i,t}$	0.092***	0.022***
	(3.84)	(3.39)
$\text{Small Bank}_{i,t}$	-0.006	-0.008
	(-0.54)	(-0.65)
$\text{Boom}_{i,t}$ * $\text{Small Bank}_{i,t}$	-0.038	-0.002
	(-1.04)	(-0.21)
$\text{Log Population}_{i,t}$	0.559***	0.546***
	(3.15)	(3.07)
Year FE_t	Yes	Yes
County FE_i	Yes	Yes
R^2	0.659	0.659
N	6,382	6,382

Table 5: Effect of Shale Boom on Economic Outcomes

This table reports the results of regressions which test the influence of shale booms on economic outcomes. The dependent variable in these regressions is log establishments in county *i* , industry j, year *t* . The explanatory variables are different shale boom variables, which have previously been defined. Log of county population is also included as a control, as are county-industry and industry-year (industry trends) fixed effects. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

$$Log\ Establishments_{i,j,t} = \alpha + \beta_1 Log\ Population_{i,t} + \beta_2 Boom_{i,t} + IndustryTrends\ FE_{j,t} + CountyIndustry\ FE_{i,j} + \varepsilon_{i,j,t}$$

	Boom = Dummy	Boom = Log Total Shale Wells
	(1)	(2)
$Boom_{i,t}$	0.022***	0.005**
	(2.94)	(2.28)
Log $Population_{i,t}$	0.691***	0.689***
	(16.49)	(16.31)
IndustryTrends $FE_{j,t}$	Yes	Yes
CountyIndustry $FE_{i,j}$	Yes	Yes
R^2	0.261	0.260
N	12,764	12,764

Table 6: Effect of Credit Supply Shock on Economic Outcomes: Differences-in-Differences
(Boom vs. Non-Boom, High Ext Finance Dependence vs. Low Ext Finance Dependence)

This table reports the results of regressions measuring the effect of a credit supply shock on different industry groups, one which has high dependence on external finance (Ind = High) and one which has low dependence on external finance (Ind = Low). The dependent variable in these regressions is log establishments. Panel A documents the differential effect of the credit supply shock on industries with high dependence on external finance and those with low dependence on external finance. The unit of observation in Panel A is log establishments of either Ind = High or Ind = Low in in county *i* , year *t* . Panel B represents a regression form of differences-in-differences and tests whether the credit supply shock differentially affects industries with high external finance dependence. Log of county population is also included as a control, as are county-industry and industry-year (industry trends) fixed effects. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for $High_j$ is not reported in Panel B, because it is subsumed by CountyIndustry $FE_{i,j}$.

$$\textit{Log Establishments}_{i,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \textit{Time FE}_t + \textit{County FE}_i + \varepsilon_{i,t}$$

Panel A

	Boom = Dummy		Boom = Log Total Shale Wells	
	Ind = High	Ind = Low	Ind = High	Ind = Low
	(1)	(2)	(3)	(4)
$Boom_{i,t}$	0.045***	-0.002	0.010***	-0.001
	(3.26)	(-0.19)	(2.90)	(-0.55)
Log $Population_{i,t}$	0.688***	0.695***	0.682***	0.696***
	(13.14)	(17.39)	(12.97)	(17.18)
Year FE_t	Yes	Yes	Yes	Yes
County FE_i	Yes	Yes	Yes	Yes
R^2	0.230	0.328	0.229	0.328
N	6,382	6,382	6,382	6,382

$$\textit{Log Establishments}_{i,j,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_3 \textit{High}_j + \beta_4 \textit{Boom}_{i,t} * \textit{High}_j$$
$$+ \textit{IndustryTrends FE}_{j,t} + \textit{CountyIndustry FE}_{i,j} + \varepsilon_{i,j,t}$$

Panel B

	Boom = Dummy	Boom = Log Total Shale Wells
	(1)	(2)
$Boom_{i,t}$	-0.001	-0.001
	(-0.18)	(-0.52)
$Boom_{i,t}$ * $High_j$	0.046***	0.012***
	(2.74)	(2.68)
Log $Population_{i,t}$	0.691***	0.689***
	(16.48)	(16.31)
IndustryTrends $FE_{j,t}$	Yes	Yes
CountyIndustry $FE_{i,j}$	Yes	Yes
R^2	0.263	0.263
N	12,764	12,764

Table 7: Effect of Credit Supply Shock on Economic Outcomes: Subdivided by Small Bank Market Share
(Boom vs. Non-Boom, High External Finance Dependence vs. Low External Finance Dependence)

This table reports a regression form of differences-in-differences for two different county groups, one county group which has high small bank market share (Bank = High Small Bank Mkt Share), and one county group with low small bank market share (Bank = Low Small Bank Mkt Share). The definition of small bank in these regressions is any bank with less than $500 million in assets adjusted for inflation. For this regression high small bank market share counties are defined to be counties in the top quartile of small bank branch market share. The dependent variable in these regressions is log of establishments. The explanatory variables are different shale boom variables, which have previously been defined. Additionally, an interaction between boom variables and the "High" external finance dependence dummy is included, this is the differences-in-differences coefficient of interest (β_4). Log of county population is also included as a control, as are county-industry and industry-year (industry trends) fixed effects. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for $High_j$ is not included separately, because it is subsumed by CountyIndustry $FE_{i,j}$.

$$\textit{Log Establishments}_{i,j,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_3 \textit{High}_j + \beta_4 \textit{Boom}_{i,t} * \textit{High}_j$$
$$+ \textit{IndustryTrends FE}_{j,t} + \textit{CountyIndustry FE}_{i,j} + \varepsilon_{i,j,t}$$

	Boom = Dummy		Boom = Log Total Horizontal Wells	
	High Small Bank Market Share	Low Small Bank Market Share	High Small Bank Market Share	Low Small Bank Market Share
	(1)	(2)	(3)	(4)
$Boom_{i,t}$	-0.016	0.006	-0.007	0.001
	(-0.70)	(0.91)	(-0.96)	(0.54)
$Boom_{i,t}$ * $High_j$	0.135***	0.025*	0.036**	0.007**
	(2.60)	(1.93)	(2.33)	(2.12)
Log $Population_{i,t}$	0.412***	0.717***	0.405***	0.714***
	(3.79)	(13.47)	(3.57)	(13.38)
IndustryTrends $FE_{j,t}$	Yes	Yes	Yes	Yes
CountyIndustry $FE_{i,j}$	Yes	Yes	Yes	Yes
R^2	0.124	0.398	0.120	0.398
N	3,370	9,394	3,370	9,394

Table 8: Effect of Bank Size and Credit Supply on Outcomes: Differences-in-Differences-in-Differences Regression

(Boom vs. Non-Boom, High Ext Finance Dependence vs. Low Ext Finance Dependence, High Small Bank Market Share vs. Low Small Bank Market Share)

This table reports results for a regression form of differences-in-differences-in-differences, where the coefficient of interest is the triple interaction term. The dependent variable in these regressions is log establishments in county i, year t, industry group j. The explanatory variables are different boom variables, which have previously been defined. The definition of small bank in these regressions is any bank with less than \$500 million in assets adjusted for inflation. These specifications provide results for two different measures of Small $\text{Bank}_{i,t}$. One measure is a dummy variable, set to 1 if a county is in the highest quartile of small bank branch market share in any given year and 0 otherwise (Small Bank = Dummy), while the other measure is the ratio of branches which belong to small banks relative to the total number of bank branches in a county (Small Bank = Ratio). Additionally, a set of fully saturated interactions between Boom variables, Small Bank variables, and the High external finance dependence dummy are included. The key coefficient of interest for the differences-in-differences-in-differences regression is the triple interaction term β_8. Log of county population is also included as a control, as are county-industry and industry-year (industry trends) fixed effects. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for High_j is not reported, because it is subsumed by CountyIndustry $\text{FE}_{i,j}$.

$$\textit{Log Establishments}_{i,j,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_3 \textit{High}_j + \beta_4 \textit{Small Bank}_{i,t} + \beta_5 \textit{Boom}_{i,t} * \textit{High}_j + \beta_6 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t}$$
$$+ \beta_7 \textit{High}_j * \textit{Small Bank}_{i,t} + \beta_8 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t} * \textit{High}_j + \textit{IndustryTrends FE}_{j,t} + \textit{CountyIndustry FE}_{i,j} + \varepsilon_{i,j,t}$$

	Boom = Dummy		Boom = Log Shale Wells	
	Small Bank = Dummy	Small Bank = Ratio	Small Bank = Dummy	Small Bank = Ratio
	(1)	(2)	(3)	(4)
$\text{Boom}_{i,t}$	0.010	0.034**	0.002	0.012***
	(1.44)	(2.15)	(1.05)	(3.01)
Small $\text{Bank}_{i,t}$	-0.002	-0.016	0.003	-0.008
	(-0.27)	(-1.07)	(0.44)	(-0.51)
$\text{Boom}_{i,t}$ * High_j	0.016	-0.065*	0.005	-0.015*
	(1.22)	(-1.82)	(1.64)	(-1.85)
$\text{Boom}_{i,t}$ * Small $\text{Bank}_{i,t}$	-0.042**	-0.055*	-0.015***	-0.021***
	(-2.02)	(-1.94)	(-2.71)	(-3.04)
Small $\text{Bank}_{i,t}$ * High_j	0.003	0.023	-0.005	0.011
	(0.25)	(0.94)	(-0.41)	(0.45)
$\text{Boom}_{i,t}$ * Small $\text{Bank}_{i,t}$ * High_j	0.109**	0.172***	0.027**	0.043***
	(2.30)	(2.60)	(2.27)	(2.73)
Log $\text{Population}_{i,t}$	0.692***	0.694***	0.688***	0.689***
	(16.45)	(16.39)	(16.28)	(16.22)
IndustryTrends $\text{FE}_{j,t}$	Yes	Yes	Yes	Yes
CountyIndustry $\text{FE}_{i,j}$	Yes	Yes	Yes	Yes
R^2	0.267	0.268	0.267	0.267
N	12,764	12,764	12,764	12,764

Table 9: Alternative Economic Outcome Measures

This table reports results for the regressions estimated in Table 6 (Differences-in-Differences) and Table 8 (Differences-in-Differences-in-Differences) with different economic outcome measures. Specification (1) uses Log Employment as an economic outcome variable, while (2) and (3) use establishments per capita (per 10,000 people), and (4) and (5) use log establishments per capita. Note, a differences-in-differences-in-differences specification is not reported for Log Employment because 62% of employment data is suppressed counties with high small bank market share. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for $High_j$ is not reported, because it is subsumed by CountyIndustry $FE_{i,j}$.

	Boom = Log Shale Wells				
	Small Bank = Dummy				
	Outcome = Log Employment	Outcome = Establishments Per Capita		Outcome = Log Establishments Per Capita	
	(1)	(2)	(3)	(4)	(5)
$Boom_{i,t}$	0.004	-0.402*	-0.259	-0.004*	-0.001
	(0.32)	(-1.73)	(-1.14)	(-1.68)	(-0.57)
Small $Bank_{i,t}$			-0.968		-0.001
			(-1.09)		(-0.14)
$Boom_{i,t}$ * $High_j$	0.035*	0.974***	0.575*	0.012***	0.005
	(1.65)	(2.85)	(1.87)	(2.60)	(1.59)
$Boom_{i,t}$ * Small $Bank_{i,t}$			-0.643		-0.011*
			(-1.13)		(-1.94)
Small $Bank_{i,t}$ * $High_j$			0.836		-0.006
			(0.79)		(-0.44)
$Boom_{i,t}$ * Small $Bank_{i,t}$ * $High_j$			1.778**		0.030**
			(2.16)		(2.22)
Log $Population_{i,t}$	0.797***				
	(17.96)				
IndustryTrends $FE_{j,t}$	Yes	Yes	Yes	Yes	Yes
CountyIndustry $FE_{i,j}$	Yes	Yes	Yes	Yes	Yes
R^2	0.142	0.113	0.117	0.096	0.102
N	10,112	12,764	12,764	12,764	12,764

Table 10: Sensitivity of Results to Industries

Panel A: This panel reports regression results of the key interaction coefficients of interest when excluding specific industries from the regression results originally reported in Table 6 Panel B (column (3) below) and Table 8 (column (4) below). The definition of Boom variable used in these regressions is log total shale wells, and the definition of Small Bank Share is a dummy variable for counties in the highest quartile of small bank branch market share. Additionally this table reports the asset beta for each industry, using two different methodologies.

Panel B: This panel reports regression results of the key interaction coefficients of interest when excluding the two highest asset beta industries from the regression results originally reported in Table 6 Panel B (column (3) below) and Table 8 (column (4) below). The definition of Boom variable used in these regressions is log total shale wells, and the definition of Small Bank Share is a dummy variable for counties in the highest quartile of small bank branch market share. Columns (6) and (7) report the average asset beta for each industry group when the two highest asset beta industry groups are excluded.

Panel C: This panel reports regression results of the key interaction coefficients of interest when including the industries which had been excluded from previous regressions due to the potential direct impact of shale booms on their businesses. The definition of Boom variable used in these regressions is log total shale wells, and the definition of Small Bank Share is a dummy variable for counties in the highest quartile of small bank branch market share.

Panel A

Industries		Excluding Industry			Asset Beta By Industry	
(1) Two Digit NAICS	(2) Two Digit NAICS Name	(3) $Boom_{i,t}$ * $High_j$	(4) $Boom_{i,t}$ * Small $Bank_{i,t}$ * $High_j$	(5) Ext Finance Dependence Flag	(6) β_{Asset1}	(7) β_{Asset2}
62	Health Care and Social Assistance	0.011***	0.024**	0	0.64	0.61
42	Wholesale Trade	0.012***	0.028**	0	0.79	0.77
11	Agriculture, Forestry, Fishing and Hunting	0.011***	0.027**	0	0.40	0.42
61	Educational Services	0.012***	0.027**	0	0.86	0.83
81	Other Services (except Public Administration)	0.011**	0.029**	0	0.51	0.49
44	Retail Trade	0.013***	0.032**	0	0.82	0.77
22	Utilities	0.010**	0.024**	0	0.21	0.21
56	Administrative and Support and Waste Management and Remediation Services	0.011***	0.023**	1	0.82	0.80
48	Transportation and Warehousing	0.009**	0.024*	1	0.60	0.58
31	Manufacturing	0.012***	0.022**	1	0.51	0.50
72	Accommodation and Food Services	0.013***	0.028**	1	0.61	0.59
71	Arts, Entertainment, and Recreation	0.013***	0.028**	1	0.69	0.64
54	Professional, Scientific, and Technical Services	0.012***	0.034***	1	1.18	1.19
51	Information	0.012***	0.030**	1	1.39	1.38
	Average for Low Dependence on External Finance				0.60	0.59
	Average for High Dependence on External Finance				0.83	0.81

Panel B

Industries		Excluding Industry			Asset Beta By Industry	
(1) Two Digit NAICS	(2) Two Digit NAICS Name	(3) $Boom_{i,t}$ * $High_j$	(4) $Boom_{i,t}$ * Small $Bank_{i,t}$ * $High_j$	(5) Ext Finance Dependence Flag	(6) β_{Asset1}	(7) β_{Asset2}
51-54	Two Highest Beta Industries (Codes 51 and 54)	0.012**	0.038***	1		
	Average Asset Beta of Low Dependence (Exclude Codes 51 and 54)				0.60	0.59
	Average Asset Beta of High Dependence (Exclude Codes 51 and 54)				0.65	0.62

Panel C

Industries		Adding Industry		
(1) Two Digit NAICS	(2) Two Digit NAICS Name	(3) $Boom_{i,t}$ * $High_j$	(4) $Boom_{i,t}$ * Small $Bank_{i,t}$ * $High_j$	(5) Ext Finance Dependence Flag
21	Mining (Includes Oil and Gas Exploration)	0.018***	0.038***	1
52	Finance and Insurance	0.010**	0.030**	0
99	Other	0.012***	0.026**	1
53	Real Estate and Rental and Leasing	0.012***	0.028**	1
23	Construction	0.008**	0.019*	0

Table 11: Falsification Test of Pre-Boom Trends

This table reports results of falsification tests for the regressions in Tables 6 and 8. Specifically, dummy variables are inserted for the two years prior to the beginning of shale well activity. The dependent variable in these regressions is log establishments in county i, year t, industry group j. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for $High_j$ is not included separately, because it is subsumed by CountyIndustry $FE_{i,j}$.

$$(1)\ \textit{Log Establishments}_{i,j,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_{2F} \textit{False Boom}_{i,t} + \beta_3 \textit{High}_j + \beta_4 \textit{Boom}_{i,t} * \textit{High}_j + \beta_{4F} \textit{False Boom}_{i,t} * \textit{High}_j + \textit{IndustryTrends FE}_{j,t} + \textit{CountyIndustry FE}_{i,j} + \varepsilon_{i,j,t}$$

$$(2)\ \textit{Log Establishments}_{i,j,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_{2F} \textit{False Boom}_{i,t} + \beta_3 \textit{High}_j + \beta_4 \textit{Small Bank}_{i,t} + \beta_5 \textit{Boom}_{i,t} * \textit{High}_j + \beta_{5F} \textit{False Boom}_{i,t} * \textit{High}_j + \beta_6 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t} + \beta_{6F} \textit{False Boom}_{i,t} * \textit{Small Bank}_{i,t} + \beta_7 \textit{High}_j * \textit{Small Bank}_{i,t} + \beta_8 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t} * \textit{High}_j + \beta_{8F} \textit{False Boom}_{i,t} * \textit{Small Bank}_{i,t} * \textit{High}_j + \textit{IndustryTrends FE}_{j,t} + \textit{CountyIndustry FE}_{i,j} + \varepsilon_{i,j,t}$$

	False Boom = Dummy, Boom = Log Total Shale Wells	
	Small Bank = Dummy	
	(1)	(2)
$Boom_{i,t}$	-0.001	0.002
	(-0.55)	(0.86)
False $Boom_{i,t}$	-0.001	-0.000
	(-0.30)	(-0.08)
Small $Bank_{i,t}$		0.005
		(0.58)
$Boom_{i,t}$ * $High_j$	0.013***	0.006*
	(2.81)	(1.78)
False $Boom_{i,t}$ * $High_j$	0.007	0.004
	(1.37)	(0.74)
$Boom_{i,t}$ * Small $Bank_{i,t}$		-0.015***
		(-2.76)
False $Boom_{i,t}$ * Small $Bank_{i,t}$		-0.006
		(-0.77)
Small Bank $Share_{i,t}$ * $High_j$		-0.009
		(-0.76)
$Boom_{i,t}$ * Small $Bank_{i,t}$ * $High_j$		0.030**
		(2.43)
False $Boom_{i,t}$ * Small $Bank_{i,t}$ * $High_j$		0.021
		(1.25)
Log $Population_{i,t}$	0.689***	0.688***
	(16.28)	(16.25)
IndustryTrends $FE_{j,t}$	Yes	Yes
CountyIndustry $FE_{i,j}$	Yes	Yes
R^2	0.263	0.267
N	12,764	12,764

Table 12: Falsification Test Using Non-Shale Growth Shocks

This table reports results of falsification tests for the regressions in Tables 6 and 8 using non-shale growth shocks in states adjacent to the seven shale states in this study. Specifically, dummy "Growth Shock" variables are inserted after high growth county-years such that the number of False Boom county years is approximately the same proportion of Shale Boom county years obtained in the main sample (roughly 5% of all county-years). The objective of this specification is to test whether general growth shocks differentially affect a particular industry group. The dependent variable in these regressions is log establishments in county i, year t, industry group j. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for $High_j$ is not included separately, because it is subsumed by CountyIndustry $FE_{i,j}$.

	Growth Shock = Dummy	
	Small Bank Share = Dummy	
	(1)	(2)
Growth $Shock_{i,t}$	0.152***	0.155***
	(10.52)	(9.26)
Small $Bank_{i,t}$		0.005
		(0.96)
Growth $Shock_{i,t}$ * $High_j$	0.002	0.011
	(0.07)	(0.53)
Growth $Shock_{i,t}$ * Small $Bank_{i,t}$		-0.008
		(-0.36)
Small $Bank_{i,t}$ * $High_j$		-0.002
		(-0.26)
Growth $Shock_{i,t}$ * Small $Bank_{i,t}$ * $High_j$		-0.027
		(-0.85)
Log $Population_{i,t}$	0.821***	0.818***
	(26.91)	(26.86)
IndustryTrends $FE_{j,t}$	Yes	Yes
CountyIndustry $FE_{i,j}$	Yes	Yes
R^2	0.307	0.308
N	22,932	22,932

Table 13: Retail Sales Changes in Boom Counties with Different Local Bank Sizes

This table reports the results of regressions which measure the effect of different boom variables on retail sales. The dependent variable in these regressions is the log of total retail sales by establishments in county i in year t. The regressions test whether there is a differential effect on retail sales based on local bank size. The Small Bank Dummy variable used in is equal to 1 if a county is in the top quartile of small bank branch market share in a given year, and 0 otherwise. The definition of small bank in these regressions is any bank with less than $500 million in assets adjusted for inflation. Retail sales data is from the U.S. Bureau of the Census Economic Census in 2002 and 2007 (conducted every 5 years). Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

$$\begin{aligned} \textit{Log Retail Sales}_{i,t} &= \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_3 \textit{Small Bank}_{i,t} \\ &+ \beta_4 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t} + \text{Year } \textit{FE}_t + \textit{County FE}_i + \varepsilon_{i,t} \end{aligned}$$

Effect of Shale Boom on Retail Sales: Counties With Different Bank Sizes

	Small Bank = Dummy	
	Boom = Dummy	Boom = Log Total Shale Wells
	(1)	(2)
$Boom_{i,t}$	0.048*	0.016***
	(1.65)	(2.67)
$Small\ Bank_{i,t}$	0.021	0.015
	(0.90)	(0.57)
$Boom_{i,t}$ * $Small\ Bank_{i,t}$	-0.064	0.003
	(-0.76)	(0.16)
$Log\ Population_{i,t}$	0.669***	0.661***
	(8.34)	(8.22)
$Year\ FE_t$	Yes	Yes
$County\ FE_i$	Yes	Yes
R^2	0.670	0.673
N	1,263	1,263

Table 14: Alternative Thresholds for Small Bank Size

(Boom vs. Non-Boom, High Ext Finance Dependence vs. Low Ext Finance Dependence, High Small Bank Market Share vs. Low Small Bank Market Share)

This table reports results for a regression form of differences-in-differences-in-differences, where the coefficient of interest is the triple interaction term. The dependent variable in these regressions is log establishments in county *i* , year *t* , industry group *j* . This table reports specifications using different small bank definitions. The regressions in this table report results for small bank cutoffs at \$200 million, \$500 million, and \$1 billion in assets adjusted for inflation. Additionally, the specifications report results for assets measured both at the bank level and at the bank holding company level. Small $\text{Bank}_{i,t}$ is a dummy variable, set to 1 if a county is in the highest quartile of small bank branch market share in any given year and 0 otherwise (Small Bank = Dummy). A set of fully saturated interactions between Boom variables, Small Bank variables, and the High external finance dependence dummy are included. The key coefficient of interest for the differences-in-differences-in-differences regression is the triple interaction term β_8. Log of county population is also included as a control, as are county-industry and industry-year (industry trends) fixed effects. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for High_j is not reported, because it is subsumed by CountyIndustry $\text{FE}_{i,j}$.

$$\textit{Log Establishments}_{i,j,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \beta_3 \textit{High}_j + \beta_4 \textit{Small Bank}_{i,t} + \beta_5 \textit{Boom}_{i,t} * \textit{High}_j + \beta_6 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t}$$
$$+ \beta_7 \textit{High}_j * \textit{Small Bank}_{i,t} + \beta_8 \textit{Boom}_{i,t} * \textit{Small Bank}_{i,t} * \textit{High}_j + \textit{IndustryTrends FE}_{j,t} + \textit{CountyIndustry FE}_{i,j} + \varepsilon_{i,j,t}$$

	Boom = Log Shale Wells				
	Small Bank = \$500M Dummy	Small Bank = \$200M Dummy		Small Bank = \$1B Dummy	
	Holding Company = Yes	Holding Company = No	Holding Company = Yes	Holding Company = No	Holding Company = Yes
	(1)	(2)	(3)	(4)	(5)
$\text{Boom}_{i,t}$	0.002	0.001	0.001	0.002	0.002
	(1.13)	(0.58)	(0.48)	(1.05)	(1.13)
Small $\text{Bank}_{i,t}$	0.003	0.001	-0.001	0.003	0.003
	(0.43)	(0.19)	(-0.11)	(0.44)	(0.43)
$\text{Boom}_{i,t}$ * High_j	0.004	0.007*	0.007*	0.005	0.004
	(1.34)	(1.92)	(1.95)	(1.64)	(1.34)
$\text{Boom}_{i,t}$ * Small $\text{Bank}_{i,t}$	-0.016***	-0.009**	-0.009**	-0.015***	-0.016***
	(-2.98)	(-2.25)	(-1.98)	(-2.71)	(-2.98)
Small $\text{Bank}_{i,t}$ * High_j	-0.009	-0.003	-0.017	-0.005	-0.009
	(-0.78)	(-0.22)	(-1.44)	(-0.41)	(-0.78)
$\text{Boom}_{i,t}$ * Small $\text{Bank}_{i,t}$ * High_j	0.032***	0.018**	0.019**	0.027**	0.032***
	(2.76)	(2.09)	(1.98)	(2.27)	(2.76)
Log $\text{Population}_{i,t}$	0.689***	0.689***	0.689***	0.688***	0.689***
	(16.28)	(16.30)	(16.29)	(16.28)	(16.28)
IndustryTrends $\text{FE}_{j,t}$	Yes	Yes	Yes	Yes	Yes
CountyIndustry $\text{FE}_{i,j}$	Yes	Yes	Yes	Yes	Yes
R^2	0.268	0.265	0.265	0.267	0.268
N	12,764	12,764	12,764	12,764	12,764

Table 15: Pre-Boom Banking Structure Robustness

This table reports results for a regression form of differences-in-differences-in-differences, where the coefficient of interest is the triple interaction term. This regression is similar to the regression results reported in Table 8, however, the bank size structure of a county is kept the same as the year prior to experiencing a boom. The dependent variable in these regressions is log establishments in county *i* , year *t* , industry group *j* . These specifications provide results for two different measures of Small Bank$_{i,t}$. One measure is a dummy variable, set to 1 if a county is in the highest quartile of small bank branch market share in any given year and 0 otherwise (Small Bank = Dummy), while the other measure is the ratio of branches which belong to small banks relative to the total number of bank branches in a county (Small Bank = Ratio). Additionally, a set of fully saturated interactions between Boom variables, Small Bank variables, and the High external finance dependence dummy are included. The key coefficient of interest for the differences-in-differences-in-differences regression is the triple interaction term β_8. Log of county population is also included as a control, as are county-industry and industry-year (industry trends) fixed effects. Standard errors are clustered by county, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level. Note: A coefficient for High$_j$ is not reported, because it is subsumed by CountyIndustry FE$_{i,j}$.

$$Log\ Establishments_{i,j,t} = \alpha + \beta_1 Log\ Population_{i,t} + \beta_2 Boom_{i,t} + \beta_3 High_j + \beta_4 Small\ Bank_{i,t} + \beta_5 Boom_{i,t} * High_j + \beta_6 Boom_{i,t} * Small\ Bank_{i,t}$$
$$+ \beta_7 High_j * Small\ Bank_{i,t} + \beta_8 Boom_{i,t} * Small\ Bank_{i,t} * High_j + IndustryTrends\ FE_{j,t} + CountyIndustry\ FE_{i,j} + \varepsilon_{i,j,t}$$

	Boom = Dummy		Boom = Log Shale Wells	
	Small Bank = Dummy	Small Bank = Ratio	Small Bank = Dummy	Small Bank = Ratio
	(1)	(2)	(3)	(4)
Boom$_{i,t}$	0.006	0.041**	0.001	0.012***
	(1.05)	(2.34)	(0.54)	(2.79)
Small Bank$_{i,t}$	0.009	-0.002	0.012	0.003
	(0.98)	(-0.15)	(1.25)	(0.18)
Boom$_{i,t}$ * High$_j$	0.020	-0.085**	0.006*	-0.019**
	(1.49)	(-2.42)	(1.94)	(-2.25)
Boom$_{i,t}$ * Small Bank$_{i,t}$	-0.025	-0.059**	-0.008	-0.020***
	(-1.23)	(-1.98)	(-1.55)	(-2.64)
Small Bank$_{i,t}$ * High$_j$	-0.003	0.021	-0.009	0.010
	(-0.21)	(0.82)	(-0.59)	(0.38)
Boom$_{i,t}$ * Small Bank$_{i,t}$ * High$_j$	0.082*	0.182***	0.019*	0.045***
	(1.91)	(2.92)	(1.70)	(2.81)
Log Population$_{i,t}$	0.692***	0.694***	0.689***	0.691***
	(16.47)	(16.31)	(16.25)	(16.11)
IndustryTrends FE$_{j,t}$	Yes	Yes	Yes	Yes
CountyIndustry FE$_{i,j}$	Yes	Yes	Yes	Yes
R^2	0.266	0.267	0.265	0.267
N	12,764	12,764	12,764	12,764

Appendix A: Bank Regressions for Banks with All Branches in One County

This table reports the results of regressions which estimate the effect of different shale boom variables on bank outcomes for banks that have all of their branches in a single county. The unit of observation in this panel is county *i* , bank *j* , year *t* . Data on banks was compiled from Call Reports and Summary of Deposit reports. A bank is in the sample if all of its branches are in a single county in given year, treatment banks are those banks which are in shale boom county-years, while control banks are single county banks in non-shale boom county-years. C&I loans are the total amount of commercial and industrial loans a bank reports on its Call Report. Interest income is the total interest income a bank generates in a year, divided by its average total loans. Deposit Interest Rate is the interest paid on all deposits divided by the average amount of deposits a bank has in a given year. Both interest rate and deposit interest rate variables were winsorized at 1% and 99%. Log of county population is included as a control, as are year fixed effects and bank fixed effects. Standard errors are clustered by bank, with t-statistics reported in parentheses below coefficient estimates, where * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

$$\textit{Bank Outcome}_{i,j,t} = \alpha + \beta_1 \textit{Log Population}_{i,t} + \beta_2 \textit{Boom}_{i,t} + \textit{Year FE}_t + \textit{Bank FE}_j + \varepsilon_{i,j,t}$$

Panel A

	Boom = Dummy			
	Outcome = Log Deposits	Outcome = Log C&I Loans	Outcome = Interest Rate	Outcome = Deposit Interest Rate
	(1)	(2)	(3)	(4)
$Boom_{i,t}$	0.0919***	0.1007**	0.0004	0.0004
	(4.11)	(2.32)	(0.13)	(1.09)
Log $Population_{i,t}$	0.2465***	0.3920***	-0.0020	0.0015***
	(4.94)	(4.73)	(-0.94)	(2.92)
Year FE_t	Yes	Yes	Yes	Yes
Bank FE_j	Yes	Yes	Yes	Yes
R^2	0.399	0.118	0.399	0.872
N	8,176	8,176	8,176	8,176

Panel B

	Boom = Log Total Shale Wells			
	Outcome = Log Deposits	Outcome = Log C&I Loans	Outcome = Interest Rate	Outcome = Deposit Interest Rate
	(1)	(2)	(3)	(4)
$Boom_{i,t}$	0.0265***	0.0297**	-0.0007	0.0000
	(4.22)	(2.55)	(-0.86)	(0.29)
Log $Population_{i,t}$	0.2457***	0.3911***	-0.0020	0.0015***
	(5.08)	(4.71)	(-0.89)	(2.92)
Year FE_t	Yes	Yes	Yes	Yes
Bank FE_j	Yes	Yes	Yes	Yes
R^2	0.401	0.119	0.399	0.872
N	8,176	8,176	8,176	8,176

Appendix B: Differences-in-Differences-in-Differences Regression, Table 8 Fixed Effect Robustness

This table contains the same regression specification as Table 8, except for the types of fixed effects used. This table provides results for various combinations of IndustryTrends $FE_{j,t}$, Year FE_t, CountyIndustry $FE_{i,j}$, County FE_i, CountyYear $FE_{i,t}$.

	Economic Outcome = Log Establishments					
	Boom = Log Total Shale Wells					
	Small Bank = Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$Boom_{i,t}$	-0.000	-0.008***	0.005	0.002		
	(-0.09)	(-3.80)	(1.42)	(1.05)		
Small Bank $Share_{i,t}$	0.064***	0.008	0.062***	0.003		
	(5.04)	(0.99)	(4.94)	(0.44)		
$Boom_{i,t}$ * $High_j$	0.010*	0.025***	0.000	0.005	0.010*	0.000
	(1.89)	(8.23)	(0.05)	(1.64)	(1.89)	(0.05)
$Boom_{i,t}$ * Small Bank $Share_{i,t}$	-0.023***	-0.016***	-0.022***	-0.015***		
	(-2.87)	(-3.00)	(-2.76)	(-2.71)		
Small Bank $Share_{i,t}$ * $High_j$	-0.126***	-0.015	-0.123***	-0.005	-0.126***	-0.123***
	(-5.81)	(-1.13)	(-5.68)	(-0.41)	(-5.81)	(-5.68)
$Boom_{i,t}$ * Small Bank $Share_{i,t}$ * $High_j$	0.044***	0.031***	0.042***	0.027**	0.044***	0.042***
	(2.97)	(2.63)	(2.81)	(2.27)	(2.97)	(2.81)
Log $Population_{i,t}$	0.688***	0.688***	0.688***	0.688***		
	(16.29)	(16.29)	(16.28)	(16.28)		
Year FE_t	Yes	Yes	No	No	No	No
IndustryTrends $FE_{j,t}$	No	No	Yes	Yes	No	Yes
County FE_i	Yes	No	Yes	No	No	No
CountyIndustry $FE_{i,j}$	No	Yes	No	Yes	No	No
CountyYear $FE_{i,t}$	No	No	No	No	Yes	Yes
R^2	0.794	0.227	0.797	0.267	0.817	0.82
N	12,764	12,764	12,764	12,764	12,764	12,764