[Job Market Paper]
Bank Heterogeneity and Capital Allocation: Evidence from 'Fracking' Shocks

Matthew C. Plosser*

November 28, 2011

Abstract

This paper empirically investigates the role of bank size on investment allocation decisions of banks over the business cycle. I generate exogenous shocks to bank deposits by exploiting the development of new oil and gas fields using 'fracking' technologies that result in significant cash windfalls to local landowners and large increases in local deposits. I trace the investment decisions of small- and medium-sized banks in response to these shocks and find that in normal times both types lend in response to a supply shock. The composition of their lending portfolio varies, with smaller banks investing more in 'soft' information intensive assets relative to larger banks. However, during the crisis period small banks invest significantly less of their incremental deposits in loans, hoarding liquidity instead. My findings suggest that during adverse times, even in a developed economy like the U.S., capital may become trapped inside several small banks.

*University of Chicago, Booth School of Business, mplosser@chicagobooth.edu. I am grateful to my Ph.D. advisors Anil Kashyap (Chair), Doug Diamond, Steven Kaplan, and Amit Seru for their generous guidance. I also thank Nina Boyarchenko, Alice Chen, Anna Costello, Eugene Fama, Valentin Haddad, Erik Loualiche, Gregor Matvos, Charles Plosser, Raghu Rajan, Shrihari Santosh, Amir Sufi, and seminar participants at the Chicago Booth Finance Workshop for helpful comments.

1 Introduction

It is widely accepted that the banking sector plays an important role in the allocation of capital, channeling funds from savers to borrowers. Banking theory has emphasized the role of bank size in accomplishing this task. Diamond (1984), for instance, argues the costs of a bank's delegated monitoring duties are minimized by diversification – a characteristic that larger banks with greater scope can deliver. However, a parallel stream of work suggests that small banks have a comparative advantage when lending to borrowers with more 'soft' information (Stein 2002). Hence, the theoretical literature implies very different investment responses by large and small banks to credit supply shocks. In this paper, I empirically investigate how supply shocks translate into investment decisions for these types of banks. Doing so allows me to make inferences on how an economy like the U.S., with a mix of small and large banks, reacts to supply shocks over the business cycle.

Estimating these effects is challenging. An examination of changes in bank asset allocations must address the endogeneity of investment opportunities and bank financing. Banks may raise financing in response to changes in their specific investment opportunities. In addition, latent economic factors may influence both the demand for credit and the supply of bank financing. These issues are particularly thorny when analyzing small banks where deposits and lending are geographically concentrated. This paper attempts to address these concerns by exploiting an exogenous shock to the credit supply of different banks and tracing their investment allocation decisions.

I use an unsolicited deposit shock to compare the marginal allocation decisions of different sized banks. Since 2001, rising energy prices and innovations in drilling technology have resulted in the development of several new oil and gas fields throughout the United States. The success of these 'fracking' fields has resulted in windfalls to local landholders who receive royalties and leasing payments from drillers. Landowners then deposit a large share of these proceeds in their local bank. The accidental nature of the deposit shock is plausibly exogenous to banks seeking deposits to advance bank investment opportunities. I compare the allocation choices of impacted small- and medium-sized banks and attribute the differences in the allocation decisions to technological differences in their ability to transform deposits into loans.

During the period 2001-2009, I identify ten new oil or gas fields that affect fifty-three 'treatment' counties. Using drilling and production data from various state agencies, I estimate the cash-flow shock to local landowners. I find that the payments to local landowners are large, with some counties receiving as much as \$300-\$400 million per year.

Using the county-level deposit data from the FDIC, I establish that local deposits are positively related to estimated cash payments (Figure 2). I conclude that these oil and gas shocks are economically meaningful and have a sizable impact on local bank deposits. I use this variation to generate a bank-level instrument for deposits by calculating the excess deposit growth in treatment counties weighted by the bank's share of deposits in the county. I then employ this instrument to estimate the allocation choices of small- and medium-sized banks in response to a change in deposits.

To estimate the impact of this shock on bank investment decisions, I construct a panel of bank balance sheet data using U.S. regulatory filings. I find that the shock influences one-hundred-sixty-five different banks from 2001-2009. To compare the response difference between small and medium banks, I interact the change in deposits with a measure of bank size, the log of total assets. I implement this non-linear IV by instrumenting for both the change in deposits and the interaction term between deposits and size. The interaction term is the coefficient of interest, as it estimates the variation in capital allocation with bank size.

The results indicate that there are significant differences in how banks allocate these incremental funds, both in the cross-section and over time (Table 11). In good times (2003-2006), small- and medium-sized banks invest most of their incremental deposits in loans. Relative to small banks, larger banks allocate more of investment to residential real estate loans and less to small business loans. Once the financial crisis set-in (2007-2009), incremental lending on average declined and the difference between banks widened. Larger banks lent even more of the incremental deposits (20% more of each new deposit) than small banks. In addition, during this period I find that small bank loan-to-deposit ratios and return-on-assets decline relative to their peers and relative to larger banks during the downturn. While these last two relationships are statistically weak, they are likely due to the allocation of funds to liquid assets, like securities and the federal funds market, rather than more profitable loans.

An alternative interpretation of my findings is that, despite my empirical design and several controls, the variation in supply I exploit could still be correlated with investment opportunities. In particular, it is possible that the heterogeneity in investment allocation across bank size over the business cycle is driven by differences in credit demand by the particular customers serviced by these banks. For instance, during the crisis period commercial real estate under-performed. Small banks are particularly exposed to demand for credit in this sector (Agarwal, Genay and McMenamin 2010). However, my results are robust across loan categories. Notably, small banks invest less in small business loans compared to their larger peers during the crisis. As prior research has shown, small businesses tend to borrow locally (Petersen and Rajan 2002). As a result, it is unlikely that small business credit demand contracted just for small banks since larger banks operating in the same areas extended more small business loans. We may also be concerned that small banks are investing less because they are relatively capital constrained during the crisis and the shock we are examining is a positive credit or liquidity shock. However, the results are robust to including Tier 1 capital ratios as a control.

Given these findings, I conclude that bank size is an important factor in determining how banks allocate their funds and that the role of bank size is sensitive to the economic conditions. Relative to larger banks, small banks are less likely to loan incremental funds in general, but more likely to lend to small businesses. This is consistent with theories of banking that argue for different technologies in the lending of large and small banks. Importantly, I find that small banks also appear to be more sensitive to downturns where they are more likely to invest in liquid assets, possibly in anticipation of a liquidity crisis. This suggests that an economy like the U.S. could find capital trapped inside several small banks when a negative shock hits the economy. Even in a well-developed economy, it seems that the availability of liquid funds in the inter-bank markets is not enough to prevent small banks from hoarding liquidity. These actions can have direct implications on the real side of the economy since capital does not flow as freely to users.

This paper contributes to our understanding of several topics related to banking, primarily by identifying important heterogeneity in the behavior of banks over the business cycle and in the cross-section. The first is the literature on the role of bank size on banks' ability to make loans and the types of loans they make (Diamond (1984), Stein (2002)). The growth period results imply that the smallest banks allocate more of the incremental funds from the deposit shock to 'soft' information investments (i.e. small businesses) and larger banks are more likely to allocate these funds toward more 'hard' information loans (i.e. real estate, individual credit). These results are consistent with empirical evidence on bank lending behavior following bank mergers (Berger, Saunders, Scalise and Udell 1998) and the types of small businesses associated with large and small banks (Berger, Miller, Petersen, Rajan and Stein 2005). However, my focus on *marginal* investment decisions, rather than average investment allocations, allows me to observe this behavior at various points in time. I find that small banks are not universally more likely to invest incremental funds in small businesses. During the crisis period their marginal lending allocations to small business is lower than larger banks, consistent with the notion that scale enhances a bank's ability to lend.

The second is the impact of liquidity or capital constraints. Bank lending behavior has been shown to be sensitive to supply shocks, dating to Peek and Rosengren (1997) and more recently Khwaja and Mian (2008) and Paravisini (2008). Whereas these recent results focus on developing markets, my findings suggest external financing frictions impact investment in a developed financial system where we would expect supply shocks to matter less given the access of banks to several sources of short term liquid financing. In addition, the heterogeneity in the impact of deposits, over time and across bank size, suggests that the technology of banks and the state of the world influence the degree to which a firm appears constrained. The literature has largely interpreted a sensitivity to financing as evidence of credit constraints; however, during the crisis period the relative lack of lending by small banks could reflect concern with future access to liquidity, mirroring empirical results that show firms without credit lines bypassed positive investments during the crisis (Campello, Giambona, Graham and Harvey 2011). Or, small bank behavior may be due to greater risk aversion among small bank owners (Saunders, Strock and Travlos 1990). The findings are of particular use when considering policy responses meant to ease financing constraints.

Third, this analysis relates to the macroeconomics literature on the role of banks in the business cycle (Bernanke and Gertler (1989), Rajan (1994)) and in particular the role of liquidity in financial crises (Diamond and Rajan 2005). Small bank lending is much more sensitive to the financial crisis, as the smallest banks appear to hoard these incremental funds in liquid assets. The unique differences in small bank lending do not appear to be explained by a paucity of investment opportunities, suggesting that it may be a result of these firms being more liquidity constrained during the crisis period (Brunnermeier 2008). The results are complementary to other empirical work focused on the cross-sectional behavior of banks in response to monetary policy (Kashyap and Stein 2000) and financial crises (Ivashina and Scharfstein (2010), Ennis and Wolman (2011)). While the real impact of this behavior may be small for a country like the U.S., with a wide-array of financing options, the lessons have a wider application to bank lending in less developed environs and the potential role of the lender of last resort.

The paper is organized as follows. Section 2 summarizes the empirical strategy and provides institutional background on shale gas development. Section 3 introduces the data and variable construction. Section 4 discusses the results and Section 5 concludes.

2 Empirical Strategy

This section presents an overview of my identification strategy followed by some background information on shale gas development.

2.1 Identification

The goal is to estimate the impact of bank size on the allocation of bank deposit shocks. I will examine the following relation between the change in loans and the change in deposits.

$$\Delta L_i^t = \lambda_D(s_i^{t-1})\Delta D_i^t + \varepsilon_i^t \tag{2.1}$$

Here, ΔL_i^t denotes the change in loans, L, for bank i from time $t-1$ to t. $\lambda_D(s_i^{t-1})$ is the share of the deposit change, ΔD_i^t, that is allocated to the left-hand side variable. The share is a function of the size of the bank prior to the change in deposits, s_i^{t-1}. The final term, ε_i^t, contains other factors that may impact changes in loans, including shocks to demand and other bank characteristics related to the allocation decision. Although 2.1 formulates the relation between loans and deposits, I will consider alternative uses for deposits by varying the left-hand side variable.

I transform this relation in order to estimate a linear model that is comparable across banks of different size. First, I scale the change in deposits and the left-hand side variable by the bank's total assets at time $t-1$, A_i^{t-1}, and denote them using lowercase designations (for example, Δd_i^t for deposits). Next, I parameterize the lambda term by interacting the change in deposits with the demeaned log of bank assets, $\tilde{s}_i^{t-1}$. I use this additional term to estimate the different allocation decision between large and small banks.[1] Finally, I include year fixed-effects, τ_t, to capture the impact of aggregate fluctuations, and a vector of controls for bank characteristics, $\mathbf{X_i^{t-1}}$ which includes size, s_i^{t-1} . My estimating equation is,

$$\Delta l_i^t = \beta_D \Delta d_i^t + \beta_{D,s}(\tilde{s}_i^{t-1} * \Delta d_i^t) + \tau_t + \mathbf{X_i^{t-1}}\psi + \varepsilon_i^t \tag{2.2}$$

The parameters of interest are β_D, the average percentage of a deposit change allocated to loans, and $\beta_{D,s}$, the variation in this allocation percentage with size. It is this second term that reveals the impact of size on the allocation decision.

[1] Demeaning bank size allows for easier interpretation of estimates but does not impact significance or magnitude of the interaction coefficient. I have found similar results using alternative specifications, including a dummy variable as an indicator for small banks (those below the median asset size).

Unfortunately, credit demand is difficult to observe and is likely correlated with changes in deposits. Hence, OLS estimation of Equation 2.2 will result in biased estimates. Why might credit demand and deposit supply be correlated? First, banks with more investment opportunities may seek deposits. They can solicit deposits by issuing Certificates of Deposit (CDs) at competitive rates, bidding for brokered deposits, expanding their branch network, and competing more aggressively at existing branches. This effect biases OLS estimates of the allocation of deposits into loans. I refer to this as 'selection' bias. Second, latent economic conditions may drive households and businesses to change both their desired level of deposits and their demand for credit. I call this 'local demand' bias. A bank's exposure to this bias is related to its geographic footprint.

In order to address these sources of bias, I introduce a variation in deposits, Δd_i^{t*}, which is not sought by banks. The variation is estimated using a bank's exposure to shale gas counties where deposits increase significantly. The construction of this variable will be explained in detail in Section 3. I then estimate Equation 2.2 using two-stage-least-squares (2SLS) where I use these unsolicited local deposit inflows as an instrument for deposit changes. I will argue that these deposit shocks are exogenous to bank selection. If these shocks are also exogenous to local demand shocks, we can estimate Equation 2.2 without being concerned about either source of bias.

One concern is that the shale shock impacts local credit demand, positively or negatively. This will bias an estimate of the average treatment effect (β_D). However, the comparison of banks exposed to a similar shock allows me to estimate the role of bank size ($\beta_{D,s}$), even if the estimated average treatment effect is not well identified. The intuition of this approach is analogous to a difference-in-difference (DiD) estimation. The first 'difference' is between treated banks exposed to shale gas shocks and similar untreated banks. My actual estimation method differs from DiD in that I measure this exposure using a continuous variable, Δd_i^{t*}, rather than a discrete variable. The effect I identify is how a bank allocates financing in response to this particular deposit shock relative to similar unexposed banks. The second 'difference' compares treated banks based on bank size, estimating the difference across banks exposed to the same treatment shock. The resulting estimate is a comparison of how banks respond differently to the same unsolicited increase in deposits.

In addition to accounting for contemporaneous shocks to supply and demand, there may be other sources of endogeneity that influence my estimates. I am most concerned with factors that might impact small banks differently relative to larger banks, thereby biasing the comparison between the two. First, some bank characteristics may vary with size and be correlated with changes in lending. I identify a number of these variables and include them as controls. Second, the deposit shock may not be allocated to small and large banks in proportion to their market share, with more of the local shock allocated to one or the other. I believe my estimation procedure accounts for this possibility by allowing the instrument's impact on deposits to vary with size in the first stage. Third, local economic shocks can impact credit demand differentially for large banks relative to small banks, especially if smaller banks are more dependent on local lending. To address this concern I compare allocation decisions within categories of loans, like small business lending or real estate. Credit demand shocks should affect both small and medium banks similarly within a category that is driven by local borrowing, differences can be attributed to something other than local credit demand. Each of these issues will be discussed in the results. I also conduct a series of robustness tests to address alternative explanations.

2.2 Background on 'Fracking'

Since 1999, rising energy prices and technological innovation have allowed drillers to recover oil and gas from geologic formations that were previously considered inaccessible or uneconomical. This development has resulted in significant payments to local landowners. The most common of these formations are gas shales (some shale formations actually produce crude oil).[2] As recently as 2000, gas shale was considered an inconsequential component of recoverable natural gas resources in the U.S. According to the Energy Information Administration (EIA), in 2011 there were over 22 major shale plays spread over 20 states and containing 827 trillion cubic feet (Tcf) of recoverable natural gas (or 'shale gas'). Shale gas now represents approximately one-third of the United States' recoverable natural gas reservoirs and at a recent price of $4.25 per thousand cubic-feet (Mcf) is worth approximately $3.5 trillion dollars at market.

Two advances in drilling technology have been critical to increasing recoverable resources: horizontal drilling and hydraulic fracturing. Shale rock is relatively impermeable and non-porous, reducing the flow of natural gas (or oil) into a traditional well. In addition, these formations tend to run in relatively narrow bands, limiting the effectiveness of a vertical drilling. Horizontal drilling allows a well to extend horizontally into the rock

[2] In addition to gas shales, there are 'tight' gas formations, typically made up of limestone or sandstone. Much like gas shale, these formations require advanced drilling techniques.

formation. Advances in the precision of this method have lowered the cost of drilling horizontally and increased yields. Once the well is drilled horizontally, hydraulic fracturing (or 'fracking') is used to increase the permeability of the formation. Fracking injects a slurry of water, sand and chemicals into the well, fracturing the surrounding rock and allowing trapped energy resources to flow into the well. The recent spate of new and successful fields can be attributed to these methods.

The discovery of a major shale gas field results in large cash payments to local landowners. Typically, developers drill a number of small test wells to ascertain the output of a formation in a particular area – this can take up to a year. Once they are confident of the output, they aggressively expand the acreage they lease from local landowners. Landowners receive payments in the form of an an upfront signing bonus based on the number of acres leased and a royalty on extracted resources. The signing bonus can vary from \$10 to \$20,000 per acre. The royalty can range from 10% to upwards of 25%.[3] Generally, these terms vary depending on the established reserves of the field, the desirability of the location, and the latest energy prices. In Section 3.1 I estimate the annual size of these payments and find that they can be as large as \$400 million or 50% of the deposit base for some counties. I also estimate that about 25% of these payments are channeled to local deposits.

The development of shale oil and gas fields is dominated by large corporations. Costs, as well as the growing technological sophistication, have limited the role of unaffiliated 'wildcatters' or entrepreneurs. A traditional, vertical well can cost as little as \$800,000, whereas a modern fracking well costs \$2 to \$8 million. Also, these wells require more operational expertise and have higher ongoing expenses. Major shale drillers include ExxonMobil, EOG Resources, Chesapeake Energy, Devon Energy, Pioneer Natural Resources, ConocoPhillips, Cabot Oil and Gas, among others. This is important to note because it suggests that this activity is financed at the national level, not by local banks.

Nevertheless, this shock is not entirely exogenous to local credit demand. While the primary financing for development and royalties comes from outside the local market, shale gas development influences local labor markets and households. A transient workforce generally drills wells, but local crews remain behind to maintain them. In response to greater employment and a sizable wealth shock, household credit demand can respond positively or negatively, with some recipients increasing their purchases and others choosing to pay-down debt.[4]

[3]Some states, like Pennsylvania, set a minimum royalty rate of 1/8th of output.

[4]This article notes examples of both types of behavior: Nossiter, Adam. "Gas Rush Is On, Louisianians

In addition, drilling can increase demand for ancillary businesses. Oil and gas must be transported to pipelines or rail terminals and fracturing requires firms to handle the processing and disposal of fluid used for fracking. A typical well requires over 5 million gallons of fluid, much of which must be processed after to use to remove additives that are harmful to the water supply. Incremental demand for labor and complementary businesses can generate a positive credit demand shock that co-varies with my deposit shock. Much of the analysis is concerned with comparing the impact on banks exposed to similar shocks. However, if large and small banks are differentially exposed to local demand shocks the result will be biased. The discussion of the results and robustness will consider these potential complications.

2.2.1 Shale Fields

I identify ten of the largest U.S. shale fields to implement my empirical strategy. Seven of the fields are predominantly natural gas fields, three are oil fields.[5] It is important for my analysis that the deposit shock is large and unsolicited, therefore I am careful to establish that the output of the field is significant and the date it was discovered. I will exclude banks that have followed development into a county rather than preceded it. Below I briefly discuss each field, its size, and when it began development. Table 1 contains a summary of the critical information.

The three oil fields tap the same underlying formation of shale (the Bakken Formation) that stretches from Canada to as far south as South Dakota and includes large parts of Montana and North Dakota. The Elm Coulee, Red River, and Bakken fields initiated development in the 2001-2004 time-frame. The Elm Coulee in Montana was the earliest and became a proving ground for shale oil extraction from the Bakken Formation.[6] Horizontal drilling spread to Southeast Montana (Red River, 2002) and then Western and Central North Dakota (Bakken, 2004).[7] As recently as 1995 the U.S. Geological Survey (USGS) estimated that there were 151 million barrels (Mbbl) of recoverable oil in the

Cash In." *The New York Times* July 29, 2008

[5]Wells primarily produce gas or oil, but there are oil byproducts from gas wells (condensate), and vice versa – most wells produce both products and both are collected and sold.

[6]Fialka, John. "Wildcat Producer Sparks Oil Boom on Montana Plains." *Wall Street Journal* April 5, 2006

[7]The dates for Red River and Bakken fields are determined by looking at drilling data and observing when the first horizontal well was drilled in these areas.

Bakken Formation. In 2008 a revised report raised that estimate to 3.0-4.3 billion barrels (Bbbl) of technically recoverable oil in Montana and North Dakota .[8] At a recent price of $75 a barrel this represents $275 billion in incremental value primarily spread over these three fields.

The Barnett field in Texas was the first of the natural gas fields. In 2000 drillers began experimenting with more sophisticated techniques to increase production from existing wells. After several years they found that the combination of horizontal drilling and 'fracking' could greatly increase the recoverable natural gas from the shale formation. By 2004 the USGS had determined there were 26.7 trillion cubic feet of recoverable gas in the Barnett shale or about $115bn worth of gas. Over 98% of this gas was recoverable because of new drilling methods. Following this success, drillers expanded exploration to shale formations throughout the country.

The remaining gas fields began between 2004-2008; they are Fayetteville (AR), Woodford (OK), Haynesville/Cotton Valley (TX), Haynesville (LA), Marcellus (PA), and Eagle Ford (TX). The Woodford shale in Oklahoma and the Cotton Valley in East Texas both began in 2004 as Barnett shale operators sought new areas. I establish start dates using the earliest horizontal drilling activity for these fields. The USGS believes the Woodford holds 24 Tcf of natural gas. The Cotton Valley shale actually lies above a deeper, and more recent discovery, the Haynesville shale. A number of companies began horizontal drilling in the Cotton Valley formation in 2004 and have more recently sought to drill the Haynesville shale beginning in 2008. Southwestern Energy believes this area of Eastern Texas contains approximately 20 Tcf of natural gas.[9] The Fayetteville field was revealed via a press release by Southwestern Energy in 2004. Southwestern estimates that there are 5 Tcf in reserves in this field.[10] The Haynesville field in Louisiana (2008) and the Eagle Ford field in Texas (2008) were publicized after companies established cheap footholds in an area.[11] In each of these cases, the announcements are the first real sign of a coming deposit boom.

The largest of all the natural gas formations is the Marcellus Shale. The Marcellus

[8] For reference, the U.S. consumes approximately 7 billion barrels of oil a year (EIA).

[9] Southwestern Energy Website, East Texas

[10] Southwestern Energy. "Southwestern Energy Announces Fayetteville Shale Play in Arkoma Basin." *News Release*. August 17, 2004, Southwestern Energy Website, Fayetteville Shale

[11] Chesapeake Energy. "Chesapeake Announces Haynesville Shale Discovery." *News Release*. March 24, 2008, Petrohawk. "Petrohawk Announces New Shale Gas Field Discovery" *News Release*. October 21, 2008

Shale extends from New York State South through Pennsylvania and Ohio to West Virginia. The Department of Energy and USGS estimate the Marcellus contains anywhere from 43-262 Tcf of natural gas. Thus far the bulk of the development has been restricted to Pennsylvania. The initial horizontal wells were drilled in 2007 and the success of these wells became well publicized in early 2008.[12] These ten fields constitute the largest 'finds' using unconventional drilling methods. The banks that operate in counties affected by these fields will be the focus of my empirical investigation.

3 Data Sources and Variable Definitions

This section describes the construction of the data. First, I introduce the county-level data necessary to evaluate the impact of shale gas development. Using this data, I identify counties impacted by shale drilling and demonstrate that drilling has a significant impact on local deposits. Then, I construct a measure of the local deposit shock that I can use as an instrument to estimate the equation of interest, 2.2. Finally, I describe banking data used in my primary estimation.

3.1 County Data and The Shale Shock

In order to determine which counties are most impacted by shale gas development, I obtain county-level oil and gas production from various state agencies. In some states, data is directly attributed to the underlying formation. For these fields I can easily identify oil and gas production related to new shale drilling. In other states, only aggregate production for the county is provided and it is not possible to directly attribute output to new shale wells. For these states, I determine a base year prior to shale gas discoveries and assume that incremental production is attributable to shale output.[13] Using the production and permit data, I identify 53 counties as shale counties or 'treatment' counties. I exclude counties where development is nascent and production is insufficient to impact deposits. Table 1 lists the data sources and lists the number of treatment counties in each field. Note that I only consider the first five years of development as treatment years, inclusive of the start year. This restriction only binds for the earliest fields and is intended to focus

[12] Gold, Russell. "Gas Producers Rush to Pennsylvania." *The Wall Street Journal* April 2, 2008

[13] States that require this method are Louisiana and Montana.

on the initial shock that landowners and banks are likely to anticipate. Results are robust to this exclusion.

Figure 5 illustrates the location of these counties. As we can see the treatment counties are clustered in three regions. The first is Montanana/North Dakota, where the Bakken oil field is located. The second is Texas/Oklahoma/Arkansas/Louisiana. There are three fields in Texas, the Barnett, Haynesville/Cotton Valley, and Eagle Ford and one field in each of the other three states. The final region, Pennsylvania, contains the Marcellus shale.

I obtain county-level deposit data from the Federal Deposit Insurance Corporation (FDIC) *Summary of Deposits* (SOD) database. SOD provides deposits by bank and location as of June 30. I aggregate deposits at the county-level and calculate two-year growth rates, from June to June. I use a two-year period to compare deposit growth in the county to calendar-year variables, like oil and gas production, that occur in the intervening year. Note that my timing conventions throughout the paper will refer to the calendar-year shock. For example, the year 2007 refers to calendar-year shale production in 2007 (t), but the corresponding two-year change in deposits is measured from June 30, 2006 to June 30, 2008 ($t-1$ to $t+1$), allowing me to capture the impact of the calendar-year variable in deposit growth. This will also be true in the bank sample described in Section 3.2.

I also add demographic, employment, and business data on counties. I obtain demographic data from the U.S. Census, including the population of the county by year, the percent of the population with a high-school degree in 2000, the share of the population that is Hispanic, the share that is black, and the share of the population that is working age (15-64). Data on economic activity comes from the Quarterly Census of Employment and Wages (QCEW) and County Business Patterns (CBP) at the Census Bureau. These data sets provide me with county measures of wages paid, employment, and small business establishments by NAICS industry. Small businesses from the CBP are defined as those businesses with fewer than 50 employees.

The resulting panel of U.S. counties has over 3,000 observations per year covering a nine year span. In contrast, my treatment group of 204 county-year observations is quite small. The treatment counties have fewer deposits, fewer employees, and a lower population density relative to the average of the full sample. Shale counties are also concentrated in three geographic regions which may have their own unique shocks and economic conditions. To generate an appropriate comparison group to the set of treatment counties, I create a matched sample using propensity score matching (Rubin (1973), Rubin (1979)). p-scores are calculated using a logit of a treatment indicator on log deposits, the population density, log population, county level demographic data, employment share by industry, year fixed effects and regional fixed effects. Each of these controls are lagged to reflect values prior to the start of the calendar year. I match each treatment observation with its five closest neighbors on propensity scores to generate a sample of untreated counties. The matched counties are weighted by the number of times they are selected with the sum of weights equaling the number of treatment observations (204). In effect, I have created a single untreated observation for each treatment observation.

Table 2 summarizes key characteristics for my treatment group of counties, the matched sample, and the full sample. The matched sample more closely resembles the treatment sample on these observables. They are smaller, have fewer deposits, and their employment composition more closely resembles that of the treatment group. The two-year deposit growth for the treatment growth is noticeably larger than its matched peers, 16.9% versus 10.0%. Treatment counties also have more growth in employment, wages and small businesses.

3.1.1 Impact of Shale Cash-Flow on Deposits

Before analyzing bank responses to shale gas development, I first establish that local deposits vary with estimated local payments to landowners. I construct an estimate of the cash-flow paid to local landowners using the drilling permit and production data. I use this estimate to determine if deposit growth is sensitive to local payments and to gauge if the relationship is reasonable. Cash-flow from shale gas development in county j at time t, CFS_j^t, is comprised of two terms, the royalty on sales and the upfront leasing bonus.

$$CFS_j^t = (BBL_j^t * P_{Oil}^t + MCF_j^t * P_{Gas}^t) * Royalty + Permits_j^t * Acreage * Bonus/Acre \quad (3.1)$$

The first term calculates royalties paid to landowners. The production of oil is in barrels, BBL, and the production of gas is in thousand cubic feet, MCF. I use the West Texas Intermediate Crude (WTI) series for the price of oil, P_{Oil}, and wellhead gas prices from the relevant region for natural gas prices, P_{Gas}. Both prices are provided by the U.S. Energy Information Administration (EIA). I assume a royalty rate of 20% based on the midpoint of reported royalties.[14]

The second term in Equation 3.1 is the upfront bonus paid to landowners. I do not observe lease signings or their bonus terms. However, leases must be held before drilling permits can be granted so I use drilling permits, *Permits*, as a proxy for new lease agreements. I then assume the average number of acres leased per permit, *Acreage*, and the cash bonus paid per acre, *Bonus/Acre*. According to the DOE, wells are generally allocated in areas of 60-160 acres,[15]; therefore, I assume 100 acres per permit and a bonus of $7,500 per acre.[16] These assumptions are not critical, but they allow me to present the shale gas activity as a single dollar index to estimate its relation with deposit growth.

The impact of cash-flows varies during the sample period, Table 3. In the early periods (2001-2004) the average cash-flows are small, about 5% of the deposit base. Few counties are impacted and they are in the earliest stages of development. Drilling activity intensifies by 2005-2006 as more counties enter the productive second and third year of development and energy prices rise. During this period cash-flows average 10% and are as high as 100% in some counties. The 2008-2009 period has the highest average impact relative to deposits. High prices and increased drilling activity deliver average cash-flows that are 30% of the existing deposit base. 2009 has a noticeable drop-off in the number of treatment counties from the high 30's in earlier periods to 16 counties. The decline in 2009 is the result of earlier fields exiting the sample as I only consider the first four years of development.

Figure 2 illustrates how treatment county deposit growth increases as shale gas fields develop and cash-flows to landowners rise. The deposit growth in the chart is demeaned by year. Excess deposit growth is approximately zero prior to the start of drilling (event year zero). In the years after a discovery, cash-flows as a percent of existing deposits increase and deposit growth rises. By year four, the average inflow of cash to landowners is almost 40% of the deposit base in the county and deposits are growing at 12% faster than the annual average. If I assume the excess deposits are the result of shale cash, by the fourth year roughly 25% of the cash payments to landowners are deposited in local banks.

[14]These terms are generally not published. I surveyed press reports on shale gas development as well as message boards and websites where landowners discuss the terms of their leases with other interested landowners who may be negotiating a lease.

[15]*Modern Shale Gas Development in the United States: A Primer* The U.S. Department of Energy

[16]Bonuses rose with prices after 2005. Therefore I only include this cash bonus for fields developed more recently (Woodford, Haynesville (LA), Marcellus, and Eagle Ford).

I corroborate the impact of the cash-flow shock on deposits by running a pooled cross-sectional regression of two-year county level deposit growth, $\%\Delta D_j^t$, on the cash-flow shock scaled by lagged deposits, and various controls.

$$\%\Delta D_j^t = \beta(CFS_j^t/D_j^{t-1}) + \psi Controls_j^t + \varepsilon_j^t \tag{3.2}$$

Deposit growth is trimmed at the 1% level to reduce the impact of extreme observations. Controls include log deposits, share of the population that is Hispanic, share of the population that is black, percent of the population with a high school degree, share of the population that is of working age (15-64), log of the population and the population density. I estimate this regression in the full and matched samples. When I consider the matched sample, I include the *p*-score, the predicted probability of being in the treatment group, from the matching probit outlined earlier. Standard errors are clustered by state to account for arbitrary serial correlation of county errors and cross-sectional correlation of errors within a region (Bertrand, Duflo and Mullainathan (2004), Petersen (2009)).

Table 4 contains the results of the estimation. As the results are consistent in both significance and magnitude across the two samples, I will not discuss them independently. Columns 1-3 demonstrate that the cash-flow shock is significant at the 1% level and robust to the inclusion of county-level controls and year fixed effects. A coefficient of 0.22-0.25 is economically reasonable and consistent with Figure 2. I interpret this to mean approximately 23% of cash-flow from these shocks is deposited in local bank accounts.

If deposits can be explained by economic activity rather than my estimate of cash-flow, it is difficult to argue that the deposits are largely exogenous to credit demand. In Column 4, I include growth in wages paid and growth in the number of small businesses as additional controls. Both are positively related to deposit growth, but they do not mitigate the magnitude or significance of the shale shock coefficient. Columns 5 and 6 test the importance of the shock in subperiods pre- and post-financial crisis. Again, the coefficients are significant at the 1% level. The coefficient is slightly smaller in the earlier period, 0.19, versus the alter period, around 0.25. I conclude from Table 4 that shale drilling activity is positively related to local deposit growth, that the impact is not explained by contemporaneous increases in employment and small businesses, and that the excess growth is above and beyond what similar counties experience at the time.

3.1.2 Constructing the Instrument

I wish to exploit county-level variation in shale gas exposure to generate a bank-level instrument that can be used in my empirical investigation. One candidate for this variation is the predicted value from a regression of deposit changes on my estimate of cash-flow shocks, i.e. Eq. 3.2. However, my estimate of cash-flows is a noisy measure of true cash-flows. I make a number of assumptions that may vary across fields and over time. Also, the linear prediction assumes a homogenous impact of cash-flows on deposits. But, the portion of the cash-flow that is deposited may depend on a number of local characteristics that vary in the cross-section. Rather than relying on this estimate, I calculate excess deposit growth in treatment counties by differencing deposit growth for these counties, $\%\Delta D_j^{Treat,t}$, with the mean of the five nearest neighbors from the matching exercise above, $\overline{\%\Delta D_j^{Match,t}}$.

$$\%\Delta D_j^{t*} = \%\Delta D_j^{Treat,t} - \overline{\%\Delta D_j^{Match,t}}$$

I assume that this excess deposit growth is attributable to incremental deposits from shale gas payments. The deposit shock resembles the pattern of cash-flows over time (see Table 3). I find that the average excess deposit growth is negative in the early years, but by 2005 is on average 5.7% with a maximum of 26.4%. The average excess deposit gets as high as 12.9% in 2009. The excess deposit growth impacts a range of counties over time and is large enough to affect local banks.

Bank allocation decisions will depend on the permanence of the deposit inflows. If banks anticipate deposit inflows will reverse, they will invest in more liquid, shorter maturity assets. Table 5 shows the mean excess deposit growth over event time where the event is the first year of development. I only use the first four event years as treatment years, but I extend the sample for this table in order to gauge how deposit growth behaves at longer horizons. The observed time-series for many of these fields is short; however, the average excess deposit growth is consistently positive, suggesting that on average the level shock to deposits is persistent. Also, although a decline in energy prices in 2009 reduced anticipated royalties and bonus pay-outs, mean excess deposit growth in shale counties remains positive (see Table 3). The unsolicited deposit shock appears relatively permanent, limiting the need for banks to invest in liquid, short-term assets.

I link the counties to banks using the SOD. The FDIC data set provides deposit locations by bank. For each bank, I aggregate the positive realizations of treatment county excess deposit growth by weighting the county excess deposit growth, $\%\Delta D_j^{t*}$, by the share of the bank's total deposits held in the county and summing across all of the counties in which the bank has deposits. This generates a bank-level variable for bank i, from the county level, j, shocks.

$$\%\Delta D_i^{t*} = \sum_{j=1}^{N} \%\Delta D_j^{t*} \left(\frac{D_{i,j}^{t-1}}{D_i^{t-1}} \right) \tag{3.3}$$

I set excess deposit growth to zero for non-shale counties or shale counties where the estimated excess growth is negative. I exclude banks that open branches in a treatment county after the initial year. I transform the estimate of deposit growth into an estimate of dollar changes in deposits scaled by prior period assets.

$$\Delta d_i^{t*} = \frac{\Delta D_i^{t*}}{A_i^{t-1}} = \%\Delta D_i^{t*} \frac{D_i^{t-1}}{A_i^{t-1}}$$

The result is a bank-level measure of exposure to unsolicited deposit shocks. This will be the primary instrument in my estimation strategy.

A concern is that the deposit shocks will mechanically have less impact on bigger banks as the sum of shocks is scaled by a larger deposit base. Table 6 summarizes the average bank-level shock across a range of bank sizes to illustrate the variation in exposure at various size percentiles. There is a meaningful impact, up to 5% of assets, through the 99th percentile (or about \$1bn in assets). For the largest banks, the maximum observed shock falls dramatically. How is it possible that the shock impacts such a wide array of sizes so similarly? First, as assets grow, banks tend to have a greater share of local deposits, increasing their exposure to the shock. Second, as banks branch out they are more likely to be exposed to multiple treatment counties. Another important takeaway from Table 6 is that the variation provided by the shale gas shock is too small to impact the largest banks, therefore my discussion will focus on small- and medium-sized banks with less than \$1 billion in assets.

3.2 Bank Data

I construct a panel of bank financials by combining data from various regulatory agencies for the relevant period, 2001-2009. Chartered commercial banks must provide detailed financials to the FDIC on a quarterly basis in *Call Reports of Income and Condition* (FFIEC Form 031). Bank holding companies file similar reports with the Federal Reserve.[17] Many commercial banks are subsidiaries of bank holding companies. As banks

[17] I do not consider thrifts due to constraints on the available data.

have been shown to establish internal capital markets (Houston, James and Marcus 1997), I am careful to attribute subsidiary banks to their highest holder. When available I attribute ownership to the financial high-holder (RSSD9364), otherwise I use the regulatory direct holder (RSSD9379). Because some bank holding company holders are themselves subsidiaries, I iterate on this process until I identify each bank's ultimate parent.

If the high-holder is a bank holding company with assets more than \$250m, the bank holding company is required to report consolidated financials on form *FR Y-9C*. I use the consolidated entity's financials for these firms. However, if the high holder is a small bank holding company ($<$ \$250m) or a non-bank financial institution, the high-holder may not report consolidated financials. For these, I consolidate accounts across the parent bank's unconsolidated balance sheets and their subsidiaries to generate a consolidated high-holder balance sheet. I focus my analysis on the second quarter (June 30) report date and discard the other quarterly reports. The June 30th report matches bank balance sheet data to the timing of the local deposit data. The second quarter report also contains additional data on banks' small-business lending. I exclude banks that have more than \$1.2bn in assets, have been in the sample less than two years, have made an acquisition or sale of assets in the past two years, or whose assets are predominantly credit cards or real estate.[18]

To estimate the equation of interest (Eq. 2.2) I calculate dollar differences in balance sheet quantities, like deposits or loans, and then scale by the prior period assets. Differences in balance sheet quantities are calculated over two-year intervals ($t-1$ to $t+1$) to compare changes in bank balance sheets to calendar-year variables. t references the calendar-year (as discussed in Sec. 3.1). I consider several loan categories: real estate, small business, commercial and industrial (C&I), and individual. In addition, I examine other uses of incremental liquidity, like cash, securities, the federal funds market or debt pay-down. Appendix A contains a summary of key variables and their construction.

As at the county level, I have a surfeit of untreated bank observations that do not share a common support with the treatment group. There are 438 bank-year observations in the treatment sample and more than 46,000 observations in the entire sample from 2001-2009. As with the counties, I generate a matched sample to serve as a counterfactual to the treatment group. I match each bank to its five nearest neighbors using propensity score matching. I calculate the propensity score using a logit of a treatment dummy on

[18] I identify bank mergers and major asset sales using the most recent Merger Information file provided by the Bank of Chicago Federal Reserve.

bank financial controls, county-level characteristics, year fixed-effects, and region fixed effects. The financial controls include log assets, ROA, the reserve ratio, the tier 1 capital ratio, the balance sheet composition (for example, percent of assets in real-estate loans), and indicators for the type of high-holder. The county characteristics are weighted by the percent of deposits the bank has in the county and summed to create a bank level exposure to county controls. County controls include log population, population density, the percent of the population with a high-school degree, the share of the population that is Hispanic, the share that is black, and the share of the population that is working age.

Table 7 summarizes the financial characteristics of the treatment group of bank observations, the matched group and the full sample. The treatment group on average invests 59% of its balance sheet in loans with the bulk of that being categorized as real estate loans (36%). Small business lending consists of both real estate and C&I loans with a principal smaller than \$1m. Treatment banks allocate 20% of their balance sheet to these smaller loans. Outside of lending, the largest allocation category is securities, federal funds sold (FFS) and repurchase agreements, with 28% of the balance sheet. Note that the dollar change in deposits scaled by assets, Δd_i^t, is 20% relative to the matched sample of 12.8%. We would expect this given we selected these banks because of their shale gas exposure. The average estimated impact of the shale gas shock on deposits, Δd_i^{t*}, is 7%, consistent with the surplus change in deposits. The matched sample has a much closer asset composition to the treatment group. It also has similar size (\$230m) and capital ratios, relative to the full sample. In addition, region fixed effects and county characteristics (not reported in this table) ensure that matched banks face similar opportunities outside of the dimensions impacted by shale gas activity.

The matching exercise generates a sample of similar banks, not impacted by the shock. This counterfactual sample allows me to estimate how the treatment banks responded differently than their peer group. However, I am primarily interested in comparing banks of varying sizes. The matching exercise matches small banks to untreated small banks and large banks to untreated large banks, but because I want to compare small banks to large banks, I need to consider other factors that co-vary with size.

Table 8 summarizes bank characteristics for banks above and below the median asset size in two time periods, before (2001-2006) and during the financial crisis (2007-2009).[19] On average a medium bank in this sample has \$350m in assets versus a small bank average of \$55m. The large banks lend a greater fraction of their assets, 62% to 55% for

[19]Note that I choose 2007 as a crisis year because bank balance sheet variables are as of June 30, 2008.

small banks, with the difference primarily in incremental real estate loans. Small banks maintain higher Tier 1 capital ratios, 16-17% versus 12-13%. A lower proportion of small banks are bank holding companies, 70-77% versus 92-93% of large banks. Finally, their propensity scores differ. p-scores estimate the likelihood of a bank being in a treatment county based on all the characteristics included in the matching exercise and are therefore a parsimonious way of controlling for a number of factors. Based on these differences I include lagged loan share of assets, Tier 1 capital ratios, an indicator for organization type, and p-scores as controls in the following regressions.

4 Analysis and Results

First, I estimate the average response to the unsolicited deposit shock. I find that instrumenting for deposits significantly impacts estimates of bank allocation decisions. I also discover that banks allocate these incremental funds differently over the business cycle. Second, I investigate the role of size using an interaction term. I show that bank size plays an important role in bank allocation decisions, particularly during the crisis period.

4.1 Average Allocation of Deposits

Before considering the impact of size on investment decisions, I estimate the average impact of the unsolicited deposit shock on bank deposits and asset allocations. I compare the OLS results to 2SLS estimates where I use the change in excess deposits for shale counties as an instrument. The first stage regresses the change in a bank's total deposits on the instrument that measures a bank's exposure to county-level shale deposit shocks. The second stage estimates how the deposit change is correlated with a specific balance sheet account (the left-hand side variable).

I estimate the following pooled cross-sectional regressions using 2SLS.

$$\text{First Stage:} \quad \Delta d_i^t = \pi_0 + \pi_D \Delta d_i^{t*} + \tau_t + \pi' \mathbf{X_i^{t-1}} + \epsilon_i^t \tag{4.1}$$

$$\text{Second Stage:} \quad \Delta l_i^t = \beta_0 + \beta_D \Delta d_i^t + \tau_t + \psi' \mathbf{X_i^{t-1}} + \varepsilon_i^t \tag{4.2}$$

Here, Δd_i^t is the two-year change in deposits scaled by assets at $t-1$. I interpret β_D as the average share of the deposit change allocated to the dependent variable, in this case loans, Δl_i^t. Year fixed effects, τ_t, control for aggregate variation over time. The vector of controls, $\mathbf{X_i^{t-1}}$, includes lagged observations of log real assets, loan share of assets, the Tier 1 capital ratio, an indicator for organization type, and the propensity score. I instrument for bank deposits using the unsolicited deposit shock, Δd_i^{t*}. Standard errors are robust to heteroskedasticity and clustered by county. Treatment banks are assigned to the county cluster where they experience the largest shale shock. Untreated banks are clustered together separately. This form of clustering accounts for arbitrary serial correlation within a shock county (and by extension within a bank) and cross-sectional correlation in a shock county (Bertrand et al. 2004).

Table 9 summarizes the results of the OLS and 2SLS regressions. In addition to total loans on the left-hand side, I consider other uses: liquid assets (cash, securities and short-term loans), non-deposit borrowing of the bank, commercial and industrial loans (C&I), and real estate loans. Panel A contains the OLS estimates of β_D (Eq. 4.2). I expect banks with investment opportunities to seek financing, therefore lending and deposits should be positively correlated. Consistent with this expectation, the estimate of β_D implies that for every dollar change in deposits, there is a 0.72 change in loans, Column A.2 (Panel A, Column 2). This is in excess of the average allocation of assets in loans of 60%. The allocation of deposit changes can be attributed to three balance sheet accounts: Loans, liquid assets, and non-deposit borrowing. The allocation share, β_D, for these three items should roughly sum to one, as deposit changes must be offset elsewhere on the balance sheet. On average 35% of deposit change is allocated to liquid assets and 2% to other borrowing (Columns A.3 - A.4). Note that the positive correlation on other borrowing suggests that deposit changes are positively related to non-deposit borrowing. This is consistent with the notion that deposit changes are endogenous, as banks with more (fewer) investment opportunities are raising (lowering) funds by all available means. Columns A.5 and A.6 present results for the two largest loan sub-categories, C&I loans and real estate loans. The coefficients on these two categories, 0.15 for C&I and 0.50 for real estate, sum to 0.65 and explain a large portion of the overall change in loans (0.72).

Contrast these results to Panel B, which contains the first- and second-stage estimates using the unsolicited deposit shock as an instrument. The first stage finds that the exposure to shale gas deposit shocks is highly correlated with bank changes in deposits (Column B.1), with a coefficient of 0.56 on the instrument. The second-stage estimates find the portion of deposits allocated to loans is 57% (Column B.2), 16% less than the OLS estimate. The difference can be explained by the increased allocation to liquid assets, 48% (Column B.3). This is consistent with the presumed endogeneity – OLS estimates are biased toward lending because banks raise (lower) deposits in response to increased (decreased) lending opportunities. By instrumenting using a shock that I believe is unsolicited by banks, I find that the deposit change is less likely to be allocated to lending. In addition, borrowing decreases by 4.2% of the change in deposits (Column B.4). Compared to the positive relation estimated via OLS, banks impacted by this shock are not on average seeking funding from other sources; they are reducing their non-deposit borrowing.

I repeat the estimation of Eq. 4.2 for two sub-periods: 2003-2006, when the economy was growing, and 2007-2009, the crisis period. I find the average impact of the deposit shock varies over the business cycle. Table 10 summarizes the results. In Columns 1 and 4, the OLS estimates of the relation between deposits and loans, β_D, are very similar across the two time periods, with coefficients of 0.72 in the earlier period and 0.73 in the later period. However, the 2SLS estimates vary. The first stage impact of the instrument on deposits is 0.45 in the earlier period and 0.90 in the later period (Columns 2 and 5). In both cases the coefficients are significant at the 1% level. In the boom period, the second stage estimate of the allocation of deposits to loans, 0.79, is within one standard deviation of the OLS estimate of 0.72. However in the bust period, the estimate is 0.28, significantly lower than the OLS estimate of 0.73. In unreported results, I verify that there is a change of similar magnitude and opposite sign to the allocation of liquid assets. I infer from this that there are important state dependencies in how banks respond to a positive liquidity shock. The dramatic change in the marginal propensity to lend might be attributable to decreased demand for credit, weaker capital positions at banks, or reduced access to liquidity. These issues will be explored further when when I compare banks of similar size in the next section.

4.2 Impact of Bank Size on the Allocation of Deposits

To estimate the role of size I interact changes in deposits with demeaned log of real assets,$\tilde{s}_i^{t-1}$.[20] The interaction of the endogenous variable with an exogenous variable creates a non-linear IV. I address this by instrumenting for both the change in deposits, Δd_i^t, and the interaction term between deposits and demeaned bank size, $(\tilde{s}_i^{t-1} * \Delta d_i^t)$,

[20]Demeaning the interaction term does not impact estimates of $\beta_{D,s}$, however it does simplify interpretation of magnitudes by estimating β_D at the average bank size. β_D can therefore be interpreted as the allocation of the deposit change to the left-hand side variable for the average bank. $\beta_{D,s}$ is the deviation from this average for banks of varying size.

using the unsolicited deposit shock, Δd_i^{t*}, and an interaction between demeaned bank size and the unsolicited deposit shock, $(s_i^{t-1} * \Delta d_i^{t*})$.[21] Recall that the identification strategy is premised on the instrument being exogenous to bank demand for deposits. If there is a corresponding increase in local credit demand that is correlated with the instrument, it will bias the average allocation coefficient, β_D, but as long as this bias does not also co-vary with bank size, it will allow me to identify the impact of size heterogeneity, $\beta_{D,s}$.

I estimate the following pooled cross-sectional regressions using 2SLS.

$$\text{First Stage:} \quad \Delta d_i^t = \pi_{0,1} + \pi_{D,1}\Delta d_i^{t*} + \pi_{D\theta,1}(\tilde{s}_i^{t-1} * \Delta d_i^{t*}) + \tau_t + \pi_{\mathbf{1}}'\mathbf{X_i^t} + \epsilon_{i,1}^t \tag{4.3}$$

$$(\tilde{s}_i^{t-1} * \Delta d_i^t) = \pi_{0,2} + \pi_{D,2}\Delta d_i^{t*} + \pi_{D\theta,2}(\tilde{s}_i^{t-1} * \Delta d_i^{t*}) + \tau_t + \pi_{\mathbf{2}}'\mathbf{X_i^t} + \epsilon_{i,2}^t \tag{4.4}$$

$$\text{Second Stage:} \quad \Delta l_i^t = \beta_0 + \beta_D \Delta d_i^t + \beta_{D,s}(\tilde{s}_i^{t-1} * \Delta d_i^t) + \tau_t + \psi'\mathbf{X_i^t} + \varepsilon_i^t \tag{4.5}$$

Variables are the same here as in Equations 4.1 and 4.2. The primary coefficient of interest, $\beta_{D,s}$, estimates how the average allocation varies with size. Year fixed effects, τ_t control for aggregate variation over time. The vector of controls, $\mathbf{X_i^t}$, includes lagged observations of log real assets, s_i^{t-1}, loan share of assets, the Tier 1 capital ratio, an indicator for the parent organization type, and the propensity score. Standard errors are robust to heteroskedasticity and clustered by primary county as outlined in Section 4.1.

Table 11 contains a summary of the estimation results for two time periods, 2003-2006 (Panel A), and 2007-2009 (Panel B). Columns 1 and 2 summarize the first stage of the two-stage estimation. In all four cases the F-stat for tests of instrument exclusion are greater than 10. Note that the first stage allows the prediction to vary with bank size by regressing the actual change in deposits on exposure to the unsolicited shock and the interaction of the shock with size. If large banks receive a greater share of the shale deposit shock on average, the interaction term in Equation 4.3, Column 1, will have a positive coefficient. Therefore the procedure allows for a heterogeneous impact of the deposit shock on banks, mitigating concern that systematic variation in who receives these deposits will bias the results. Note that in B.1, the interaction term has a positive and significant coefficient, suggesting that exposure to the shale gas shock increased deposits more for large banks versus small banks during the crisis period. I can only speculate as

[21] An alternative method is to assume a functional form for the endogenous variable (usually linear) and then to use a control function as a proxy. While the method I use is more robust to model misspecification, the latter can be more precise. Given that the endogeneity is most likely *not* linear, I have chosen the more robust method.

to why this is the case, but depositors may have had greater comfort depositing funds in larger institutions.

In the second stage I estimate the impact of the shock and the interaction term on loans, liquid assets, and non-deposit borrowing. Note that interaction term is zero for the average firm because interacted bank size is demeaned. To put the magnitude of the interaction term in context, average log real assets in thousands of dollars is approximately 11.8 in sample (see Table 7), with a standard deviation of 1. In dollars this equates to \$50m in assets for a bank one s.d. below the average and \$360m for a bank one s.d. above the average. A one standard deviation increase in size increases the allocation towards the left-hand side variable by the estimate of $\beta_{D,s}$. For example, Column B.3 has an interaction coefficient of 0.21. A bank one standard deviation above the average will invest 21% more of the change in deposits in loans relative to the average bank which invests 37%, the estimate of β_D. In general, when I discuss 'larger' banks, I am referring to a one standard deviation change in log assets. Similar to the average allocation coefficients, the sum of the interaction coefficients for loans, liquid assets, less non-deposit borrowing should equal zero – if a bank has 21% more invested in loans than the average bank, they should have that much less allocated to liquid assets or reductions in borrowing.

For the early period, the second stage estimates (Columns A.3-A.6) suggest very little difference in the allocation of loans between small and medium banks. The second-stage coefficients on the interaction term are small, and insignificantly different than zero. Point-wise, large banks invest slightly less of their incremental deposits in loans, -.024, and liquid assets, -.079, and more in reducing other borrowing, -.088. However none of these estimates approach statistical significance at meaningful levels. In the later period, Panel B, I discover that larger banks are more likely to invest incremental deposits in loans, 0.21, and less likely to invest in liquid assets, -0.19. These results are both significant at the 10% level. Moreover, these magnitudes are quite large. This suggests that the observed change in the average loan allocation can be partially attributed to the reduction in lending by the smallest banks.

4.2.1 Types of Loans And Liquid Assets

To better understand the variation in lending behavior between banks, I repeat the 2SLS estimation procedure from Equation 4.5, but with specific categories of loans and types of liquid assets to identify what may be driving the differential response between small- and medium-sized banks. Table 12 reports these additional results. Columns 1-3 consider three types of lending: C&I loans, real estate loans, and small business loans. Small business loans include all C&I and commercial real estate loans for less than $1 million. In the earlier period, larger banks invest slightly less in C&I and real estate loans, but the results are not statistically significant. I find that larger banks invest 29% less in small business loans (Column A.3), this result is significant at the 10% level. Given that large and small banks both invest about 75% of incremental deposits in loans, the smaller proportion allocated to small businesses must be offset in other categories. I corroborate this in unreported results and find larger banks invest significantly more in non-small business real estate loans, particularly residential real estate. The results are consistent with work that has found large banking organization are better able to invest in 'hard' information assets typified by observable information, like credit scores, and less suited to lending to 'soft' information projects, like small businesses (Berger et al. 2005).

The bust period tells a different story, I find that larger banks lend significantly more in total and this is seen across a range of categories including C&I loans, 0.09, real estate loans, 0.13, and small business loans, 0.14. Therefore, it is unlikely that small bank lending is down in response to reduced demand for specific products as they reduce lending in all categories relative to large banks. We can also rule out unique local demand conditions, as small business lending is driven locally (Petersen and Rajan 2002) and medium-sized banks successfully lend to small businesses during the crisis. This suggests that the different allocation decisions between banks is driven by a bank factor, like demand for liquid assets, rather than a differences in lending opportunities.

If small banks are lending significantly less of the deposit shock, what are they investing in? Columns 4-6 repeat the estimation for investments in specific liquid asset classes. In the pre-crisis period large and small banks lent the same, however the larger banks were more likely to reduce borrowing given a deposit shock (Tab. 11 Column A.5) and small banks purchased more in liquid assets, 7.9% (Tab. 11 Column A.4). In Table 12 I find that the bulk of the additional small bank investment in liquid assets was allocated toward securities, 5.9%. This is consistent with the fact that larger banks on average rely on other forms of borrowing (see Table 8); therefore, given a positive liquidity shock they reduce non-deposit borrowing. Smaller banks do not have other debt to reduce, possibly because they are constrained in these markets, and instead invest incremental funds in liquid assets, principally securities.

In the crisis period, larger banks invest less in all three liquid asset classes: cash, securities, and short-term loans.[22] On average a one standard deviation increase in bank size resulted in banks allocating 19% less of their incremental deposit shock towards liquid assets (Tab. 11, Column B.4). More than half of this difference, 13%, is allocated to securities with 3% allocated to cash and 3% to short-term loans. By allocating more toward liquid assets, small banks create a buffer stock that can insure them against future liquidity shocks.

Columns 7 and 8 look at the changes in loan-to-deposit ratios and the return on assets. In the crisis period, I find that both variables are increasing with bank size, but with only marginal significance. Nevertheless, the signs are consistent with my findings that smaller banks are investing in lower return assets (securities) while larger banks are lending. The change in small bank behavior from the boom period may be attributable to a number of factors. These banks may lack demand for their services, they may lack capital, or they may be anticipating the need for liquidity in the future.

4.3 Robustness

In this section I consider additional factors that may influence the results. I focus my robustness checks on the crisis period when bank responses differ. First, areas of the country with more small banks may have been more exposed to the economic downturn. I include fixed effects by field or county in order to limit identification to variation within a geographic area. Second, the coincident shock to local economic activity may be impacting small banks differently than large banks. I proxy for local economic activity and interact this term with bank size to determine if this might explain small bank behavior in 2007-2009. Third, small banks may be capital constrained due to unrecognized losses that have not yet been reflected in their capital ratio. Small banks are known to invest heavily in commercial real estate and losses in this sector have been associated with reduced lending. I consider exposure to this sector to see if it explains small bank retrenchment. Finally, I consider whether lagged bank profitability (ROA) can explain the variation. Robustness checks fail to explain the greater propensity for small banks to invest the deposit shock in liquid assets.

For brevity, I focus robustness tests solely on changes in lending. The empirical methodology is the same as Table 12, but with alternative control variables. Table 13

[22] Short-term loans consists of excess reserves placed with other banks and securities that have been purchased subject to an agreement to resell (repos).

summarizes the results of these robustness checks. The first two tests, Columns 1 and 2, include fixed effects by field and by primary county, respectively. I find that the magnitude of the coefficient on the interaction term between size and deposits, $\beta_{D,s}$, is consistent with earlier results. The field fixed effects specification, 0.21, is significant at the 10% level while the county fixed effect coefficient, 0.20, has a t-stat of 1.5. I conclude from these results that the variation in bank lending exists within geographic areas.

Small banks may be more sensitive to local economic activity. In Column 3 I include a proxy for local business activity, small business growth, and an interaction term with size to allow for small banks to be more sensitive to the proxy. I construct each bank's exposure to small business growth, $\%\Delta smallbusiness_i^t$, by taking the sum of county-level small-business growth weighted by the share of the bank's deposits in the county at $t-1$.[23] While the coefficient on small businesses is positive, 0.29, and significant at the 5% level, the interaction term is small and indistinguishable from zero. I find that this change does not impact the estimate of the average bank allocation to loans, β_D, nor does it impact the estimate of the interaction coefficient, 0.19. It appears that local small business activity does not have a differential impact on lending and does not explain the differences in small and medium-sized banks during this time period.

While I have controlled for Tier 1 capital ratios throughout these analyses, bank capital ratios may not reflect the true strength of the balance sheet as banks postpone write-downs on illiquid assets. Unrecognized write-downs could explain small bank reluctance to invest. Small banks have been shown to invest heavily in commercial real estate during the boom period and these investments performed poorly during the downturn, resulting in reduced lending (Agarwal et al. 2010). In order to account for unrealized losses in commercial real estate, the specification in Column 4 includes the lagged share of assets lent to commercial real estate projects as a control, $\%recom_i^{t-1}$. The share of assets in commercial real estate is weakly positively correlated with lending and has no substantive effect on the coefficients of interest. The final column includes lagged ROA as a control variable. ROA may proxy for the risk of a bank or the bank's skill in investing. I find that ROA is negatively correlated with lending and that ROA does not explain the variation in marginal lending between banks of different size.

[23]This is the same weighting procedure outlined in Section 3.1.2 to construct the deposit shock. Note that the results are robust to several alternative weighting procedures, including equal-weighting for each county in which a bank has branches.

4.3.1 Placebo Test

I also conduct a placebo test of the shale gas instrument. I draw 50 random county-years from the full sample of counties between 2001 and 2009. I label the drawn county-year and the four subsequent years as placebo shale shocks observations, resulting in 204 county-year placebos.[24] I match the counties to peers on propensity score, calculate excess deposit growth, and construct a bank-level estimate of excess deposits as outlined in Section 3.1.2. The random sample has fewer positive excess deposit shocks and those it does have are smaller. In the absence of shale gas development there is not a consistent catalyst for large increases in deposits.

I estimate the impact of the placebo instrument on lending using 2SLS. I also consider a specification that includes the interaction effect between the placebo and bank size. I do not present these results for brevity. The first stage results are significant and correlated with deposits. This is not surprising, as I am linking excess growth in counties to the banks that hold deposits there. However, a test for excluded instruments suggests the placebo instrument is weak (F-stat < 10). Second stage estimates in the early period are somewhat similar in magnitude to my estimates and the OLS estimates; however, interaction terms are statistically indistinguishable from zero. The instrument does not replicate the significant decline in marginal lending observed in the crisis period nor does it replicate the large interaction with bank size in this period. I conclude from these results that the influx of deposits from shale counties are unique from general deposit increases and the results do not mechanically result from the construction of the instrument.

5 Conclusion

Theory has identified size as an important characteristic in the ability of banks to lend (Diamond 1984), the types of loans they make (Stein 2002) and their access to external financing (Stein 1998). In this paper I empirically investigate the role bank size plays in the allocation of an unsolicited deposit shock. I find that prior to the financial crisis, both small- and medium-sized banks allocate these funds predominantly towards loans. During this time period, small banks invest more funds in small business loans, consistent with

[24] The exact match to the number of treatment county-years in my sample is coincidental, as the total number of observations depends on how many in the initial assignment are located in later years that are censored by the end of the sample.

theories emphasizing the role of organizational size in the type of projects they invest in (Stein 2002). However, I determine that small banks are more sensitive to the financial crisis, choosing to forego lending across a wide array of loan categories to invest in liquid assets. This final result appears robust to a wide-array of robustness tests.

Given these results, I conclude that bank size is an important characteristic in determining how banks allocate funds and that the role of bank size is state dependent. The response of small banks during the financial crisis could reflect concern with future access to liquidity or may be the result of greater risk aversion on the part of small bank ownership during this time period. Regardless, these actions can have direct implications on the real side of the economy by reducing capital available to borrowers. The fact that marginal allocation decisions reveal a preference for liquidity implies that small banks in general desired a much more liquid balance during the crisis period. These results can be particularly useful when considering policy interventions meant to ease bank distress and increase lending. While the macroeconomic impact of this behavior in the U.S. is likely small, the results may inform additional work on the role of bank size in downturns or financial crises.

References

Agarwal, S., H. Genay, and R. McMenamin, "Why arent banks lending more? the role of commercial real estate," *Chicago Fed Letter*, 2010, (Dec).

Anna, L.O., R.M. Pollastro, S.B. Gaswirth, M.D. Lewan, P.G. Lillis, L.N.R. Roberts, C.J. Schenk, R..R Charpentier, T.A. Cook, and T.R. Klett, "Assessment of undiscovered oil and gas resources of the Williston Basin Province of North Dakota, Montana, and South Dakota," *US Geological Survey Fact Sheet*, 2008, *3092.*

Berger, Allen N., Anthony Saunders, Joseph M. Scalise, and Gregory F. Udell, "The effects of bank mergers and acquisitions on small business lending," *Journal of financial Economics*, 1998, *50* (2), 187–229.

___ **, Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein**, "Does function follow organizational form? Evidence from the lending practices of large and small banks," *Journal of Financial Economics*, 2005, *76* (2), 237–269.

Bernanke, Ben S. and Mark Gertler, "Agency costs, net worth, and business fluctuations," *The American Economic Review*, 1989, pp. 14–31.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How much should we trust differences-in-differences estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Brunnermeier, Marcus K., "Deciphering the 2007-08 liquidity and credit crunch," *Journal of Economic Perspectives*, 2008, *23* (1), 77–100.

Campello, Murillo, Erasmo Giambona, John R. Graham, and Campbell R. Harvey, "Liquidity management and corporate investment during a financial crisis," *Review of Financial Studies*, 2011, *24* (6), 1944–1979.

Diamond, Douglas W., "Financial intermediation and delegated monitoring," *The Review of Economic Studies*, 1984, *51* (3), 393.

___ **and Raghuram G. Rajan**, "Liquidity shortages and banking crises," *The Journal of Finance*, 2005, *60* (2), pp. 615–647.

Dubiel, R.F., P.D. Warwick, Sharon Swanson, Lauri Burke, L.R.H. Biewick, R.R. Charpentier, J.L. Coleman, T.A. Cook, Kris Denne, Colin Doolan, Catherine Enomoto, P.C. Hackley, A.W. Karlsen, T.R. Klett, S.A. Kinney, M.D. Lewan, Matt Merrill, Krystal Pearson, O.N. Pearson, J.K Pitman, R.M. Pollastro, E.L. Rowan, C.J. Schenck, and Brett Valentine, "Assessment of undiscovered oil and gas resources of the Bend arch–Fort Worth basin province of north-central Texas and southwestern Oklahoma, 2003," *US Geological Survey Fact Sheet*, 2004, *3022* (2).

Ennis, Huberto M. and Alexander L. Wolman, "Large excess reserves in the U.S.: A view from the cross-section of banks," *Working paper*, 2011.

Houston, Joel, Christopher James, and David Marcus, "Capital market frictions and the role of internal capital markets in banking," *Journal of Financial Economics*, 1997, *46* (2), 135 – 164.

Ivashina, Victoria and David Scharfstein, "Bank lending during the financial crisis of 2008," *Journal of Financial Economics*, 2010, *97* (3), 319–338.

Kashyap, Anil K. and Jeremy C. Stein, "What do a million observations on banks say about the transmission of monetary policy?," *American Economic Review*, 2000, pp. 407–428.

Khwaja, Asim I. and Atif Mian, "Tracing the impact of bank liquidity shocks: Evidence from an emerging market," *The American Economic Review*, 2008, pp. 1413–1442.

Paravisini, Daniel, "Local bank financial constraints and firm access to external finance," *The Journal of Finance*, 2008, *63* (5), 2161–2193.

Peek, Joe and Eric Rosengren, "The international transmission of financial shocks: The case of Japan," *The American Economic Review*, 1997, pp. 495–505.

Petersen, Mitchell A., "Estimating standard errors in finance panel data sets: Comparing approaches," *Review of Financial Studies*, 2009, *22* (1), 435–480.

____ and Raghuram G. Rajan, "Does distance still matter? The information revolution in small business lending," *The Journal of Finance*, 2002, *57* (6), 2533–2570.

Pollastro, R.M., R.J. Hill, T.A. Ahlbrandt, R.R. Charpentier, T.A. Cook, T.R. Klett, M.E. Henry, and C.J. Schenk, "Assessment of undiscovered oil and gas resources of the Bend arch–Fort Worth basin province of north-central Texas and southwestern Oklahoma, 2003," *US Geological Survey Fact Sheet*, 2004, *3022* (2).

Rajan, Raghuram G., "Why bank credit policies fluctuate: A theory and some evidence," *The Quarterly Journal of Economics*, 1994, *109* (2), 399.

Rubin, Donald B., "The use of matched sampling and regression adjustment to remove bias in observational studies," *Biometrics*, 1973, pp. 185–203.

—, "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies," *Journal of the American Statistical Association*, 1979, *74* (366), pp. 318–328.

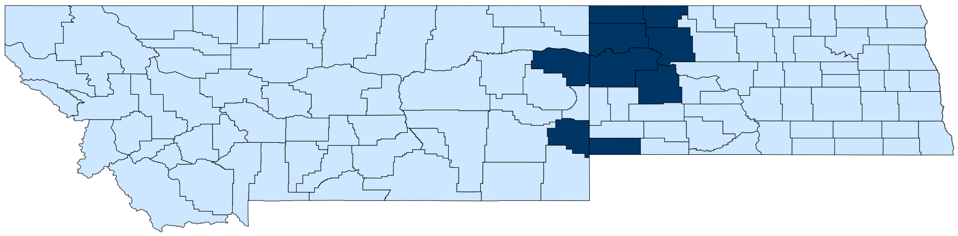
Saunders, Anthony, Elizabeth Strock, and Nickolaos G. Travlos, "Ownership Structure, Deregulation, and Bank Risk Taking," *The Journal of Finance*, 1990, *45* (2), pp. 643–654.

Stein, Jeremy C., "An Adverse-Selection Model of Bank Asset and Liability Management with Implications for the Transmission of Monetary Policy," *RAND Journal of Economics*, 1998, *29* (3), 466–486.

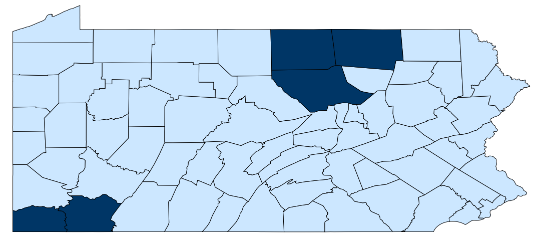
—, "Information production and capital allocation: Decentralized versus hierarchical firms," *The Journal of Finance*, 2002, *57* (5), 1891–1921.

United States. Department of Energy. Office of Fossil Energy and National Energy Technology Laboratory, *Modern shale gas development in the United States: A primer*, US Department of Energy, Office of Fossil Energy, 2009.

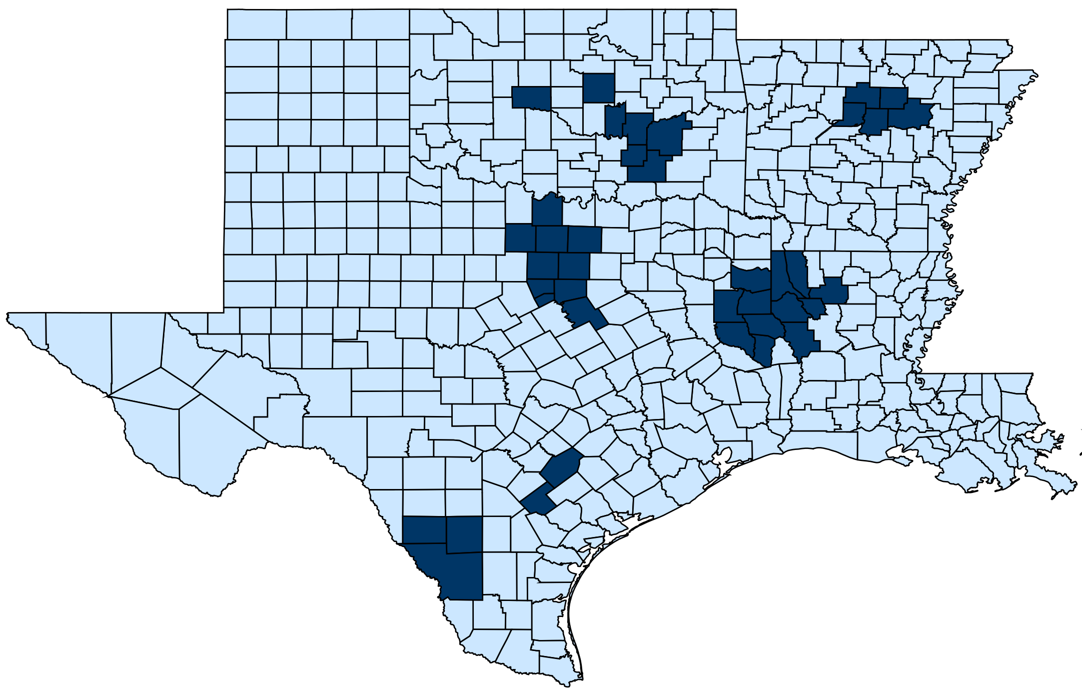
Figure 1: Maps of Treatment Counties by Region



(a) Montana and North Dakota



(b) Pennsylvania



(c) Texas, Oklahoma, Louisiana, Arkansas

Figure 5 illustrates the location of counties I identify as being impacted by recent shale gas development (shaded). I select those counties that experience significant drilling or production development in the first four years after a fields discovery or prior to 2010, whichever comes first. Fig. 2(a) includes counties exposed to the Elm Coulee (1 in MT), Red River(1 in MT, 1 in ND), and Bakken fields (6 in ND). Fig. 2(b) includes the Marcellus field (5). Fig. 2(c) includes the Barnett (10 in TX), Haynesville/Cotton Valley (6 in TX), Fayetteville (5 in AR), Haynesville (6 in LA), Woodford (7 in OK), and Eagle Ford (5 in TX).

Figure 2: County Excess Deposit Growth and Shale Gas Cash-Flow Shock

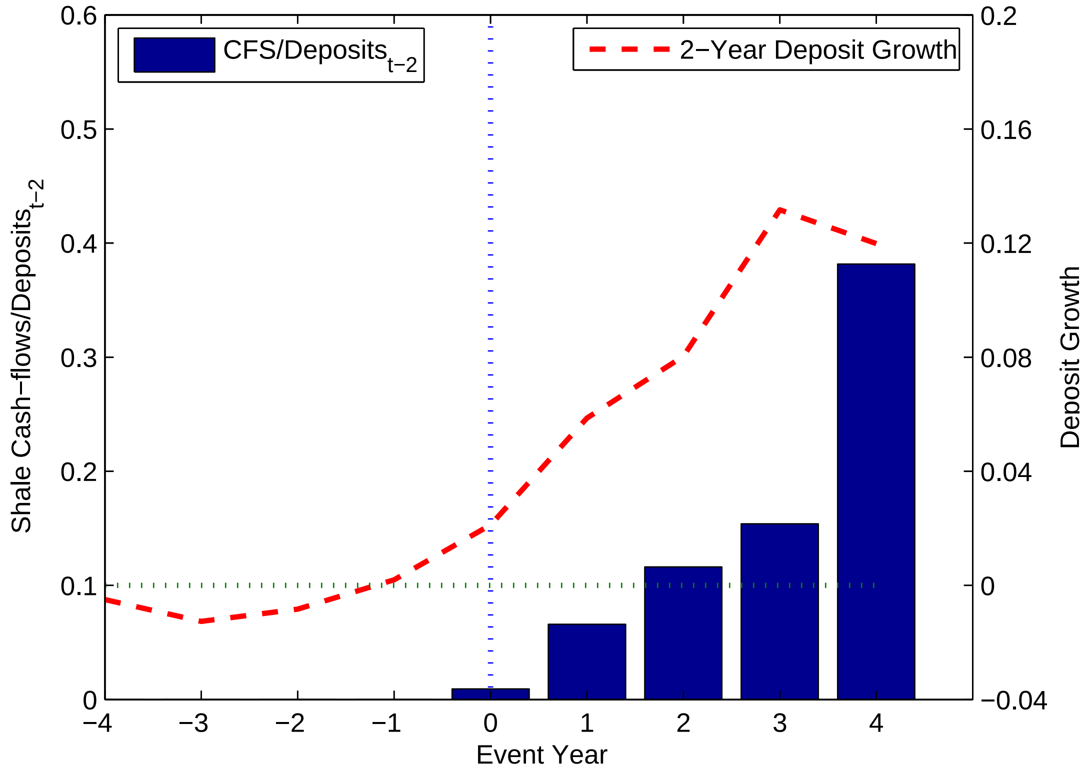


Figure 2 compares two-year county deposit growth to the intervening annual cash-flow shock from shale gas (CFS) scaled by lagged deposits. County level deposit data based on growth in commercial banking deposits from the FDIC's *Summary of Deposits*. Deposit data is as of June 30, therefore two-year deposit growth is used relative to annual gas and oil data that begins and ends during the intervening calendar year. Deposit growth demeaned by year.

Table 1: Summary of Major Unconventional Energy Fields

Field	State	Type	Start Year	Reserves	Counties	Banks	Data Source
Elm Coulee	MT	Oil	2001	3.6 Bbbl	1	2	Montana Board of Oil and Gas Conservation
Red River	MT	Oil	2002	3.6 Bbbl	2	3	Montana Board of Oil and Gas Conservation
Bakken	ND	Oil	2004	3.6 Bbbl	6	10	North Dakota Department of Mineral Resources
Barnett	TX	Gas	2003	26.7 Tcf	10	58	Railroad Commission of Texas
Fayetteville	AR	Gas	2004	4.9 Tcf	5	11	Arkansas Oil and Gas Commission
Haynesville/Cotton Valley	TX	Gas	2004	20 Tcf	6	19	Railroad Commission of Texas
Haynesville	LA	Gas	2007	60 Tcf	6	11	Louisiana Deparetment of Natural Resources: SONRIS
Woodford	OK	Gas	2006	24 Tcf	7	27	Oklahoma Corporation Commision: Oil and Gas Division
Marcellus	PA	Gas	2007	81 Tcf	5	12	Pennsylvania Department of Natural Resources: Oil and Gas
Eagle Ford	TX	Gas	2008	50 Tcf	5	12	Railroad Commission of Texas

Table 1 lists the shale gas/oil fields used to identify unsolicited deposit shocks. *Type* indicates the predominant resource, oil or natural gas, though most fields produce both. *Start Year* is the first major year of development or the first public announcement of a resource find, whichever is earliest. *Reserves* is the total technically recoverable resource from the underlying formation unless otherwise noted. Some formations are quite large and may encompass multiple fields. The top three oil fields are one example as they tap the same underlying reserve base across a wide geographic area. Reserves are sourced from USGS Fact Sheets except in the case of Fayetteville and Haynesville/Cotton Valley, which are reported by Southwestern Energy. The reserves listed for these two fields reflect what is available in that area specifically, rather than the entire formation. Tcf is trillion cubic feet of natural gas, Bbbl is billion barrels of oil. *Counties* contains the number of counties that experienced significant development during the first four years after the start year or by 2010, whichever comes first. Banks are the number of banks impacted in these counties. The final column lists data sources for drilling permits and oil/gas output.

Table 2: Summary of County-Year Statistics 2001-2009

	Treatment				**Matched Untreated**				**Full Sample Untreated**			
	N	Mean	Median	σ	N	Mean	Median	σ	N	Mean	Median	σ
Deposits ($ 000's)	204	1221.0	329.1	5966.0	886	1137.0	291.9	5472.0	27,509	1737.0	315.7	9686.0
Employment (000's)	204	22.5	5.6	77.9	886	25.0	5.4	110.9	27,508	34.9	6.5	128.1
Pop. Density (000's/sq. mile)	204	88.4	33.5	239.0	886	98.9	29.5	464.4	27,468	248.5	44.4	1721.0
2-Year Deposit Growth	204	16.9%	15.8%	12.5%	886	10.0%	9.0%	11.0%	27,509	8.7%	7.4%	10.8%
Employment Growth	203	3.0%	2.9%	6.5%	877	1.1%	1.2%	6.4%	27,270	-0.5%	-0.2%	6.3%
Non-Energy Emp. Growth	203	2.5%	2.0%	7.3%	871	0.9%	1.0%	6.5%	27,151	-0.5%	-0.3%	6.4%
Wage Growth	204	8.7%	7.7%	9.8%	878	5.7%	5.5%	8.4%	27,247	2.9%	3.1%	7.3%
Small Bus. Growth	202	1.7%	1.2%	3.5%	878	0.2%	0.2%	3.8%	27,252	0.2%	0.0%	3.8%
CFS/D^{t-1}	204	14.4%	3.6%	30.9%	886	0.0%	0.0%	0.0%	27,509	0.0%	0.0%	0.0%
Employment Share by Industry												
Agriculture	204	1.3%	0.4%	2.3%	886	1.3%	0.2%	2.7%	27,385	1.8%	0.0%	5.0%
Construction	204	6.6%	6.1%	3.8%	886	6.4%	5.7%	4.6%	27,385	6.0%	5.4%	4.4%
Manufacturing/Utilities	204	13.0%	13.4%	9.6%	886	13.3%	11.2%	11.1%	27,385	17.6%	15.4%	13.7%
Wholesale & Retail Trade	204	20.3%	20.4%	4.7%	886	20.8%	20.7%	6.3%	27,385	19.9%	19.9%	6.2%
Tranportation/Warehousing	204	5.3%	1.7%	7.6%	886	4.1%	0.2%	8.1%	27,385	1.2%	0.0%	4.5%
Business Services	204	3.9%	3.5%	3.7%	886	3.8%	2.9%	5.0%	27,385	2.5%	1.7%	3.5%
Finance/Real Estate	204	6.4%	5.2%	4.6%	886	6.3%	5.2%	5.1%	27,385	7.1%	5.5%	6.2%
Education	204	4.9%	4.6%	2.6%	886	4.9%	5.0%	2.8%	27,385	4.4%	4.3%	2.8%
Healthcare	204	0.2%	0.0%	0.4%	886	0.2%	0.0%	0.7%	27,385	0.6%	0.0%	1.2%
Misc. Services	204	8.5%	7.8%	9.7%	886	8.7%	8.3%	9.0%	27,385	8.4%	8.4%	8.5%
Energy	204	9.8%	11.8%	6.7%	886	10.1%	11.3%	6.8%	27,385	11.7%	12.4%	7.7%

Table 2 summarizes the county characteristics for the shale gas counties (treatment), a matched sample of counties excluding treatment observations and the full sample of counties excluding treatment observations. Deposit growth is two year deposit growth (June-June). All other growth rates are calendar year growth rates. Treatment county-years are limited to the first four years of development in a shale-gas county. The matched sample is constructed by matching to the five nearest counties using propensity score matching. p-scores are calculated using a logit of a treatment indicator on demographic, geographic, and economic factors. Growth rates are trimmed at the 1% level to limit the impact of extreme outliers. Statistics in the matched sample are weighted, as some county-years actually appear more than once.

Table 3: Summary of County Cash-Flow Shock and Matched Deposit Growth

			CFS_j^t (\$ '000s)		CFS_j^t/D_j^{t-1}		$\%\Delta D_j^{t*}$	
Year	Counties	Banks	Mean	Max	Mean	Max	Mean	Max
2001	1	1	1,864	1,864	1.4%	1.4%	-4.5%	-4.5%
2002	3	4	4,092	4,296	4.7%	6.3%	-7.3%	-1.4%
2003	8	48	14,652	34,687	6.4%	20.7%	2.3%	9.5%
2004	31	107	11,806	63,075	5.0%	43.7%	-0.9%	22.2%
2005	35	118	27,667	176,098	10.5%	118.3%	5.7%	26.4%
2006	36	124	32,658	218,149	10.2%	156.2%	9.7%	60.8%
2007	39	132	46,816	413,174	10.5%	79.1%	8.9%	63.4%
2008	37	98	62,958	303,021	29.5%	207.1%	9.3%	44.9%
2009	16	42	103,525	355,416	30.2%	107.6%	12.9%	49.3%

Table 3 summarizes data on the estimated cash-flow shock and excess deposit growth by year. CFS_j^t is an estimate of royalty and leasing bonuses paid to local landowners in thousands of dollars. CFS/D^{t-1} is the ratio of these payments to deposits as of June 30th in the prior year. $\%\Delta D_j^{t*}$ is the excess deposit growth in treatment counties relative to a matched sample of similar counties. The matched sample matches treatment counties to five other counties based on nearest neighbor propensity score matching and weights these matches at 1/5th the value of a shale gas county.

Table 4: Regression of Deposit Growth on Cash-Flow Shock

Table 4 contains estimates of the pooled cross-sectional regression of deposit growth on cash-flows from shale gas and various controls.

$$\%\Delta D_j^t = \beta(CFS_j^t/D_j^{t-1}) + \psi Controls_j^t + \varepsilon_j^t$$

Deposits are observed in June, hence two-year deposit growth is used to capture the impact of variables in the intervening calendar year. All other variables reflect changes during the calendar year. The cash-flow shock, CFS_j^t/D_j^{t-1} is the sum of estimated royalties and leasing bonuses paid to landowners in shale gas counties scaled by lagged deposits. $\%\Delta Wages$ is the percent change in private wages paid in the county. $\%\Delta SmallBusiness$ is the percent change in the number of small businesses ($<$ 50 employees). The full sample includes all counties. The matched sample matches treatment counties to five other counties based on nearest neighbor propensity score matching and weights each of these matches at 1/5th the value of a shale gas county. Suppressed controls include log deposits, county-level demographic data, and the propensity score (in the case of the matched sample). Standard errors reported in parentheses are clustered by state. t-stats reported in parentheses. ***, **, *indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	**Full Sample**					
	‘01-‘09	‘01-‘09	‘01-‘09	‘01-‘09	‘01-‘06	‘07-‘09
Dependent Var.:	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$
CFS/D_j^{t-2}	0.22***	0.24***	0.24***	0.22***	0.19***	0.24***
	(0.023)	(0.025)	(0.027)	(0.027)	(0.018)	(0.035)
$\%\Delta Wages$				0.061***		
				(0.018)		
$\%\Delta SmallBusiness$				0.27***		
				(0.054)		
Controls		+	+	+	+	+
Year FE			+	+	+	+
Observations	27,749	27,686	27,686	27,684	18,451	9,235
R-squared	0.006	0.058	0.080	0.096	0.125	0.052
Panel B	**Matched Sample**					
	‘01-‘09	‘01-‘09	‘01-‘09	‘01-‘09	‘01-‘06	‘07-‘09
Dependent Var.:	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$	$\%\Delta D_j^t$
CFS/D_j^{t-2}	0.24***	0.25***	0.25***	0.24***	0.19***	0.26***
	(0.022)	(0.029)	(0.032)	(0.033)	(0.023)	(0.033)
$\%\Delta Wages$				0.059		
				(0.044)		
$\%\Delta SmallBusiness$				0.18*		
				(0.10)		
Controls		+	+	+	+	+
Year FE			+	+	+	+
Observations	1,090	1,090	1,090	1,090	595	495
R-squared	0.197	0.272	0.324	0.313	0.334	0.405

Table 5: Summary of Excess Deposit Growth Relative Shale Discovery

Event Time	Counties	$\%\Delta D^*$	$\sigma(\%\Delta D^*)$
-3	50	-0.6%	8.3%
-2	51	1.0%	8.9%
-1	53	0.0%	9.8%
0	53	1.3%	10.6%
1	52	6.5%	12.0%
2	47	9.3%	9.5%
3	37	13.9%	17.0%
4	37	11.9%	15.3%
5	36	6.8%	12.5%
6	12	1.4%	16.7%
7	3	5.4%	7.6%
8	1	7.8%	–

Table 5 displays excess deposit growth in event time where 0 is the first year of development. $\%\Delta D*$ is the excess two-year deposit growth relative to a matched sample of similar counties. $\sigma(\%\Delta D*)$ is the standard deviation. Excess deposit growth does not appear to reverse, suggesting the shocks to deposits are relatively persistent.

Table 6: Summary of Bank Deposit Shock by Size

				Δd_i^{t*}	
Percentile	Obs.	Assets ($mm)	Shock Counties	Mean	Max
10	31	16.9	1.10	9.3%	55.2%
20	30	32.4	1.10	10.8%	43.2%
30	35	45.1	1.00	11.6%	51.7%
40	40	60.7	1.03	8.9%	55.8%
50	39	80.4	1.28	9.6%	54.7%
60	51	105.8	1.47	7.6%	49.3%
70	38	147.8	1.26	6.4%	51.4%
80	62	200.1	1.63	5.9%	49.5%
90	63	318.0	1.59	4.1%	41.4%
99	68	1,095.0	2.38	1.2%	6.3%
100	20	60,910.0	2.35	0.2%	1.6%

Table 6 summarizes data on the unsolicited deposit shock by bank size percentile. The results demonstrate that a wide array of bank sizes are exposed to the shock. Percentiles are determined using real asset values. Observations are bank-year observations in the size percentile. Assets are in millions and are in real terms. Shock Counties is the average number of shale counties the banks have branches in. Δd_i^{t*} is estimated bank exposure to shale gas deposit shocks. The sum of excess deposit growth in shale gas counties, weighted by the banks deposits in the county at $t-1$, scaled by bank assets at $t-1$.

Table 7: Summary of Bank-Year Statistics 2001-2009

	Treatment				**Matched (Untreated)**				**Full Sample (Untreated)**			
	N	Mean	Median	σ	N	Mean	Median	σ	N	Mean	Median	σ
Asset Composition:												
Cash	438	6.1%	4.7%	4.5%	1,585	6.2%	4.3%	6.6%	46,373	5.0%	3.8%	4.6%
Securities, FFS, & Repos	438	27.9%	26.9%	14.6%	1,585	27.6%	25.4%	15.0%	46,311	28.8%	26.6%	15.2%
Total Loans	438	59.3%	61.0%	15.1%	1,585	59.6%	61.8%	15.5%	46,308	61.5%	63.4%	15.3%
C&I Loans	438	11.0%	9.4%	6.7%	1,585	11.1%	9.5%	7.2%	46,404	9.9%	8.2%	7.6%
Real Estate Loans	438	36.3%	35.1%	14.1%	1,585	37.2%	37.4%	14.8%	46,404	39.5%	39.9%	16.2%
Loans to Individuals	438	7.6%	6.2%	4.7%	1,585	7.2%	5.7%	6.4%	46,404	6.2%	4.9%	6.0%
Small Business Loans	438	19.6%	19.2%	9.2%	1,585	20.1%	18.4%	9.9%	46,378	17.7%	16.1%	10.3%
Liabilities Composition:												
Deposits	438	84.1%	85.3%	6.2%	1,585	84.0%	85.0%	5.9%	46,404	82.7%	84.2%	7.4%
Other Borrowings	438	3.8%	1.5%	5.4%	1,585	3.8%	1.9%	4.9%	46,281	5.0%	2.8%	6.3%
Equity	438	10.7%	9.4%	4.7%	1,585	10.8%	9.6%	4.9%	46,296	11.0%	9.9%	6.9%
Other Stats:												
Assets($mm)	438	223.0	132.8	246.3	1,585	230.4	145.9	237.0	46,404	161.2	92.6	192.9
Log(Real Assets)	438	11.69	11.69	1.04	1,585	11.78	11.78	1.00	46,387	11.38	11.35	1.03
ROA	425	1.1%	1.0%	0.8%	1,585	1.1%	1.0%	1.2%	44,756	0.9%	0.9%	1.1%
Tier 1 Capital Ratio	438	13.9%	11.9%	8.0%	1,585	14.1%	12.6%	7.5%	46,360	17.4%	14.2%	14.7%
Loan/Deposit Ratio	438	70.8%	72.1%	18.5%	1,585	71.2%	73.6%	19.1%	46,308	80.7%	75.9%	1256.0%
Δd_i^t	438	20.0%	16.7%	20.0%	1,585	12.8%	9.6%	16.9%	47,078	16.0%	9.4%	29.0%
Δd_i^{t*}	438	7.0%	4.3%	9.7%	1,585	0.0%	0.0%	0.0%	47,078	0.1%	0.0%	1.2%
% BHC	438	85.8%	100.0%	34.9%	1,585	84.7%	100.0%	36.1%	46,404	75.2%	100.0%	43.2%
p-score	417	8.8%	6.9%	7.8%	1,585	8.7%	6.8%	7.8%				

Table 7 summarizes the full and matched sample of bank-year observations. Excludes banks larger than $1.2bn in assets, de novos, acquirors/targets, credit card banks, real-estate banks, or non-deposit financed banks. Log Real Assets is the log of bank real assets in thousands of dollars. Asset and Liabilities Composition are calculated at $t-1$ by dividing by total assets. Securities, FFS, & Repos is the sum of security holdings, federal funds sold (FFS) and securities purchased with an agreement to resell. % BHC is the percent of banks that are bank holding companies. Δd_i^t is the dollar change in deposits from $t-1$ to $t+1$ divided by assets at $t-1$. Δd_i^{t*} is the weighted average of excess deposit growth in shale gas counties scaled by deposit composition. The matched sample matches treatment counties to five nearest neighbor banks using propensity score matching. The matched sample only contains observations from 2003-2009 as those are the only years with positive deposit shocks. p-scores calculated using year fixed-effects, region fixed effects, bank characteristics, location demographics, and dummy variables for type of high-holder.

Table 8: Bank-Year Statistics 2003-2009 By Size – Matched Sample

	2003-2006				**2007-2009**			
	Small		Medium		Small		Medium	
	Mean	σ	Mean	σ	Mean	σ	Mean	σ
Observations	454		706		351		491	
Asset Composition:								
Cash	8.4%	7.6%	5.4%	4.0%	6.7%	7.3%	4.8%	3.2%
Securities, FFS, & Repos	28.9%	15.3%	26.2%	13.9%	32.2%	16.8%	25.9%	13.5%
Total Loans	56.6%	15.3%	62.2%	14.3%	54.3%	16.8%	61.6%	14.4%
C&I Loans	11.1%	7.2%	11.6%	6.9%	9.9%	6.8%	10.8%	6.3%
Real Estate Loans	31.1%	13.8%	39.6%	13.5%	31.5%	14.7%	41.1%	13.5%
Loans to Individuals	8.9%	5.1%	7.4%	6.4%	7.6%	4.8%	6.1%	5.0%
Small Business Loans	19.9%	10.6%	20.8%	9.2%	18.4%	10.3%	19.1%	8.3%
Liabilities Composition:								
Deposits	85.4%	5.7%	84.2%	5.7%	84.3%	6.0%	82.6%	6.6%
Other Borrowings	1.8%	3.3%	4.7%	5.2%	2.4%	3.9%	5.0%	6.0%
Equity	11.8%	4.9%	9.6%	4.1%	12.1%	5.3%	10.5%	4.8%
Other Stats:								
Assets($mm)	53.3	23.6	314.2	247.1	61.8	28.9	364.8	265.9
Log(Real Assets)	10.67	0.55	12.34	0.65	10.75	0.59	12.45	0.63
ROA	1.0%	0.8%	1.2%	0.6%	1.1%	1.0%	1.2%	1.5%
Tier 1 Capital Ratio	15.7%	8.5%	13.2%	6.3%	17.0%	10.8%	11.9%	4.9%
Loan/Deposit Ratio	66.5%	18.0%	74.2%	18.1%	64.5%	19.8%	74.8%	17.8%
Δd_i^t	15.5%	21.6%	15.8%	16.2%	19.3%	22.8%	15.3%	14.2%
Δd_i^{t*}	5.4%	10.6%	2.7%	7.2%	4.6%	7.7%	2.2%	3.9%
% BHC	70.7%	45.6%	93.4%	24.8%	76.7%	42.4%	92.2%	26.9%
p-score	6.2%	5.9%	10.4%	8.0%	5.7%	5.3%	10.5%	8.9%

Table 8 compares small banks to larger banks in the matched sample. Small banks are defined as banks with below the median in asset values ($< \$80m$ in assets). Asset and Liabilities Asset and Liabilities Composition are calculated at $t-1$ by dividing by total assets. Log Real Assets is the log of bank real assets in thousands of dollars. Securities, FFS, & Repos is the sum of security holdings, federal funds sold (FFS) and securities purchased with an agreement to resell. % BHC is the percent of banks that are bank holding companies. Δd_i^t is the dollar change in deposits from $t-1$ to $t+1$ divided by assets at $t-1$. Δd_i^{t*} is the weighted average of excess deposit growth in shale gas counties scaled by deposit composition. The matched sample matches treatment counties to five nearest neighbor banks using propensity score matching. p-scores calculated using year fixed-effects, region fixed effects, bank characteristics, location demographics, and dummy variables for type of high-holder.

Table 9: OLS & 2SLS: Allocation of Deposit Growth into Loans/Liquid Assets

Table 9 contains estimates of the impact of changes in deposits, Δd_i^t on changes in bank balance sheet accounts, Δl_i^t. Both changes are scaled by assets at time $t-1$.

$$\Delta l_i^t = \beta_0 + \beta_D \Delta d_i^t + \psi' \mathbf{X_i^{t-1}} + \varepsilon_i^t$$

Panel A contains OLS estimates. Panel B contains 2SLS estimates using the exposure to deposits in shale counties as an instrument, Δd_i^{t*}, where Column 1 includes the results from the first stage regression. F-stat for test of instrument exclusion equal to 62.3. *loans* denotes total loans, *liquid* is cash, securities and short-term loans, *borrow* denotes non-deposit borrowings from the liability side of the bank's balance sheet including repurchase agreements, *c&i* is commercial and industrial loans, and *re* real estate loans. Suppressed controls include year fixed effects and lagged observations of log real assets, loan share of assets, the Tier 1 capital ratio, an indicator for the parent organization type, and the propensity score. Standard errors reported in parentheses are clustered by primary county. ***, **, * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A		**OLS**				
Dependent Var.:		$\Delta loans_i^t$	$\Delta liquid_i^t$	$\Delta borrow_i^t$	$\Delta c\&i_i^t$	Δre_i^t
Δd_i^t		0.72***	0.35***	0.024*	0.15***	0.50***
		(0.047)	(0.039)	(0.013)	(0.013)	(0.053)
Observations		2,002	2,002	1,997	2,002	2,002
R-squared		0.687	0.377	0.078	0.336	0.585
Panel B	**First**	**Second Stage**				
Dependent Var.:	Δd_i^t	$\Delta loans_i^t$	$\Delta liquid_i^t$	$\Delta borrow_i^t$	$\Delta c\&i_i^t$	Δre_i^t
Δd_i^{t*}	0.56***					
	(0.070)					
Δd_i^t		0.56***	0.49***	-0.043*	0.097***	0.46***
		(0.099)	(0.085)	(0.025)	(0.036)	(0.098)
Observations	2,002	2,002	2,002	1,997	2,002	2,002
R-squared	0.128					

Table 10: OLS & 2SLS: Sub-Period Allocation of Deposit Growth into Loans, 2003-2006 and 2007-2009

Table 10 contains estimates of the impact of changes in deposits, Δd_i^t on changes in balance sheet accounts, Δl_i^t, for two separate time periods. Both changes are scaled by assets at time $t-1$.

$$\Delta l_i^t = \beta_0 + \beta_D \Delta d_i^t + \psi' \mathbf{X_i^{t-1}} + \varepsilon_i^t$$

Both changes are scaled by assets at time $t-1$. Columns 1 and 4 contain OLS estimates. Columns 2 and 5 contain the first stage estimates of changes in deposits using the exposure to deposits in shale counties as an instrument, Δd_i^{t*}. F-stats for tests of instrument exclusion exceed 10 in both cases. Columns 3 and 6 contain the second stage estimates. *loans* denotes total loans. Suppressed controls include year fixed effects and lagged observations of log real assets, loan share of assets, the Tier 1 capital ratio, an indicator for the parent organization type, and the propensity score. Standard errors reported in parentheses are clustered by primary county. ***, **, * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	2003-2006			**2007-2009**		
	OLS	**First**	**Second**	**OLS**	**First**	**Second**
Dependent Var.:	$\Delta loans_i^t$	Δd_i^t	$\Delta loans_i^t$	$\Delta loans_i^t$	Δd_i^t	$\Delta loans_i^t$
Δd_i^{t*}		0.45***			0.90***	
		(0.046)			(0.13)	
Δd_i^t	0.72***		0.78***	0.73***		0.27***
	(0.035)		(0.075)	(0.073)		(0.10)
Observations	1,160	1,160	1,160	842	842	842
R-squared	0.721	0.152		0.656	0.131	

Table 11: 2SLS: The Impact of Size on Asset Allocation, 2003-2006 and 2007-2009

Table 11 contains 2SLS estimates of the impact of changes in deposits, Δd_i^t and the interaction with size ($\tilde{s}_i^{t-1} * \Delta d_i^t$) on changes in balance sheet accounts, Δl_i^t.

$$\Delta l_i^t = \beta_0 + \beta_D \Delta d_i^t + \beta_{Ds}(\tilde{s}_i^{t-1} * \Delta d_i^t) + \beta_s s_i^{t-1} + \tau_t + \psi' \mathbf{X_i^t} + \varepsilon_i^t$$

I instrument for both the change in deposits, Δd_i^t, and the interaction term between deposits and demeaned bank size, ($\tilde{s}_i^{t-1} * \Delta d_i^t$), using the unsolicited deposit shock, Δd_i^{t*}, and an interaction between demeaned bank size and the unsolicited deposit shock, Δd_i^{t*}. Columns 1 and 2 contain the results from the first stage regressions. F-stats for tests of instrument exclusion exceed 10 in all four cases. Columns 3-5 contain the second stage estimates. Panel A contains estimates for the '03-'06 period, Panel B the '07-'09 period. *loans* denotes total loans, *liquid* is cash, securities and short-term loans, *borrow* denotes non-deposit borrowings from the liability side of the bank's balance sheet including repurchase agreements. Suppressed controls include year fixed effects and lagged observations of log real assets, loan share of assets, the Tier 1 capital ratio, an indicator for the parent organization type, and the propensity score. Standard errors reported in parentheses are clustered by primary county. ***, **, * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)
Panel A: 2003-2006	**First Stage**		**Second Stage**		
Dependent Var.:	Δd_i^t	$(\tilde{s}_i^{t-1} * \Delta d_i^t)$	$\Delta loans_i^t$	$\Delta liquid_i^t$	$\Delta borrow_i^t$
Δd_i^{t*}	0.43***	-0.093			
	(0.072)	(0.095)			
$(\tilde{s}_i^{t-1} * \Delta d_i^{t*})$	-0.059	0.40***			
	(0.12)	(0.13)			
Δd_i^t			0.77***	0.28***	-0.075
			(0.17)	(0.097)	(0.068)
$(\tilde{s}_i^{t-1} * \Delta d_i^t)$			-0.024	-0.079	-0.088
			(0.19)	(0.12)	(0.063)
Observations	1,160	1,160	1,160	1,160	1,156
R-squared	0.152	0.423			
Panel B: 2007-2009	**First Stage**		**Second Stage**		
Dependent Var.:	Δd_i^t	$(\tilde{s}_i^{t-1} * \Delta d_i^t)$	$\Delta loans_i^t$	$\Delta liquid_i^t$	$\Delta borrow_i^t$
Δd_i^{t*}	1.11***	0.083			
	(0.14)	(0.15)			
$(\tilde{s}_i^{t-1} * \Delta d_i^{t*})$	0.33**	0.85***			
	(0.13)	(0.20)			
Δd_i^t			0.37***	0.58***	-0.052
			(0.12)	(0.13)	(0.051)
$(\tilde{s}_i^{t-1} * \Delta d_i^t)$			0.21*	-0.19*	0.015
			(0.11)	(0.10)	(0.046)
Observations	842	842	842	842	841
R-squared	0.138	0.527			

Table 12: 2SLS: The Impact of Size on Types of Loans and Liquid Assets, 2003-2006 and 2007-2009

Table 12 contains 2SLS estimates of the impact of changes in deposits, Δd_i^t and the interaction with size $(\tilde{s}_i^{t-1} * \Delta d_i^t)$ on changes in balance sheet accounts, Δl_i^t.

$$\Delta l_i^t = \beta_0 + \beta_D \Delta d_i^t + \beta_{Ds}(\tilde{s}_i^{t-1} * \Delta d_i^t) + \beta_s s_i^{t-1} + \tau_t + \psi' \mathbf{X_i^t} + \varepsilon_i^t$$

I instrument for both the change in deposits, Δd_i^t, and the interaction term between deposits and demeaned bank size, $(\tilde{s}_i^{t-1} * \Delta d_i^t)$, using the unsolicited deposit shock, Δd_i^{t*}, and an interaction between demeaned bank size and the unsolicited deposit shock, Δd_i^{t*}. F-stats for tests of instrument exclusion exceed 10 in all four cases. Panel A contains estimates for the '03-'06 period, Panel B the '07-'09 period. *c&i* denotes commercial and industrial loans, *re* real estate loans, *sbl* small business loans, *cash* cash and balances due, *sec* securities, *ffs&repos* federal funds sold and short-term agreements to resell securities, *ROA* return-on-assets, and *loan/dep* the loan-deposit ratio. Suppressed controls include year fixed effects and lagged observations of log real assets, loan share of assets, the Tier 1 capital ratio, an indicator for the parent organization type, and the propensity score. Standard errors reported in parentheses are clustered by primary county. ***, **, * indicate statistical significance at 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A	**Second Stage – 2003-2006**							
Dependent Var.:	Δci_i^t	Δre_i^t	Δsbl_i^t	$\Delta cash_i^t$	Δsec_i^t	$\Delta ffs\&repos_i^t$	ΔROA_i^t	$\Delta loan/dep_i^t$
Δd_i^{t*}	0.080	0.64***	0.14	0.078*	0.071	0.17**	-0.0024	-0.036
	(0.081)	(0.13)	(0.12)	(0.044)	(0.11)	(0.082)	(0.0050)	(0.13)
$(\tilde{s}_i^{t-1} * \Delta d_i^{t*})$	-0.10	-0.025	-0.29*	0.010	-0.059	-0.016	-0.0031	0.053
	(0.10)	(0.16)	(0.15)	(0.047)	(0.14)	(0.12)	(0.0075)	(0.15)
Observations	1,160	1,160	1,160	1,147	1,158	1,160	1,160	1,159
Panel B	**Second Stage – 2007-2009**							
Dependent Var.:	Δci_i^t	Δre_i^t	Δsbl_i^t	$\Delta cash_i^t$	Δsec_i^t	$\Delta ffs\&repos_i^t$	ΔROA_i^t	$\Delta loan/dep_i^t$
Δd_i^t	0.073**	0.28**	0.098	0.072**	0.50***	0.0036	0.014**	-0.24**
	(0.033)	(0.12)	(0.067)	(0.033)	(0.098)	(0.060)	(0.0067)	(0.10)
$(\tilde{s}_i^{t-1} * \Delta d_i^t)$	0.091**	0.13	0.14**	-0.029	-0.13	-0.033	0.0046	0.086
	(0.038)	(0.086)	(0.065)	(0.084)	(0.11)	(0.063)	(0.011)	(0.12)
Observations	842	842	842	839	842	842	842	842

Table 13: 2SLS Robustness Tests: Impact of Size on Allocation to Lending, 2007-2009

Table 12 contains 2SLS estimates of the impact of changes in deposits, Δd_i^t and the interaction with size $(\tilde{s}_i^{t-1} * \Delta d_i^t)$ on changes in loans, Δl_i^t.

$$\Delta l_i^t = \beta_0 + \beta_D \Delta d_i^t + \beta_{Ds}(\tilde{s}_i^{t-1} * \Delta d_i^t) + \beta_s s_i^{t-1} + \tau_t + \psi' \mathbf{X_i^t} + \varepsilon_i^t$$

I instrument for both the change in deposits, Δd_i^t, and the interaction term between deposits and demeaned bank size, $(\tilde{s}_i^{t-1} * \Delta d_i^t)$, using the unsolicited deposit shock, Δd_i^{t*}, and an interaction between demeaned bank size and the unsolicited deposit shock, Δd_i^{t*}. F-stats for tests of instrument exclusion exceed 10 in each case. Column 1 includes fixed effects by gas/oil field, including a fixed effect for banks unexposed to a field. Column 2 includes fixed effects by primary county. Column 3-6 include year fixed effects. Column 3 controls for the bank's exposure to local economic activity, $\%\Delta smallbusiness_i^t$, and an interaction of local activity with size. Column 4 includes the share of bank assets invested in commercial real estate loans, $\%recom_i^{t-1}$. Column 5 includes return on assets, ROA_i^{t-1}. Suppressed controls in each specification include lagged observations of log real assets, loan share of assets, the Tier 1 capital ratio, an indicator for the parent organization type, and the propensity score. Standard errors reported in parentheses are clustered by primary county. ***, **, * indicate statistical significance at 1%, 5%, and 10%, respectively.

	Second Stage – 2007-2009				
	(1)	(2)	(3)	(4)	(5)
	$\Delta loans_i^t$	$\Delta loans_i^t$	$\Delta loans_i^t$	$\Delta loans_i^t$	$\Delta loans_i^t$
Δd_i^t	0.27	0.15	0.36***	0.39***	0.38***
	(0.18)	(0.23)	(0.13)	(0.12)	(0.12)
$(\tilde{s}_i^{t-1} * \Delta d_i^t)$	0.20**	0.21	0.19*	0.21*	0.19*
	(0.097)	(0.14)	(0.11)	(0.11)	(0.11)
$\%\Delta smallbusiness_i^t$			0.29*		
			(0.15)		
$(\tilde{s}_i^t * \%\Delta smallbusiness_i^t)$			0.024		
			(0.12)		
$\%recom_i^{t-1}$				0.094	
				(0.093)	
ROA_i^{t-1}					-0.98
					(0.68)
Fixed Effects	Field	County	Year	Year	Year
Observations	842	841	842	842	842

A Variable Descriptions and Timing Conventions

Reference for key variables and their construction. Note on timing conventions: Local deposits and bank balance sheets are observed as of June 30th of each year. Bank data is calculated over two-year intervals in order to compare changes in bank balance sheets to calendar year variables. t references the calendar year. For example, the year 2007 refers to an observation of shale gas cash-flows in 2007 (t), but a two year change in deposits is measured from June 30, 2006 to June 30, 2008 ($t-1$ to $t+1$) in order to encompass the calendar year.

County Variables: County j at time t
Primary data source include the FDIC Summary of Deposits, Census population data, the Quarterly Census of Employment and Wages, the County Business Patterns database, and permit and oil/gas output data from Table 1.

CFS_j^t – Estimated dollar value of cash-flow shock from shale gas development. Comprised of royalties to local landowners, estimated using oil/gas output, and bonus payments estimated using drilling permits issued. For analysis purposes, this variable is scaled by the level of deposits at $t-1$.

$\%\Delta D_j^t$ – Percent change in county deposits from $t-1$ to $t+1$. Deposits as of June 30th each year. Deposit balances are meant to encapsulate a calendar year. 1% of observations trimmed to remove extreme outliers.

$\%\Delta D_j^{t*}$ – Excess percent change in shale gas county deposits from $t-1$ to $t+1$. Deposits as of June 30th each year. Treatment county deposit growth (shale counties) is calculated in excess of the mean deposit growth for matched peers, $\%\Delta D_j^{t*} = \%\Delta D_j^{Treat,t} - \overline{\%\Delta D_j^{Match,t}}$. Five matched peers determined using nearest neighbor propensity score matching on economic, geographic and demographic characteristics.

$\%\Delta Wages_j^t$ – Percent change in wages paid in the county from $t-1$ to t. Calculated as of the end of the calendar year. 1% of observations trimmed to remove extreme outliers.

$\%\Delta SmallBusiness_j^t$ – Percent change in small businesses in the county from $t-1$ to t. Small businesses defined as those establishments in the CBP with fewer than 50 employees. Calculated as of the end of the calendar year. 1% of observations trimmed to remove extreme outliers.

Bank Variables: Bank i at time t

Primarily sourced from Call Reports (FFIEC 031), bank holding company filings (FR Y-9C, FR Y-9SP). Observed on June 30th of each year.

s_i^t – Key interaction variable for bank size. Log of real total assets at $t-1$. Real assets calculated using CPI-U.

Δd_i^{t*} – Bank exposure to shale gas deposit shocks. The sum of excess deposit growth in shale gas counties, weighted by the banks deposits in the county at $t-1$ $\%\Delta D_i^{t*} = \sum_{j=1}^{N} \%\Delta D_j^{t*} \left(\frac{D_{i,j}^{t-1}}{D_i^{t-1}}\right)$, this is then scaled by bank assets at $t-1$.

Δd_i^t – The dollar change in deposits from $t-1$ to $t+1$, scaled by assets at $t-1$.

$\Delta loans_i^t$ or Δl_i^t – The dollar change in loans on the balance sheet from $t-1$ to $t+1$, scaled by assets at $t-1$.

$\Delta c\&i_i^t$ – The dollar change in commercial and industrial loans on the balance sheet from $t-1$ to $t+1$, scaled by assets at $t-1$.

Δre_i^t – The dollar change in real-estate loans on the balance sheet from $t-1$ to $t+1$, scaled by assets at $t-1$.

Δsbl_i^t – The dollar change in small business loans on the balance sheet from $t-1$ to $t+1$, scaled by assets at $t-1$. Small business loans are defined as loans for amounts less than \$1m.

$\Delta cash_i^t$ – The dollar change in cash and balances due from other depository institutions on the balance sheet from $t-1$ to $t+1$, scaled by assets at $t-1$.

$\Delta ffs\&repos_i^t$ – The dollar change in federal funds sold and short-term agreements to resell securities from $t-1$ to $t+1$, scaled by assets at $t-1$.

Δsec_i^t – The dollar change in securities held for sale and held to maturity on the balance sheet from $t-1$ to $t+1$, scaled by assets at $t-1$. Small business loans are defined as loans for amounts less than \$1m.

$\Delta liquid_i^t$ – The dollar difference in cash, securities and federal funds sold. from $t-1$ to $t+1$, scaled by assets at $t-1$. This reflects uses for cash to improve liquidity.

$\Delta borrow_i^t$ – The dollar difference in federal funds purchased and other borrowed money (including notes and debentures) from $t-1$ to $t+1$, scaled by assets at $t-1$. This reflects leverage reductions.

ΔROA_i^t – The change in the return-on-assets (net income/assets) from $t-1$ to $t+1$.

$\Delta loan/dep_i^t$ – The change in the loan to deposit ratio from $t-1$ to $t+1$.

ROA_i^{t-1} – Return-on-assets (net income/assets) at $t-1$.

$\%recom_i^t - 1$ – The share of assets invested in commercial real estate loans at $t-1$.

$\%\Delta smallbusiness_i^t$ – The weighted average of small business growth (< 50 employees) for year t for each county the bank, i, has branches. Growth is weighted by the share of the bank's deposits in the county at $t-1$.