Collateral Damage: The Impact of Shale Gas on Mortgage Lending*

Yanyou Chen

James Roberts

Christopher Timmins

Ashley Vissing

PRELIMINARY AND INCOMPLETE

DO NOT CITE

January 24, 2019

Abstract

We analyze mortgage lenders' behavior with respect to shale gas risk during the period of the U.S. shale gas boom, which coincided with the U.S. housing market rise, collapse and subsequent increase in lending industry scrutiny. Shale gas operations may place affected houses into technical default such that GSE's (Fannie Mae and Freddie Mac) are unable to maintain them in their portfolios. We find that lenders did indeed increase the weight they place on shale risk relative to income risk in mortgage pricing behavior after 2010. This effect is particularly evident for groundwater dependent properties, indicating that lenders view shale activities as placing the residential value of these properties at greater risk. When we quantify the willingness to pay to avoid shale risk, we find that insurers, on average, lose around \$2,394.9, or 1.2% of profit earned on an average mortgage.

*Chen, Roberts and Timmins – Department of Economics, Duke University, PO Box 90097, Durham, N.C. 27708. *yanyou.chen@duke.edu*, *christopher.timmins@duke.edu*. Vissing – Energy Policy Institute at the University of Chicago, Saieh Hall for Economics, 5757 S. University Avenue, Chicago, IL 60637. *abvissing@uchicago.edu*. The authors would like to thank seminar participants at the Property and Environment Research Center and the Duke Energy Fellows Lunch for their helpful comments and criticisms. All remaining errors are our own.

1 Introduction

The Energy Information Administration reports that between 2007 and 2014, annual natural gas withdrawals from shale deposits in the U.S. rose from 1.99 to 13.8 million cubic feet,[1] growth that led to lower natural gas prices, higher natural gas demand, and substitution from other forms of fossil fuel consumption like coal. The recent rise of shale gas has been spurred by technological development that combines large-scale hydraulic fracturing and horizontal drilling techniques with other advances in three-dimensional surveying techniques. In addition to allowing for more efficient extraction from broad, tight-shale layers, the horizontal drilling technology increases access to large areas of shale from relatively confined surface areas, which allows firms to extract oil and natural gas stored in tight-shale formations located beneath densely populated neighborhoods.[2] Unlike traditional oil and natural gas reservoirs, settlement patterns throughout the U.S. have evolved for decades with indifference to the location of shale. Consequently, we now often find residential properties located on top of those resources, bringing shale gas activity into homeowners' backyards, and beginning in 2014, more than 15 million Americans live within one mile of an active oil or gas shale well.[3] Our paper explores how the extraction technology changes have altered mortgage markets by evaluating how lenders internalize the increased risks born by leveraged houses located near to shale oil and natural gas development.

What are the chief risks born by proximity to these large-scale, hydraulically fractured wells? Current literature is evolving to identify and measure the health, economic, and geologic consequences of proximity to these industrial activities, especially in light of increased household exposure.[4] Among the identified risks is air pollution from well production and transmission activities, which leads to increased methane emissions.[5] Waste water disposal from fractured wells is linked to surface water contamination caused by radioactive salts and metals or by the chemicals used to the treat the fractured wells.[6] Geologically, there is a growing literature linking the employment of large-scale fracturing technology and increased incidence of tremors.[7] The

[1] Energy Information Administration, `https://www.eia.gov/dnav/ng/hist/ngm_epg0_fgs_nus_mmcfa.htm`

[2] In particular, many horizontal laterals can be drilled in different directions from a single wellpad, which may comprise less than an acre of land. Horizontal laterals can extend in mile-long segments beneath suburban and urban regions.

[3] Gold and McGinty (10/25/13). "Energy Boom Puts Wells in America's Backyards." *Wall Street Journal*.

[4] Coase (1960) noted that externalities are reciprocal in nature, and that the external costs of a production process will be low if no one is around to be affected by it. In contrast, the evolved shale gas extraction technology allows for many impacts on nearby landowners.

[5] The literature measures the impact of increased methane emissions during the drilling, fracturing and production phases of well development, in particular, including Caulton et al. (2014), Brandt et al. (2014), and the increases in other particulate matter, including Colborn et al. (2011) and Roy et al. (2014), and volatile organic compounds, as in Gilman et al. (2013).

[6] Olmstead et al. (2013), Warner et al. (2013), Fontenot et al. (2013), and Hill and Ma (2017)

[7] Koster and van Ommeren (2015) and Cheung et al. (2016)

hedonic literature identifies community-wide costs and benefits of proximity to shale, whereby the costs are greatest for households accessing groundwater,[8] and more broadly, communities experience degradation of amenities through increased noise, road damage, and traffic accidents.[9] Conversely, the literature also identifies and measures economic benefits to having an active extraction industry, which includes higher wages, income, and municipal revenue.[10] Added to the economic benefits are royalties and bonus payments earned by households that own (and lease) the rights to their sub-surface minerals. We contribute to this literature by identifying and measuring an additional cost of these technologies that is internalized by the mortgage lending industry and passed on to homeowners through higher lending rates.

The fact that shale development has the potential to impact property values means that it also has the potential to interact in a variety of ways with mortgage lending practices. First and foremost, a mortgage loan is commonly secured for both surface and subsurface rights. Mortgages do not generally allow homeowners to sell or lease parts of their property without prior approval from the lender; however, mortgage experts generally report that requests for approval are rare (NYT (2011) and Law (2011)).[11] The situation is further complicated by the fact that most mortgages are not held in the portfolio of the primary lender, but are instead sold on the secondary mortgage market. Lenders participating in the secondary mortgage market, in effect, sell their loans to government sponsored enterprises (GSE's)[12] or investment banks that bundle the loans (securitize) and sell the resulting securitized assets (mortgage-backed securities) to individual investors. Participation in the secondary market increases lenders' liquidity; however, GSE's are prohibited from purchasing mortgages on properties engaged in industrial activities including transport or storage of toxic substances (chemicals, oil and gas products, or radioactive materials) (Law (2011)), and shale gas extraction involves a number of activities that have the potential to violate these rules, which would leave the borrower (without prior approval) in

[8]Among other hedonic analyses, recent research has shown that shale development can negatively impact nearby housing prices (Throupe et al. (2013), Gopalakrishnan and Klaiber (2014), James and James (2014), and Muehlenbachs et al. (2015)), but other research has reached differing conclusions, as in Delgado et al. (2014) and Boslett et al. (2016).

[9]Muehlenbachs and Krupnick (2013) and Graham et al. (2015) link increased incidence of traffic accidents to drilling activity.

[10]Mason et al. (2015) and Bartik et al. (2016) measure some of the economic benefits, and Cesur et al. (2017) documents the benefits to infant mortality that are attributable to energy generation from natural gas as opposed to coal.

[11]Nationwide, Bank of America reports receiving approximately 100 such requests per month, fewer than a dozen are sent to Fannie Mae and Freddie Mac each year. (NYT (2011)) Chesapeake Energy, one of the largest drilling companies, only seeks permission from lenders before a property is drilled, not before it is leased; this violates mortgage rules, which require approval to sign a lease. (NYT (2011))

[12]Government sponsored enterprises (GSE's) include Fannie Mae (Federal National Mortgage Association) and Freddie Mac (Federal Home Loan Mortgage Corporation).

technical default.[13,14] Because the primary lender is responsible if the borrower defaults due to shale extraction activities, lenders have strong incentives to precisely evaluate the risks incurred by lending to homeowners in regions with high levels of shale development.

We use the mortgage market to learn about additional external costs associated with the "fracking boom" by quantifying the relationship between shale drilling risks and the frequency of subprime (higher priced) lending. In particular, we use loans issued to buy homes in Tarrant County, Texas, to test whether lenders mitigate shale development risks by issuing more subprime loans to households that are susceptible to drilling externalities. We measure the extent to which lenders' preferences (distaste) for shale gas risk have shifted relative to other forms of risk since 2010 (i.e., during a period when the industry reconsidered its lending practices and shale gas activities rapidly expanded). In particular, the rise in shale extraction was coupled with the collapse of the housing market and, with it, an increased scrutiny of GSE activities leading lenders to worry they may have to assume control of the shale-exposed mortgages as a consequence of technical default. We find that it is important to estimate differential preferences for shale risk (relative to income risk) before and after this period in which technical default and lost property values become more significant concerns among lenders. Moreover, we consider how those preferences vary with water source, which is an important component of house value and may be particularly vulnerable to shale risk. In particular, we ask whether lenders treat properties that are already exposed to nearby drilling differently from those that have signed a lease to extract their minerals and are likely to be exposed hydraulic fracturing operations, compared to those that do not yet have a lease.

We use a combination of housing, lending, and drilling data to capture variation in subprime lending practices across households that are more or less susceptible to the negative externalities of nearby drilling behavior. Our data spans the growth of shale gas activity and the periods before and after the financial crisis, and we are able to estimate how lenders' preferences differed

[13] Further, there may be other discrepancies between the terms specified in an oil or natural gas lease agreement and those specified in a mortgage. For example, some states void title insurance if a property is used for any commercial venture (Law (2011)) and, without title insurance, participation in the secondary market is limited. Homeowner insurance policies will also be violated if industrial activities are present, leading to default (May (2011)).

[14] Borrowers may even find themselves in violation of their mortgage agreement through no fault of their own, but rather simply because they happen to be located in close proximity to another property that engaged in shale development. For example, news reports described a couple in Washington County, PA who were denied a new mortgage for their property by Quicken Loans because of shale development on a neighboring plot. Quicken responded by saying – "While Quicken Loans makes every effort to help its clients reach their homeownership goals, like every lender we are ultimately bound by very specific underwriting guidelines. In some cases conditions exist, such as gas wells and other structures in nearby lots, that can significantly degrade a property's value. In these cases, we are unable to extend financing due to the unknown future marketability of the property. (http://www.wtae.com/investigations/Couple-denied-mortgage-because-of-gas-drilling/12865512#.T6mu842bM44.facebook)"

across these two periods along several dimensions. In the data, we observe that lenders weight shale risk more heavily in their subprime decisions after the financial crisis. We then use a non-parametric estimator, first introduced by Frölich (2006), that captures firms' preference heterogeneity for income and shale risks and allows us to estimate the trade-offs across our two types of risk without imposing functional form assumptions. Focusing on the comparisons across water sources, we estimate a distribution of lenders' preferences for the two sources of risk and show how it differs by groundwater versus piped water. In particular, we find that after the financial crisis, lenders are willing to bear significantly more income risk to reduce their exposure to shale risk among houses accessing groundwater.

Our paper differs from the existing shale gas literature by focusing on lenders' decisions who, we argue, are more likely to internalize risks especially in light of greater scrutiny in lending practices post-financial crisis, which evolved throughout 2008 and 2009.[15] Lenders face trade-offs between different types of lending risk on a daily basis, allowing them for develop an expertise in weighing the risks of any particular loan. Conversely, households may only buy a few houses over the course of their lives, which tasks households with a large information burden when deciding how to internalize perceived shale risks. Using measures that capture shale and income risks, along with other household characteristics, we non-parametrically estimate lenders' preferences to issue high risk loans, as indicated by a subprime mortgages. Following Bajari and Benkard (2005) and Frölich (2006), we characterize lenders' heterogeneous preferences and trade-offs between two types of risks: income and shale.

The remainder of the paper proceeds as follows. We begin by fitting the analysis into additional literature that is relevant to our questions in Section 2. We then describe the secondary market for mortgages in Section 3 and our data set in Section 4, which combines information from the Home Mortgage Disclosure Act with data on property appraisals and well locations. Section 5 describes the theory used to describe lenders' relative preferences for shale and income risk and Section 6 describes the empirical framework. Section 7 reports estimates and Section 9 concludes with a discussion of policy relevance.

[15] Our study is a complement to literature exploring whether there is an observed difference in mortgage default behavior for households living in active shale regions compared to non-shale. Upton and McCollum (2016) studies the probability of default in the presence of shale gas development, examining the default behavior of landowners in a period when national default rates were rising. He finds that landowners living in shale gas regions are less likely to default on their mortgage loans, lending credence to the positive economic impacts of shale gas development. Shen et al. (2015) conduct a similar study with data describing default behavior of households in Pennsylvania located over the Marcellus Shale.

2 Other Relevant Literature

In addition to the already cited environmental economics literature identifying and measuring the costs and benefits of the growing shale extraction industry,[16] this paper contributes to literature concerned with the effects of the financial crisis on lending practices, subprime lending practices, more generally, and the relationship between shale development and foreclosure incidence.

In addition to estimating preference heterogeneity for shale and income risk, we capture differences in lending practices before and after the financial crisis, which is a period of increased foreclosures and significant changes to industry behavior that emphasizes greater scrutiny for each considered mortgage. We add to the literatures linking increased foreclosure rates with subprime lending[17] and impacts on property values and neighborhoods. Gerardi et al. (2007) study the relationship between subprime lending and the increased incidence of observed foreclosures in the context of the Massachusetts housing market. Focusing their analysis on years encompassing the U.S. financial crisis, the authors find that, along with house price depreciation, subprime lending is positively correlated with the incidence of foreclosure citing that subprime borrowers are 20% more likely to face foreclose than prime borrowers. McCarthy et al. (2002) cite the longer-term effects of foreclosures on households' access to stable and decent housing, and Simons et al. (2009) and Lin et al. (2009) capture the spillover effects of foreclosed properties on the surrounding blocks. Further, Immergluck and Smith (2006) find a positive correlation between foreclosures and violent crime levels at the neighborhood level.

In our analysis, we control for households' race and ethnicity in order to capture the oft cited relationship between subprime lending and sociodemographic characteristics. Munnell et al. (1996) is the among the papers that use data assembled in compliance with the Home Mortgage Disclosure Act to study potential discriminatory lending practices toward minority households.[18] Gerardi and Willen (2009) study Massachusetts property-level data and find that subprime lending leads to more turnover in minority neighborhoods.

The model we present captures the risk taking behavior of local and national lenders when assessing mortgages for properties that are potentially impacted by shale gas development. This is a complement to a new literature documenting the relationships between shale gas activity

[16] Our review of this literature is by no means complete, as it is a growing literature.

[17] Some of these studies also include those by Newman and Wyly (2004), Immergluck and Smith (2005), and Quercia et al. (2007).

[18] The authors build on the literature analyzing discriminatory lending practices in mortgage markets that include Black et al. (1978), King (1981), Schafer and Ladd (1981) and Ladd (1998), among many others with additional literature summarized in Yinger (1996) and LaCour-Little (1999). A more descriptive analysis of the HMDA data and sociodemographic heterogeneity can be found in Avery et al. (2006).

and capital allocation and between shale gas activity and foreclosures. It also draws upon the literature analyzing other observable factors that lead to foreclosures, or more risky loans from the perspective of banks. A nascent literature explores the relationship between unconventional energy development and the local banking sectors. Plosser (2014) studies the deposit shock resulting from landowners profiting from nearby shale gas development and the banks' resulting capital allocation decisions. Gilje (2012) uses shale discoveries as a natural experiment to study credit supply and its implication for economic outcomes especially in areas dominated by smaller banks. Gilje et al. (2016) explores whether banks export liquidity because they are exposed to positive shale gas shocks. The authors find that banks with branches located in non-shale areas increase lending in those branches, and the positive effect is more pronounced for 'difficult-to-securitize' loans. Finally, there is an older literature that employs models to capture the default probabilities for mortgages as a function of factors like loan-to-value ratios and house-price depreciation.[19] Gerardi et al. (2009) is a more recent paper exploring the effects of house price depreciation on foreclosure rates versus relaxed underwriting standards. They find that the two mechanisms are interrelated, but, without house-price depreciation, there would not have likely been such a dramatic increases in the foreclosure rate.

3 Secondary Markets for Mortgages

Lenders interested in selling their mortgages to government sponsored enterprises (GSE's), like Fannie Mae (Federal National Mortgage Association) and Freddie Mac (Federal Home Loan Mortgage Corporation), in the secondary market must meet certain criteria that may be violated by proximity to shale development. These details are noted in the following section along with a short history of GSE's and potential risks of shale gas development.

Among the largest actors in the secondary mortgage markets are the GSE's. The government established these enterprises to introduce additional liquidity into the mortgage market and promote home ownership. In the U.S., 90% of all houses are purchased with mortgage financing. Lending institutions typically do not rely on their own capital to support most of the loans that they write over the long-term. Instead, loans are bundled (securitized) and sold to investors as mortgage-backed securities. The federal government, primarily through its GSE's, is the largest of these investors, purchasing over 90% of mortgages in the US today.[20] Fannie Mae and Freddie Mac assume the credit risk for all mortgages that are re-sold as mortgage-backed securities. In

[19] Refer to Quercia and Stegman (1992) and Kau et al. (1994) for other contributions to this literature.

[20] https://smartasset.com/mortgage/everything-you-need-to-know-about-the-secondary-mortgage-market

exchange for bearing this risk, the GSE's keep the guaranty fees associated with the loans. As recently as 2008, Fannie Mae and Freddie Mac had owned or guaranteed about half of the U.S.'s $12 trillion mortgage market. This, along with the low rates at which the GSE's are typically able to borrow, make them highly profitable enterprises. After the housing market crash of 2008, lenders began to evaluate the standards under which home mortgage loans were approved and credit was subsequently tightened. Both GSE's underwent significant scrutiny, were restructured, and fell under the conservatorship of the Federal Housing Finance Agency. Since that time, both the House Financial Services and Senate Banking Committees passed reforms that would have reduced the government's footprint in housing finance (although neither was passed into law),[21] and a similar plan was proposed by the Obama administration.[22] With this as a backdrop, the continued growth of shale gas in the U.S. led to congressional hearings beginning in 2009, continuing in 2010,[23] and culminating in a flurry of activity in the fall of 2011 both inside and outside Washington.[24]

3.1 The GSE's and Shale Gas

The GSE's specify a set of criteria to which lenders must adhere if they want to be able to sell mortgages on the secondary market, and there are many areas where these criteria may conflict with standard practice in shale gas or oil development. For example, Freddie Mac guidelines 39.4(i) specify that a mortgage can only be issued on a leased property if "exercise of the rights will not result in damage to the mortgaged premises or impair the use or marketability for residential purposes." Furthermore, the guidelines prohibit "right of surface or subsurface entry within 200 feet of a residential structure," and require "comprehensive endorsement to the title insurance company that affirmatively ensures the lender against damage or loss from exercise of such rights." Practically, this requires "no structure erected on premises exceeding three stories or 35 feet," that the premises "shall not be used for storage of any material machinery, equipment, or supplies," and that the property will "not be used for commercial purposes." Furthermore, the Freddie Mac guidelines 39.4(m) require that lenders must warrant that activities on the property:

[21] https://www.nafcu.org/HousingFinanceReform/

[22] https://www.treasury.gov/initiatives/Pages/housing.aspx

[23] Bateman, C. (June 2010).“ A Colossal Fracking Mess.” Vanity Fair. http://www.vanityfair.com/news/2010/06/fracking-in-pennsylvania-201006.

[24] Urbina, I. (November 24, 2011). “Officials Push for Clarity on Oil and Gas Leases.” New York Times. http://www.nytimes.com/2011/11/25/us/officials-push-for-clarity-on-oil-and-gas-leases.html?_r=0

1. must not interfere with the use and enjoyment of any present or proposed improvements on the mortgaged premises or with the use and enjoyment of the balance of the mortgaged premises not occupied by improvements,
2. must not affect the marketability of the mortgaged premises,
3. must have no or minimal effect on the value of the mortgaged premises,
4. must be commonly acceptable to private institutional mortgage investors in the area.

Fannie Mae and Freddie Mac have a fiduciary duty to establish rules that reduce the risk of lost house value or default. As noted in the introduction, the extraction of shale gas involves a number of activities that have the potential to violate these rules, which would leave the borrower (without prior approval) in technical default. Toxic chemicals are pumped, along with million gallons of water and sand, directly under mortgaged homes. "Produced" water, which is forced back out of the well, contains brine, fracking chemicals, and even radioactive substances. It is often stored on-site, sometimes in open holding ponds. Permanent easements for truck and pipeline transport, production platforms, and storage facilities (that can spill) are common on properties with or near drilling activity. Risks to home values can be a particular problem for homes where the water supply is threatened. Finally, without title insurance (see above), secondary lenders may not be able to hold mortgages.

The primary lender is responsible if the secondary lender does not know about the lease and the house goes into technical default as a result. Fannie Mae and Freddie Mac can demand that the originating lender buy back any loans that do not meet secondary market requirements (Carpenter 2011). There is not a good measure of how many mortgages may currently be in violation of secondary mortgage market rules.

In light of these growing concerns, a primary lender who believes that a property may soon be approached for shale development may worry that the property could default, would have to be foreclosed upon, or that shale development might hamper its ability to sell the mortgage on the secondary market. A related concern might be that noncompliant mortgages already sold to the secondary market would have to be bought back. As such, that lender may charge a premium to lend when there is concern over impending shale development. Alternatively, to the extent that they are able, lenders may simply exit markets where shale gas is prevalent.[25]

In the remainder of this paper, we quantify changes in lenders' preferences by estimating

[25] NYT (2011) reports that, in 2011, at least eight local or national banks did not typically issue mortgages on properties exposed to shale gas development. In other instances, lenders began requiring drilling companies to indemnify property owners against any future losses to home value, or requiring home owners to expressly agree not to sign a lease as long as they hold the mortgage.

the changes in trade-offs between income and shale risk before and after the financial crisis and concurrent increased interest in shale gas at the federal level. This provides us with a new perspective on the costs of the risks associated with shale gas development for nearby homeowners, specifically measured via the decisions made by housing *professionals* (i.e., mortgage lenders).

4 Data

The following subsections describe our data location and sources and our variable construction methods. Our data is comprised of housing, drilling, and lending data that allows us to construct a household-level portrait of house transactions, household characteristics, and income and shale risk factors related to our dependent variables of interest, namely high-interest loans and foreclosures.

4.1 Tarrant Co., Texas

Our analysis will focus on shale gas development and its impact on property markets in Tarrant County, Texas. Tarrant Co., located in north-central Texas, is the home to approximately 1.8 million residents. It is comprised of 41 incorporated areas, including Fort Worth, which is the county seat. The population of Tarrant is approximately 27% Hispanic or Latino (of any race).[26] Tarrant Co. and the underlying Barnett shale are typically considered to be the birthplace of modern hydraulic fracturing because of innovations made there by Mitchell Energy.

We also report summary statistics and present model specifications that include Denton Co., located north of Tarrant, where there is also active drilling but in a more rural setting. Finally, we estimate counter-factual relationships using Dallas Co. data where firms are restricted to drill wells located at least 1,500 feet away from residential and commercial buildings, effectively a drilling moratorium.[27] Dallas is located directly east of Tarrant County, and since there is no drilling, our counter-factuals assume lenders need only evaluate income risk (as compared to income and shale risk).

4.2 HMDA

HMDA was established to determine whether lenders serve communities' financial needs and facilitate enforcement of fair lending laws. When buying a house, one typically fills out a form

[26] www.tarrantcounty.com/en/county/about-tarrant.html

[27] These setback rules are quite stringent compared to Fort Worth where setbacks are 600 feet. https://www.texastribune.org/2013/12/11/dallas-city-council-tightens-gas-drilling-ordinanc/

at closing that transmits information about the race, sex and income of the buyer along with the loan amount and terms. As of 2006, there were 8,850 lenders covered by the disclosure rules (approximately 80% of home lending nationwide). In 1989, the HMDA law was amended to require disclosure of loan-level information, and in 2004, it was further amended to require disclosure of information about loan pricing. Specifically, the lender is required to report the spread between the annual percentage rate and the applicable Treasury yield if it is greater than or equal to 3 percentage points for a first-lien loan. After 2009, the rule for first-lien loans was changed to require reporting if the difference between the annual percentage rate and the applicable average prime offer rate is greater than or equal to 1.5 percentage points (i.e., both the baseline and the cutoff rule changed, see Fed 2009). As such, HMDA does not specifically identify subprime loans, but rather "higher priced" loans; we use their reporting requirement in each period as the determinant of higher priced. We also follow the common practice in the literature and use these terms higher priced and subprime interchangeably. In 2004, the first year when pricing information was provided, fewer than 20% of households had higher priced loans, and higher priced loans were more common amongst black and Hispanic borrowers.

The number of loans reported as higher priced depends upon many factors, some of which have nothing to do with the borrower's riskiness. In particular, a narrowing of the difference between short and long-term interest rates can increase the the number of loans overall exceeding the higher priced threshold; we account for this in our empirical model below.

We employ the loan-to-income ratio as our measure of *income risk*. *Ceteris paribus*, given two borrowers with the same income, the borrower with the larger total loan amount will be more at risk of shocks that will prevent repayment of the loan, leading to default. HMDA reports both the size of the loan and the borrower's (self-reported) income. Further, HMDA describes whether or not the loan was securitized by a government agency or a commercial lender and whether the loan was issued by the Federal Housing or Veterans Administrations, which are important control variables included in the empirical specifications and are described in Section 7.1.

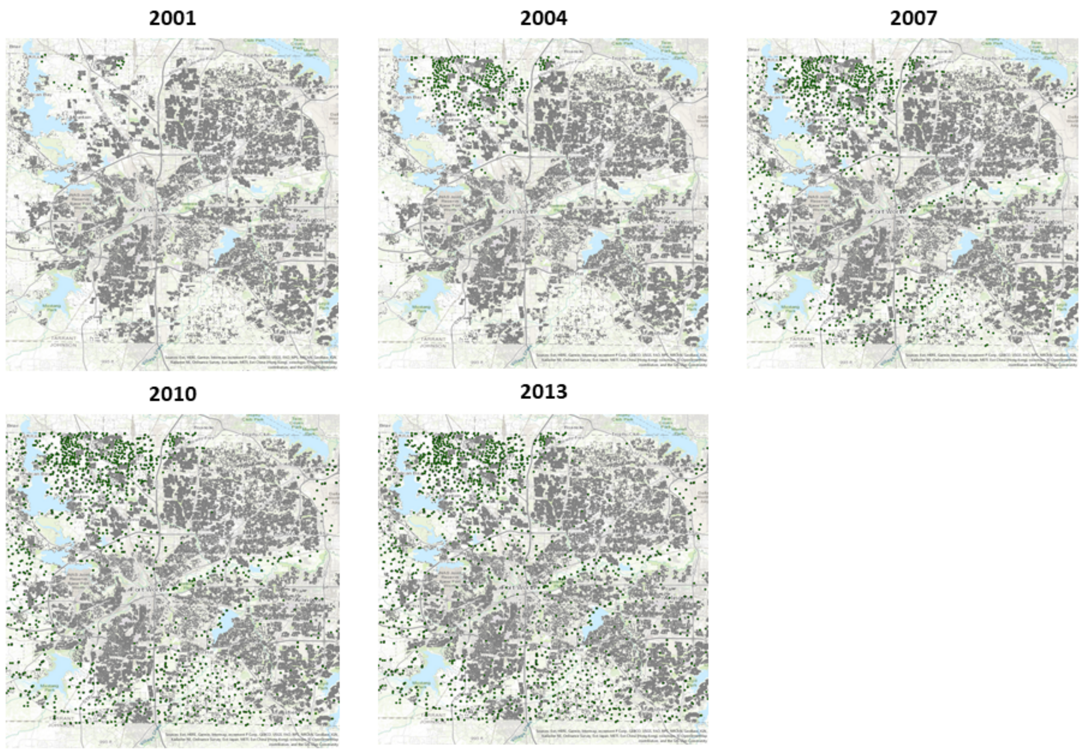
4.3 Dataquick & Corelogic

Data from the real estate data services companies, like Dataquick and Corelogic that are accessed through an a licensing agreement with the Duke University Department of Economics, are used in conjunction with information from the Tarrant Co. Assessor's Office, to measure the sale and assessed values of homes located in Tarrant, Denton, and Dallas counties of Texas, sale dates, and other house attributes like the counts of bedrooms, bathrooms, and living and land square footages. Further, these companies provide information describing the lenders and loan characteristics like loan amount, whether or not the loan was issued by the Federal Housing or Veterans Administrations, etc... Corelogic uniquely identifies whether the home is eventually foreclosed, which is a second dependent variable explored for the purposes of our "back-of-the-envelope" calculations. We connect the HMDA data to our housing data by merging with the following data fields: lender name, lender amount, zip codes, and sale dates.

4.4 Exposure to Shale Development

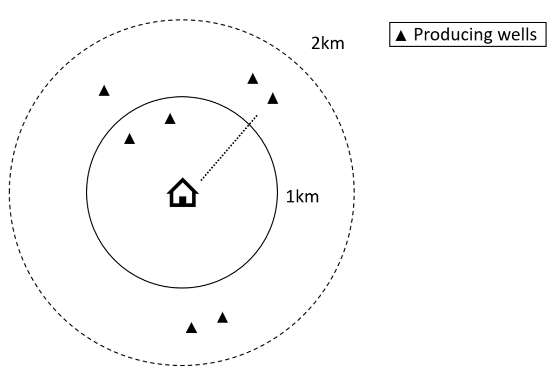
To estimate the relationship between shale risk and high-interest lending practices, we propose both a simple and complicated measure of shale risk. First, we consider two simple specifications: we count the number of producing wells located within 2-kilometers of the sold house as of the sale date, and we calculate the distance between the sold house and the nearest producing well. The logit estimates using these definitions are reported in the Appendix. However, by only looking at the number of producing wells or distance to the nearest producing well fail to accurately measure the shale risk. Figure 1 illustrates where producing wells emerge in Tarrant county in different time periods. From Figure 1, we observe that first the producing wells do not emerge in a uniform direction. From 2001-2004, the wells mostly locate on the north west region of Tarrant County, but after 2004 wells emerge from south as well. Moreover, the speed of emergence is different for different neighborhoods. Once wells start to emerge near a neighborhood, the drilling activites are become much more active in that neighborhood than other neighborhoods where no producing well exists. Therefore, any specification that only considers one characteristic (distance or number of wells within a certain distance) and specification that assumes linear relation on the characteristics will mis-measure the shale risk.

Figure 1: Patterns of Producing Activities



To better estimate shale risk, we use counts of wells and distance to the nearest producing well to estimate a duration model. In particular, shale risk is measured by the cumulative hazard function, which measures the total amount of risk that has been accumulated up to time t. In our duration model, exposure starts when a drilling well appears within the 3-km radius of the sold house and a failure occurs when a drilling well appears within the 1-km radius. For houses where a failure has not yet occurred, we consider the following attributes in calculating the cumulative hazard function: number of producing wells within 2-km radius, and distance to the nearest producing well. Figure 2 illustrates how the attributes are obtained.

Figure 2: Measure of Shale Risk



In particular, we use the proportional hazard (PH) model to calculate shale risk. Let τ be a non-negative random variable denoting the time to a failure event, and denote $S(t)$ as τ's survivor function and $h(t)$ as its hazard function. Then we have

$$
\begin{aligned}
S(t) &= 1 - F(t) = Pr(\tau > t) \\
f(t) &= \frac{dF(t)}{dt} = \frac{d}{dt}[1 - S(t)] = -S'(t) \\
h(t) &= \lim_{\Delta t \to 0} \frac{Pr(t + \Delta t > \tau > t | \tau > t)}{\Delta t} = \frac{f(t)}{S(t)}
\end{aligned}
$$

In our PH model, we assume that the hazard function $h(t)$ has a Weibull distribution, with time-varying covariates:

$$
\begin{aligned}
h(t) &= h_0(t) \exp(\beta_0 + \mathbf{x}(\mathbf{t})'\beta) \\
&= pt^{p-1} \exp(\beta_0 + \mathbf{x}(\mathbf{t})'\beta)
\end{aligned}
$$

The time-varying covariates $\mathbf{x_{it}}$ include: {number of producing wells within 2-km radius, distance to the nearest producing well}. Start of exposure is defined as when the first producing well emerges within 3-km radius. Length of exposure t is measured by total number of months since start of exposure, which is defined as:

$$
t = \text{current date} - \text{start of exposure}
$$

We estimate the parameters of this model through maximum likelihood estimation. The likelihood function is:

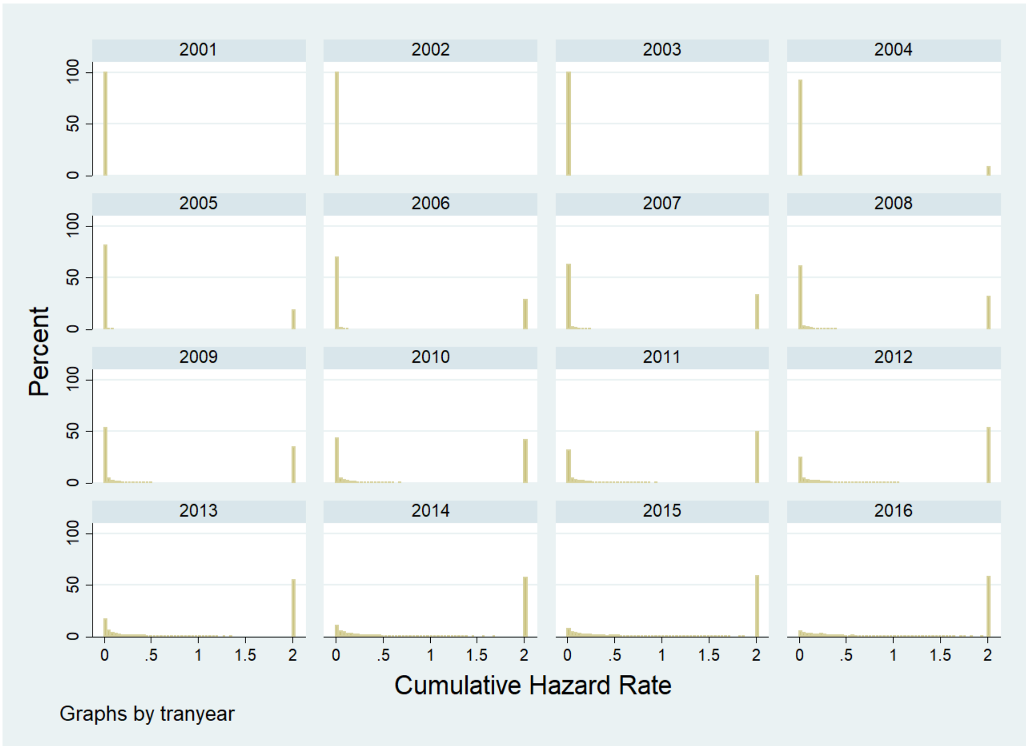
$$
\begin{aligned}
\mathcal{L}(\beta | t, \mathbf{x}, T) &= \prod_{i=1}^{k_1} f(t_i | \mathbf{x_i}(\cdot), \beta) \prod_{j=1}^{k_2} Pr(t_j > T | \mathbf{x_j}(\cdot), \beta) \\
&= \prod_{i=1}^{k_1} S(t_i | \mathbf{x_i}(\cdot), \beta) \cdot h(t_i | \mathbf{x_i}(t_i), \beta) \prod_{j=1}^{k_2} S(t_j | \mathbf{x_j}(\cdot), \beta)
\end{aligned}
$$

where T is the termination period in our sample, k_1 is the set of houses that are exposed to shale risk during the sample period, and k_2 is the set of houses that are not yet exposed to shale risk by the end of our sample period. The cumulative hazard function is then given by:

$$
H(t) = \int_0^t h(u) du
$$

Figure 3 depicts the distributions of cumulative hazard rates across years. The mass point at 0 indicates number of houses that have not yet been exposed to any shale risk. The mass point at 2 indicates houses that have had a failure (a producing well appears within 1-km radius).

Figure 3: Distribution of Cumulative Hazard Rate



The results of the PH model estimation are reported in Table 1:

Table 1: Estimation Results of PH Model

	PH Coefficient
Number of producing wells within 2000m	-0.0239***
	(-68.56)
Distance to the nearest producing well	-0.00216***
	(-134.21)
_cons	-6.204***
	(-170.32)
logP	0.615***
	(146.66)
N	82,400

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Estimation Results of PH Model (Specification 2)

	PH Coefficient
Number of producing wells within 3000m	-0.0190***
	(-50.06)
Number of producing wells within 2000m	0.0104***
	(13.92)
Distance to the nearest producing well	-0.00218***
	(-135.31)
_cons	-6.258***
	(-170.81)
logP	0.646***
	(155.11)
N	82,400

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.5 Summary Statistics

We conclude this section by describing the dependent and independent variables, the characteristics of households with and without subprime mortgage rates, and motivating the inclusion of important control variables. Particularly, it is important to control for household characteristics, including race and ethnicity, loan characteristics, whether the loans are securitized and by whom, and whether the loan is to purchase an owner-occupied house.

Table 3 summarizes the house, household, sale, and loan characteristics of all transaction occurring from 1999 to 2016 in Dallas, Tarrant, and Denton counties. Denton county is comprised of fewer minorities while Dallas is comprised of more, and, on average, Denton sells larger homes measured by both the parcel (land) size and the living space. Tarrant and Denton counties both have drilling activity, though average well exposure is greater in Tarrant County, whereby shale exposure is measured by the count of producing wells located within 1,000 meters of the house at the sale date. Sales, loan, and income values are highest in Denton followed by Dallas. The frequency of foreclosures and the average mortgage interest rates (among those reported) in Tarrant fall between Denton, reporting the lowest, and Dallas, reporting the highest. Tarrant, on the other hand, has the highest levels of FHA and VA supported mortgages (consistent with lower income and sales values.

Race & Ethnicity. It is important to control for the racial and ethnic composition of borrowers when performing empirical analyses that assess mortgage lending practices. As noted in the

Table 3: Summary Statistics by County

	Tarrant		Denton		Dallas	
	Mean	(Std. Dev.)	Mean	(Std. Dev.)	Mean	(Std. Dev.)
House Characteristics						
Beds	3.354	(0.642)	3.466	(0.652)	3.287	(0.686)
Baths	2.104	(0.583)	2.494	(0.828)	2.367	(0.914)
Living (sqft)	2121.508	(811.609)	2360.904	(853.044)	2025.629	(855.58)
Land (sqft)	10257.93	(8269.194)	10644.38	(9941.089)	9595.025	(5274.123)
GW	0.645	(0.478)				
Age	18.733	(20.044)	11.647	(11.925)	29.072	(21.915)
Shale Exp. (wells)	0.575	(1.728)	0.269	(1.173)		
Household Characteristics						
White	0.81	(0.392)	0.844	(0.363)	0.684	(0.465)
Hispanic	0.193	(0.394)	0.117	(0.321)	0.282	(0.45)
Black	0.088	(0.284)	0.052	(0.222)	0.163	(0.369)
Asian	0.05	(0.218)	0.074	(0.262)	0.061	(0.239)
Income	81531.65	(54400.54)	95641.52	(57043.27)	90929.45	(80411.15)
Transaction & Loan Characteristics						
High Interest Loan	0.104	(0.306)	0.077	(0.267)	0.135	(0.342)
Loan-to-Income	2.094	(0.779)	2.245	(0.811)	2.193	(0.819)
Sale Amount	140566.4	(75242.53)	174225.6	(83716.19)	164406.1	(114878.6)
Loan Value	150060.6	(89365.52)	189920.7	(101263.2)	171042	(135882.9)
Annualized L2I	0.166	(0.058)	0.172	(0.061)	0.181	(0.064)
Foreclosure	0.077	(0.266)	0.064	(0.245)	0.086	(0.281)
Fixed Loan	0.897	(0.303)	0.884	(0.32)	0.845	(0.362)
FHA Loan	0.294	(0.456)	0.216	(0.412)	0.266	(0.442)
Gov't Sec.	0.341	(0.474)	0.349	(0.477)	0.296	(0.457)
Comm. Sec.	0.509	(0.5)	0.523	(0.499)	0.525	(0.499)
Owner Occ.	0.925	(0.264)	0.925	(0.263)	0.922	(0.268)
Obs.	202,286		78,792		178,222	

[a] Summarizes the primary variables used to describe the house, household, sales, and loan characteristics. The first and second columns compare mean and standard deviations across the Dallas and Tarrant counties used in the analysis.

literature and data sections, HMDA was established, in part, to ensure lending practices were non-discriminatory. Consequently, it is important to control for both income, debt, race, and ethnicity when empirically testing lending behavior. In Table 4 we see that a subprime mortgage is more often associated with minority status (black or Hispanic, in particular) and with lower income. Further, minority households are more likely to be issued loans insured by the Federal Housing Administration, described next in the text, and the loans are less likely to be secured by a government entities like Freddie or Farmer Mac and Fannie or Ginnie Mae.

Table 4: Summary Statistics by Race & Ethnicity

	Black		Asian		White		Hispanic	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Transaction & Loan Characteristics								
High Interest Loan	0.243	(0.429)	0.057	(0.232)	0.105	(0.307)	0.205	(0.404)
Loan-to-Income	2.309	(0.788)	2.194	(0.911)	2.138	(0.801)	2.312	(0.789)
Income	69,513	(45,973)	92,619	(66,026)	91,589	(68,454)	56,037	(41,639)
Sale Amount	135,651	(73,152)	166,959	(91,133)	159,936	(97,796)	108,744	(60,177)
Loan Value	143,716	(80,294)	173,240	(108,902)	171,252	(116,492)	114,305	(67,467)
Annualized L2I	0.188	(0.063)	0.169	(0.067)	0.167	(0.06)	0.189	(0.06)
Foreclosure	0.117	(0.321)	0.106	(0.308)	0.074	(0.262)	0.112	(0.315)
Fixed Loan	0.806	(0.395)	0.887	(0.316)	0.883	(0.322)	0.885	(0.319)
FHA Loan	0.401	(0.490)	0.122	(0.327)	0.255	(0.436)	0.406	(0.491)
Gov't Sec.	0.257	(0.437)	0.418	(0.493)	0.314	(0.464)	0.290	(0.454)
Comm. Sec.	0.603	(0.489)	0.435	(0.496)	0.531	(0.499)	0.486	(0.500)
Owner Occ.	0.93	(0.255)	0.849	(0.358)	0.927	(0.26)	0.954	(0.209)
Obs.	55,760		29,243		385,188		31,206	

[a] Summarizes the financial and loan characteristics of all transactions by race and ethnicity. Black and Hispanic households have higher rates of high interest and FHA loans and lower incomes and sales values.

FHA Loans. The models also control for whether the loan is insured by the Federal Housing Administration or Veterans Administration. The FHA is a government agency that helps borrowers obtain mortgage loans by lowering the down payment requirements to as low as 3.5 percent down for qualified borrowers, whereas traditional lenders require up to twenty percent down, and in fact, the FHA insured more loans after the subprime crisis as demonstrated in Figure 4.[28,29] Borrowers that take advantage of FHA loans pay a mortgage insurance premium that varies with the financed amount. By around 2003, subprime lending supplanted much of the FHA loans by relaxing lending requirements further (some requiring zero down payments) and expediting the application process. After the subprime lending market crashed in 2007, FHA

[28]The administration was establish as part of the National Housing Act of 1934, and in 1965, it became part of the Department of Housing and Urban Development.

[29]FHA interest rates do not vary with the borrowers' credit scores as they do with conventional loans that typically spike if the credit score is less than 620. https://en.wikipedia.org/wiki/Federal_Housing_Administration

increased the number of approved loans, which reached 43.8% of all mortgages originated in November of 2009, and after the recession, FHA lending decreased to roughly 11%.[30] Mirroring the description of active players in the lending market, Figure 4 plots the share of loans insured by FHA or VA or are subprime. The dashed line plots those shares for regions with drilling activity while the solid lines are regions without drilling activity. The share of subprime loans is slightly greater in drilled regions of Texas post-2010, and for the whole duration, FHA insured a greater share of the loans originating in drilled regions.

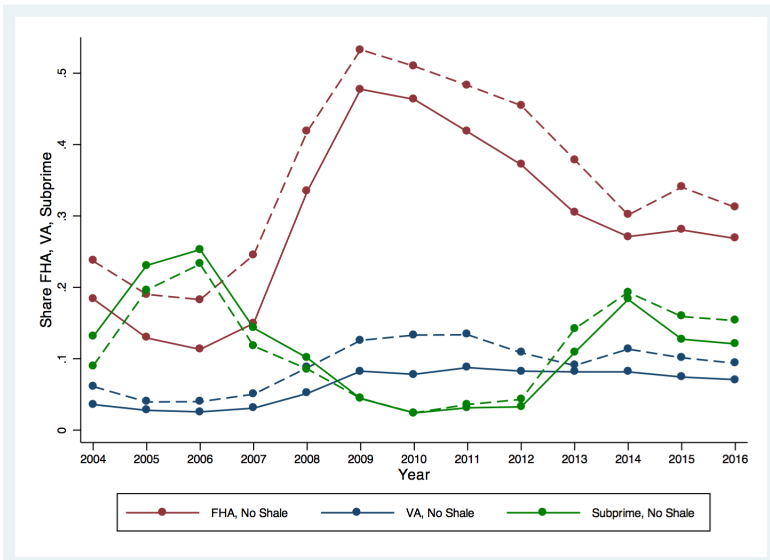
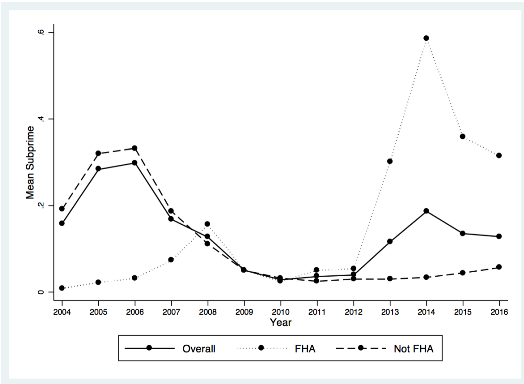


Figure 4: Frequency Originated Loan is FHA, VA, Subprime (by Shale Exposure)

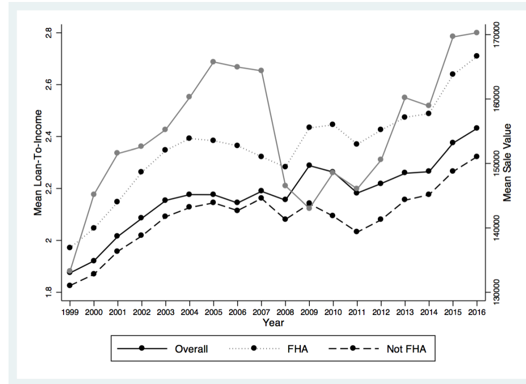
[a] Plots the share of loans issued by the FHA and VA or that are likely subprime across the state of Texas. The dashed lines indicate the share among counties with no drilling activity, and the solid lines indicate the shares among counties with drilling activity.

Figures 5a and 5b utilize the Texas-wide HMDA data to describe the frequency of high interest rate loans and mean loan-to-income values across years, stratified by whether the loan was insured by the Federal Housing Administration or not. With a decrease in FHA loans more generally, we find that FHA loans are less likely issued with high interest rates before 2007 as demonstrated in Figure 5a. After 2012, the frequency of FHA, high interest loans soars, commiserate with an overall increased dependency on FHA. Figure 5b demonstrates that FHA loans consistently insure loans issued to higher loan-to-income borrowers. The gray line demonstrates the mean sale value of homes across the same period (second y-axis), and one can

[30] https://www.clevelandfed.org/newsroom-and-events/publications/economic-trends/2015-economic-trends/et-20150414-fha-lending-rebounds-in-wake-of-subprime-crisis.aspx. This analysis explores whether FHA standards increased after the subprime lending market crashed by lending to more creditworthy borrowers. The authors find that FHA loans were extended to fewer deep subprime borrowers by the end of 2007 and to none by 2010.



(a) Mean Frequency of Subprime Interest Rates (by FHA)



(b) Mean Loan-to-Income (by FHA)

Figure 5: FHA Summary

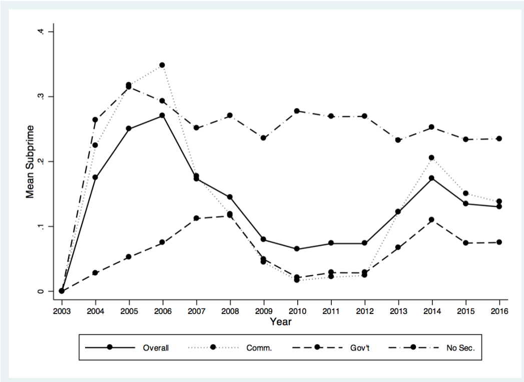
[a] 5a tracts the mean likelihood an FHA loan is issued with a high interest rate. We find that there are significantly higher interest rates among FHA loans after the subprime crisis. 5b tracts the mean loan-to-income values among all borrowers and stratified by whether the purchased home is insured by FHA. The gray line describes the mean sales values using the second y-axis.

visually see the fall in home prices in 2008 and 2009.

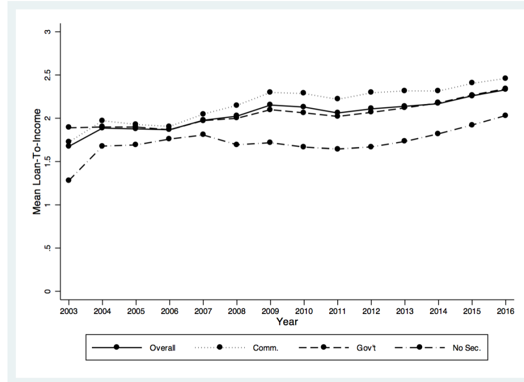
Securitized Loans. Loans may never be securitized, securitized by a government entity like Freddie or Farmer Mac and Fannie or Ginnie Mae, or they may be securitized by a commercial entity. Since primary lenders are held responsible for default, it may matter to whom they sell the originated loans. Figure 6a describes the mean frequency with which a loan is issued with high interest rate across securitized and non-securitized loans. Loans originated between 2009 and 2012 tend to have lower interest rates, and those with the highest interest rates are often not securitized. Figure 6b describes the mean loan-to-income values across securitized and non-securitized loans, and on average, the loan-to-income values are lowest among non-securitized loans. The Appendix reports summaries of the data by whether the loans are not secured, commercially secured, or government secured.

Owner Occupied Status. Research by Albanesi et al. (2017) documents that mortgage defaults were most prevalent in the middle of the credit score distribution and linked to real estate investors. Figures 7a and 7b use the Freddie Mac Single Family loan-level dataset to generate the mean credit scores and debt-to-income values among all non-FHA borrowers with loans originating between 1999 and 2016 in an effort to explore whether this is an issue for our analysis.

The mean credit scores among investors is larger than that of loans granted to purchase owner occupied units and their debt-to-income values are lower, both factors indicate greater credit-worthiness. In general, we see that non-FHA loans require higher credits scores post



(a) Mean Frequency of Subprime Interest Rates (by Securitized)



(b) Mean Loan-to-Income (by Securitized)

Figure 6: Securitized Summary

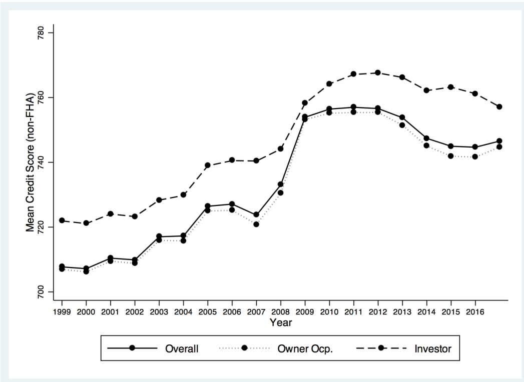
[a] 6a tracts the mean likelihood a securitized loan is issued with a high interest rate. Not securitized loans have the most consistently high interest rates, whereas the commercial loans are more likely to have higher interest rates before and after the subprime crisis. 6b tracts the mean loan-to-income values among all borrowers and stratified by whether the loan is securitized.

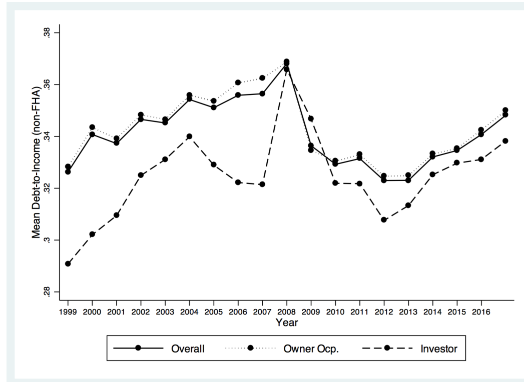
subprime crash. However, as noted before, after the crash, FHA loans comprised a greater share of all originated loans (most utilized by households with low credit scores). Among the non-FHA loans, investors' debt-to-incomes peaked near the financial crisis and then returned to the sub-owner occupied status thereafter. In general, the non-FHA loans were issued to individuals with smaller debt-to-incomes after the financial crisis, and a new upward trend begins thereafter.

Turning back to the HMDA Texas data, we find that investors are slightly more likely to have greater interest rates before the subprime crisis and lower, thereafter, as demonstrated in Figure 8a, and the average loan-to-income values are, on average, lower for investors as indicated in Figure 8b. The Texas data mirrors what is found in the Freddie Mac Single Family loan-level dataset.

National Lenders. In Table 5, we compare the number of different types of loans issued by different types of lenders. In particular, it has been suggested that "national" lenders may have simply pulled out of shale dependent areas as shale risk became an issue.[31] We report significantly more loans in the non-shale areas because the sample includes Dallas; however, the fraction of local lenders offering lower interest rate loans is similar across shale and non-shale regions and across transaction years. High interest loans issued by local lenders is slightly higher in shale regions across most years except in 2006 and 2009 when non-shale regions experienced

[31] We include the following lenders as "national": Bank of America, Wells Fargo, Starkey Mortgage, Countrywide, America's Wholesale, DHI Mortgage, CTX Mortgage, J.P. Morgan, Chase Manhatten, Chase Bank, Merrill Lynch, Credit Suisse, Fidelity, Citi Group, Citi Mortgage, Citi Bank, Citizens Bank, Coldwell, National City, Century 21, PHI Mortgage, Wachovia, Washington Mutual, Fairway Independent, H. & R. Block, Prime Lending.



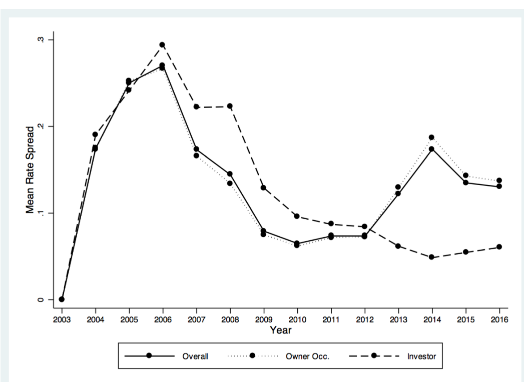


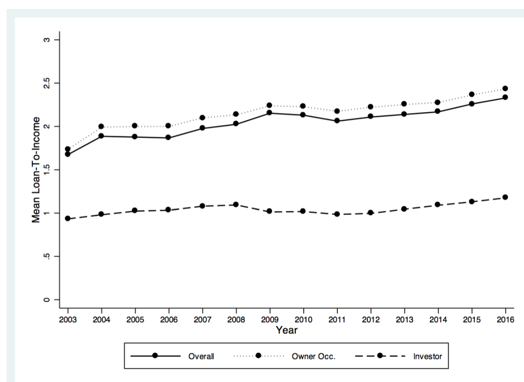
(a) Mean Credit Score

(b) Mean Debt-to-Income

Figure 7: Owner Occupied, Non-FHA Loan Summary

[a] 7b tracts the mean credit scores among all borrowers and stratified by whether the purchased home is owner occupied or not for all non-FHA loans using the Freddie Mac Single Family loan-level dataset. 7a tracts the mean debt-to-income values among all borrowers and stratified by whether the purchased home is owner occupied or not for all non-FHA loans.





(a) Mean Frequency of Subprime Interest Rates (by Occ.)

(b) Mean Loan-to-Income (by Occ.)

Figure 8: Owner Occupied Summary

[a] 8a tracts the mean likelihood a loan for an owner-occupied house is issued with a high interest rate. Investor loans are much less likely to have higher interest rates after 2009. 8b tracts the mean debt-to-income values among all borrowers and stratified by whether the loan originated for an owner-occupied house. Investors, on average, have much lower loan-to-income values.

5% more subprime, local lending.

Table 5: Types of Loans Issued by Different Lenders

	Shale Risk				No Shale Risk			
Tran.	Subprime		Non-Subprime		Subprime		Non-Subprime	
Year	Local	National	Local	National	Local	National	Local	National
2004	109 (89%)	14	598 (57%)	451	3847 (86%)	644	16496 (58%)	12157
2005	474 (84%)	89	1088 (57%)	807	8987 (85%)	1545	17070 (56%)	13220
2006	1060 (77%)	314	1908 (50%)	1931	9705 (83%)	2043	16027 (51%)	15179
2007	419 (69%)	189	2163 (48%)	2338	2539 (64%)	1420	11324 (45%)	13999
2008	218 (59%)	151	2105 (51%)	2011	1067 (57%)	803	8719 (50%)	8641
2009	113 (65%)	60	2246 (59%)	1589	490 (70%)	215	9722 (62%)	5839
2010	105 (91%)	10	2757 (63%)	1620	328 (92%)	28	8996 (64%)	4977
2011	178 (89%)	23	3360 (69%)	1483	370 (88%)	51	8296 (68%)	3867
2012	231 (91%)	23	4135 (74%)	1426	349 (91%)	33	8404 (74%)	3016
2013	477 (93%)	35	2305 (76%)	711	715 (89%)	88	4797 (75%)	1564
2014	729 (91%)	68	2459 (76%)	786	1228 (91%)	122	4445 (75%)	1474
2015	691 (91%)	67	3089 (80%)	757	887 (90%)	102	5091 (78%)	1415
2016	354 (92%)	29	1696 (81%)	389	407 (86%)	64	2680 (79%)	719

[a] Summarizes the rate of subprime loans issued by local and national lenders across shale and non-shale risk regions and transaction years.

In fact, we find that this was not the case. Figure 9 compares the difference in the percentage of all shale loans written by local and national lenders in Tarrant Co. each year to the same difference for non-shale loans, we find an almost identical time path. We do see that local lenders issued a much smaller percentage of *all* loans in the depths of the financial crisis, but that this reduction was proportional across shale and non-shale loans.

Figure 9: % Loans Local Lenders - % Loans National Lenders (Shale v. Non-Shale)

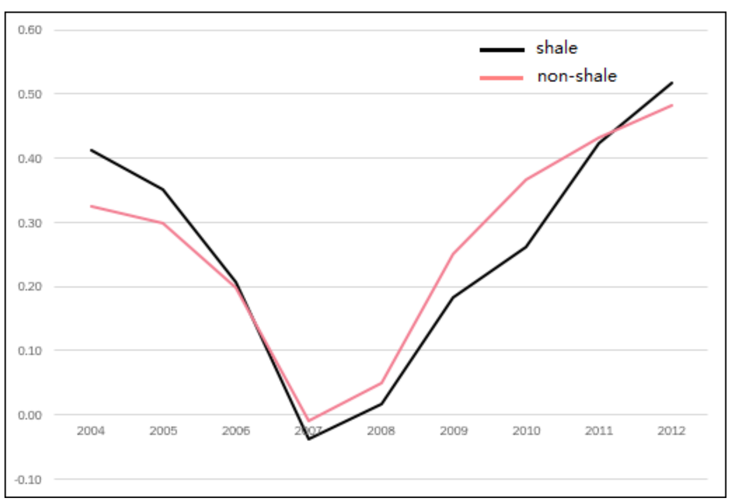
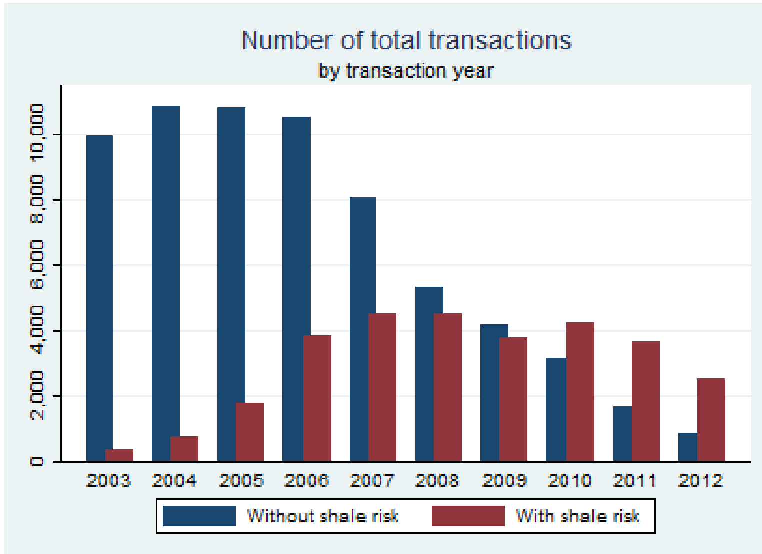


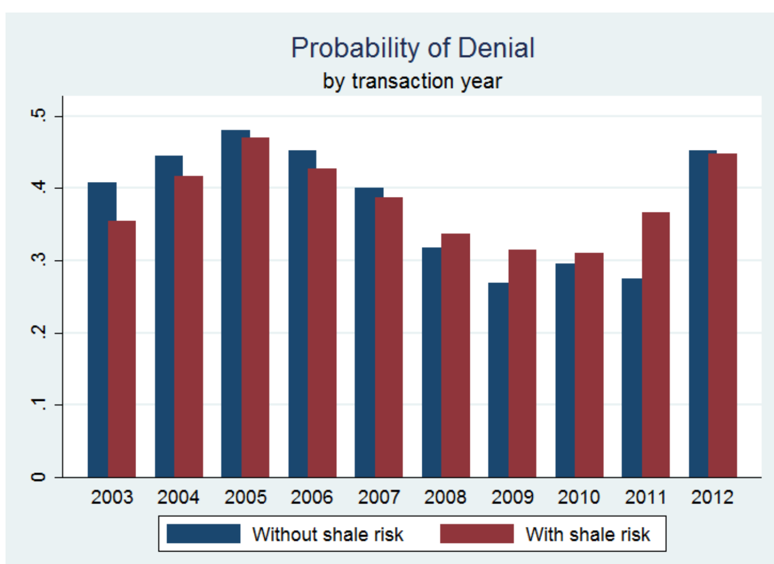
Figure 10 shows that, as shale development became more pervasive in Tarrant Co., loans with shale risk (unsurprisingly) outpaced those without.

Figure 10: Number of Total Transactions



For the remainder of the paper, we focus on loan pricing, but it is worthwhile to look briefly at whether shale gas exposure led to loan applications being rejected altogether. Figure 11 looks at the probability of loan *denials*. We see that, early in the period, a greater likelihood of denial was associated with a non-shale risk loan, while later in the period, the opposite is true. We are not, however, able to disentangle whether this is the result of lenders becoming more wary of shale gas or simply that there were more homes with shale gas exposure in the pool of mortgage applications. We therefore focus our attention on loan pricing, where we are able to differentiate between these two effects.

Figure 11: Probability of Denials



Water Source & Shale Exposure. Finally, many of our conclusions will focus on the house's water source (i.e., piped v. private groundwater well). We demonstrate in Table **??** that groundwater dependent houses are larger, newer, on bigger plots, and are occupied by higher-income individuals on average. Minority households tend to be located in more urban areas accessing district water. Loans issued to purchase houses located in the non-district water regions and with shale exposure tend to more often be securitized, and comparing whether a house is exposed to shale, those loans are often insured through the Federal Housing Administration.

Table 6: Summary Statistics by Water Source

	District Water		Non-district Water	
	Mean	Std. Dev.	Mean	Std. Dev.
House Characteristics				
Beds	3.209	(0.608)	3.433	(0.643)
Baths	1.952	(0.533)	2.185	(0.59)
Living (sqft)	1930.08	(724.53)	2221.65	(834.68)
Land (sqft)	8411.27	(5617.60)	11313.29	(9196.77)
Age	23.61	(25.26)	16.52	(16.02)
Shale Exp. (wells)	0.571	(1.488)	0.583	(1.857)
Household Characteristics				
White	0.790	(0.407)	0.823	(0.382)
Hispanic	0.248	(0.432)	0.164	(0.37)
Black	0.101	(0.301)	0.08	(0.271)
Asian	0.041	(0.199)	0.054	(0.226)
Income	74263.01	(52726.08)	84748.03	(54584.59)
Transaction & Loan Characteristics				
High Interest Loan	0.119	(0.324)	0.095	(0.294)
Loan-to-Income	2.062	(0.787)	2.108	(0.771)
Sale Amount	122284.7	(64883.77)	150010.1	(78770.94)
Loan Value	130745.5	(78082.72)	158923.7	(92549.31)
Annualized L2I	0.165	(0.057)	0.167	(0.058)
Foreclosure	0.085	(0.279)	0.074	(0.262)
Fixed Loan	0.906	(0.292)	0.891	(0.311)
FHA Loan	0.328	(0.469)	0.277	(0.448)
Gov't Sec.	0.319	(0.466)	0.351	(0.477)
Comm. Sec.	0.512	(0.5)	0.509	(0.500)
Obs.	69,937		124,361	

[a] Summarizes the primary variables used to describe the house, household, sales, and loan characteristics. The first and second columns compare the mean and standard deviations across properties likely to access district water and those likely accessing groundwater.

Table 7 summarizes the data across shale and non-shale exposed regions. The houses are slightly larger in areas with slightly greater levels of shale exposure. They are also more expensive when not located near shale, yet the loan values are lower.

Table 7: Summary Statistics by Shale Exposure

	No Shale Wells		Shale Wells	
	Mean	Std. Dev.	Mean	Std. Dev.
House Characteristics				
Beds	3.345	(0.643)	3.406	(0.619)
Baths	2.094	(0.593)	2.161	(0.488)
Living (sqft)	2098.952	(812.97)	2262.779	(776.325)
Land (sqft)	10136.71	(7612.469)	11129.84	(11610.4)
GW	0.657	(0.475)	0.552	(0.497)
Age	20.329	(20.393)	12.063	(16.666)
Shale Exp. (wells)			3.53	(2.811)
Household Characteristics				
White	0.805	(0.396)	0.84	(0.367)
Hispanic	0.199	(0.399)	0.174	(0.379)
Black	0.085	(0.278)	0.10	(0.299)
Asian	0.049	(0.216)	0.05	(0.218)
Income	80831.17	(54728.11)	82839.85	(52207.8)
Transaction & Loan Characteristics				
High Interest Loan	0.095	(0.293)	0.108	(0.31)
Loan-to-Income	2.067	(0.767)	2.203	(0.815)
Sale Amount	141524.7	(77181.46)	133294.7	(65010.93)
Loan Value	147380.7	(91162.74)	156983.1	(76204.43)
Annualized L2I	0.166	(0.058)	0.165	(0.057)
Foreclosure	0.072	(0.259)	0.109	(0.311)
Fixed Loan	0.893	(0.31)	0.936	(0.246)
FHA Loan	0.275	(0.447)	0.389	(0.487)
Gov't Sec.	0.34	(0.474)	0.337	(0.473)
Comm. Sec.	0.502	(0.5)	0.547	(0.498)
Obs.	161,030		36786	

[a] Summarizes the primary variables used to describe the house, household, sales, and loan characteristics. The first and second columns compare the mean and standard deviations across households with and without shale exposure.

5 Theory

The theory for analyzing the decisions made by lenders with respect to various risks and lending rates is similar to that used to describe the trade-offs made by workers choosing amongst risky jobs. This is the idea behind *wage hedonics*, where the compensating differentials associated with different job attributes are used to measure their values (Viscusi 1993). The idea even appears in Adam Smith's seminal work (Smith 2015), where risky or unpleasant jobs are noted as commanding a premium. The problem has a simple graphical interpretation, which is described in the following figure. Firms are described by a series of iso-profit curves, denoted by π_A and π_B. Along each curve, a lower risk must be accompanied by a lower wage in order for a constant level of profit to be maintained. Firm A is better at providing a low risk environment and is able to pay a higher wage when risks are low compared with firm B.

Workers face similar trade-offs, which are described by iso-expected utility curves EU_0 and EU_1. In particular, each worker is willing to accept a higher risk in exchange for a higher wage payment, with their willingness to do so, or marginal rate of substitution between risk and compensation, being summarized by the slope of the iso-expected utility curve.

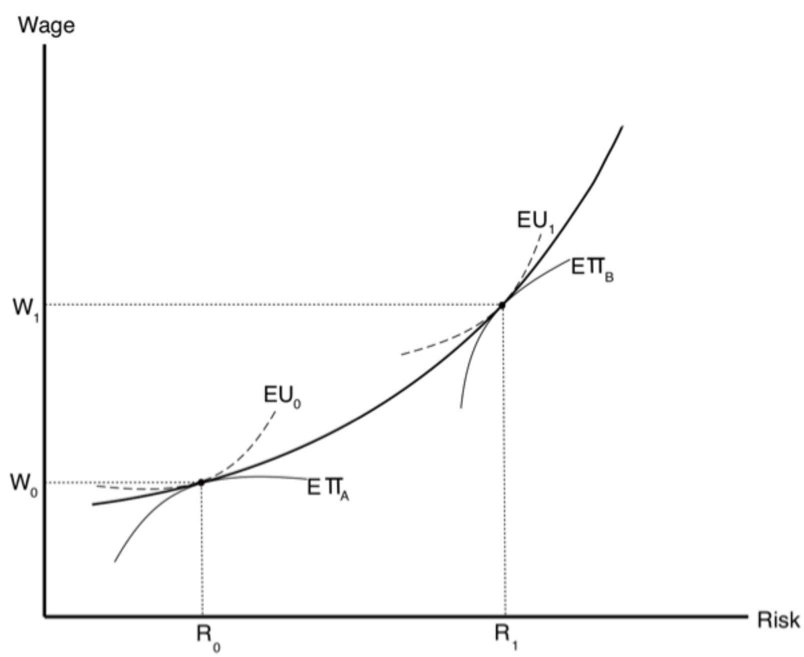


Figure 12: Iso-expected Utility Curves

As in other hedonic applications, the allocation of workers to jobs involves a sorting process whereby workers match with firms that yield them the highest expected utility, given firms' wage offers described by their iso-profit curves. In particular, the best that worker #0 (who requires relatively little in terms of compensation in exchange for taking on more risk) can do is to match with firm A, which requires relatively small increases in wages in exchange for more risk in order to hold profits constant. Worker #0 ends up netting (R_0, W_0) in hedonic equilibrium. Worker #1 requires more compensation in exchange for taking on additional risk (i.e., a steeper indifference curve) and ends up choosing to match with firm B, yielding (R_1, W_1). With continuously distributed workers and firms, the hedonic wage function describes the set of tangency points between the iso-profit functions of firms and iso-expected utility functions of workers.

We use $U(W)$ to represent the utility from wage W if a worker is in the healthy state and $V(W)$ to represent the utility from the same wage in an injured state; p represents the probability of that injury. EU represents expected utility:

$$EU = (1-p)U(W) + pV(W)$$

Taking the total differential of expected utility yields:

$$(1-p)U'(W)dW + pV'(W)dW - U(W)dp + V(W)dp = 0$$

With some re-arranging, we derive an expression for the slope of an iso-expected utility curve:

$$dW[(1-p)U'(W) + pV'(W)] + dp[V(W) - U(W)] = 0$$

$$\left.\frac{\mathrm{d}W}{\mathrm{d}p}\right|_{dEU=0} = -\frac{EU_p}{EU_W} = \frac{U(W)-V(W)}{(1-p)U'(W)+pV'(W)}$$

EU will be positively sloped if $U(W) - V(W) > 0$ (i.e., utility from a given wage is greater in the non-injured state) and if $U'(W), V'(W) > 0$ (i.e., utility is increasing in wages regardless of the injury state). 2nd order conditions require that curves have convexity as illustrated. Wage hedonic techniques use the slope of estimated relationship between risk and wage to recover $\frac{dW}{dp}|_{dEU=0}$.

In the case of banks issuing mortgages, there are two forms of risk that we consider: (i) shale (p_1) and (ii) income (p_2). Shale risk refers to the list of reasons that shale development might put a house into technical default according to the standard mortgage guidelines described in the introduction. Foremost is the risk of any outcome that might detract from the house's value as a residential structure. Income risk refers to the standard risk that liquidity constrained individuals will suffer income shocks that make it impossible for them to make timely mortgage payments.

Shale risk and income risk combine to create default risk, which is ultimately what concerns the lender. Perceived default risk faced by lender i is $D_i = D_i(p_1, p_2)$, where $0 \leq D_i \leq 1 \; \forall \, i$ and $\frac{\partial D_i}{\partial p_j} \geq 0 \; \forall \, i$ and for $j = 1, 2$. Assume that $D_i(p_1, p_2)$ is quasi-convex.

Firm i's expected profit from making a loan with risks p_1 and p_2 is given by:

$$E\Pi_i = (1 - D_i(p_1, p_2))\pi_i^G(\sigma) + D_i(p_1, p_2)\pi_i^B$$

where $\pi_i^G(\sigma)$ measures the profit associated with the 'good' state (i.e., in which there is no foreclosure) when the mortgage rate is σ (i.e., the actual rate or an indicator for subprime or not). $\pi_i^B(\sigma)$ measures expected profits in the bad state in which a foreclosure takes place and the lender takes control of the property, selling it for a loss. Taking the total derivative of expected profit with respect to the two forms of risk and rearranging, we can find the change in p_2 associated with a small change in p_1 that would hold $E\Pi_i$ fixed.

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}p_j}\right|_{dE\Pi_i=0} = \frac{\frac{\partial E\Pi_i}{\partial p_j}}{\frac{\partial E\Pi_i}{\partial \sigma}} = \frac{\frac{\partial D_i}{\partial p_j}[\pi_i^G(\sigma) - \pi_i^B(\sigma)]}{(1-D_i(p_1,p_2))\pi_j^{G\prime}(\sigma) + D_i(p_1,p_2))\pi_j^{B\prime}(\sigma)} \qquad j = 1, 2$$

$$\left.\frac{\frac{\mathrm{d}\sigma}{\mathrm{d}p_1}}{\frac{\mathrm{d}\sigma}{\mathrm{d}p_2}}\right|_{dE\Pi_i=0} = \frac{\frac{\frac{\partial E\Pi_i}{\partial p_1}}{\frac{\partial E\Pi_i}{\partial \sigma}}}{\frac{\frac{\partial E\Pi_i}{\partial p_2}}{\frac{\partial E\Pi_i}{\partial \sigma}}} = \frac{\frac{\partial E\Pi_i}{\partial p_1}}{\frac{\partial E\Pi_i}{\partial p_2}}$$

This final term represents the negative of the slope of the iso-expected profit curve. As such, we can recover firm i's willingness to take on additional income risk in exchange for a one-unit reduction in shale risk by taking the ratio of the two hedonic gradients. Under our assumptions, this will be invariant to how we define the lending rate, and we can simply use a probability of subprime rather than an exact rate, which is useful because we do not see the exact rate unless it is a subprime.[32] We therefore learn about the lender's willingness to trade-off one type of risk for another holding the overall perceived default risk constant.

6 Empirical Model

6.1 Probability of Subprime — Local Logit Regression

Using a binary indicator variable for a high-priced mortgage as our proxy for subprime, we would like to measure the way in which the likelihood of a subprime mortgage varies with different risk variables without imposing a great deal of structure. Following on the model of Bajari and Benkard (2005), we allow the data to speak to the shape of this equilibrium hedonic function (in two different dimensions of risk) and recover a flexible representation of lender preferences that characterizes the distribution of heterogeneity. We then explore how these preferences vary over time and with water source.

Parametric regression models (such as the probit and logit) are commonly used to study binary dependent variables, but these models impose restrictive functional form assumptions. Semi-parametric binary choice estimators (single-index models) relax these restrictions but they effectively reduce the heterogeneity in the X characteristics to a single dimension. This restricts the interaction between covariates – specifically, the ratio of two marginal effects does not depend on X in the single-index model. Because we are interested in heterogeneity in the tradeoffs between two types of risk (income risk and shale risk), we instead follow Frölich (2006) to perform non-parametric regression for binary dependent variables (local likelihood logit estimation).

The local likelihood logit estimator is:

$$\hat{E}[Y \mid X = x] = \frac{1}{1 + e^{-x'\hat{\theta}_x}}$$

where

$$\hat{\theta}_x = \arg\max_{\theta_x} \sum_{i=1}^{n} \Big(Y_i \ln\Big(\frac{1}{1 + e^{-X_i'\theta_x}}\Big) + (1 - Y_i) \ln\Big(\frac{1}{1 + e^{X_i'\theta_x}}\Big)\Big) K_H(X_i - x)$$

[32]Since 2004, the Federal Reserve Board required lenders to collect and report the spread between the annual percentage rate (APR) on a loan and the yield on Treasury securities of comparable maturity if the spread is equal to or greater than 3.0 percentage points for a first-lien loan. In December 2008, the Board published an amendment to this rule which requires the lenders to report the spread if it is equal to or greater than 1.5 percentage points for a first-lien loan. In order to make our estimates consistent across years, we define the threshold to be 3.0 percentage points across all our sample periods.

The regressors include shale risk (defined as cumulative hazard rate), income risk (defined as the debt-to-income ratio), housing characteristics (lot size, house size, number of bathroom, number of bedroom and year built of a property), loan attributes (yield of 30-year treasury bond, and whether the rate is fixed or not), and year dummies.

The kernel weight $K_H(X_i - x)$ is computed as:

$$K_{h,\delta,\lambda}(X_i - x) = \prod_{q=1}^{q_1} \kappa\Big(\frac{X_{q,i} - x_q}{h}\Big) \prod_{q=q_1+1}^{q_2} \delta^{|X_{q,i} - x_q|} \prod_{q=q_2+1}^{Q} \lambda^{\mathbb{1}(X_{q,i} \neq x_q)}$$

The kernel function measures the distance between X_i and x for each variable through one of three components, depending upon the particular type of variable: continuous regressors (the first term), ordered discrete regressors (the second term) and unordered discrete regressors (the third term). In our application, continuous regressors include cumulative hazard rate, debt-to-income ratio, lot size, house size, and yield of 30-year treasury bond; ordered regressors include number of bathrooms, number of bedrooms, and the age of each property; non-ordered regressors include year dummies, and whether the rate is fixed or not.

Values for the bandwidth and hyper-parameters (h, δ, λ) in the kernel function are obtained from cross-validation, and the cross-validation criterion is based on maximizing the leave-one-out fitted likelihood function:

$$CROSSVAL(h, \lambda, \delta) = \sum_{i=1}^{n} Y_i \ln g(X_i, \hat{\theta}_{-X_i|h,\delta,\lambda}) + (1 - Y_i) \ln(1 - g(X_i, \hat{\theta}_{-X_i|h,\delta,\lambda}))$$

where $g(\cdot)$ is our local likelihood logit estimator.

We are interested in recovering banks' willingness to trade income risk for shale risk, which is revealed in the ratio of the two hedonic gradients, where

$$\rho(x) = \frac{\frac{\partial \hat{E}[Y|X=x]}{\partial x_1}}{\frac{\partial \hat{E}[Y|X=x]}{\partial x_2}} = \frac{\hat{\theta}_1(x)}{\hat{\theta}_2(x)}$$

Unlike the single index models, this ratio is a flexible function of the regressors, x. The negative of this ratio defines the negative of the slope of the expected iso-profit curve drawn in (shale risk, income risk)-space. Figure 13 illustrates a simple case with linear iso-profit curves.

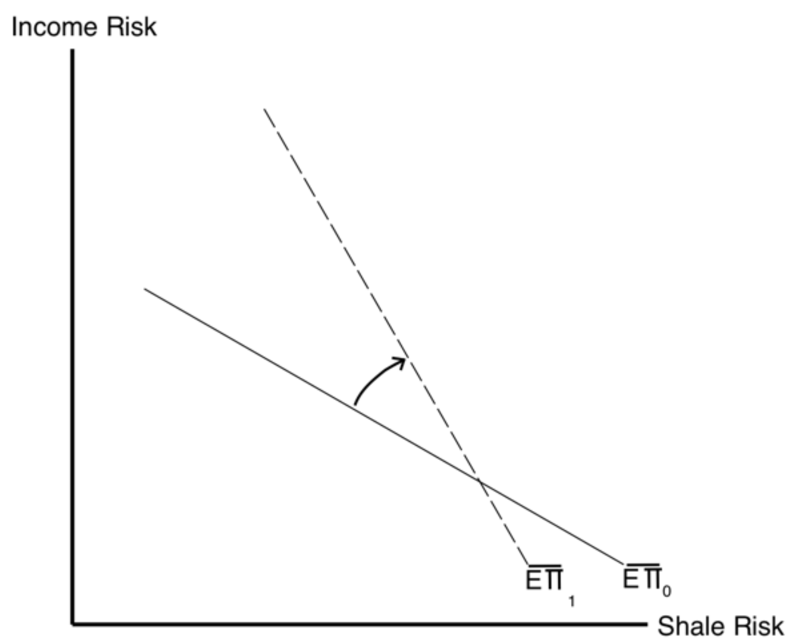


Figure 13: Iso-profit Curves

A higher value suggests that iso-curve has become steeper, implying that lenders require a larger reduction in income risk in order to accept another unit of shale risk, holding expected profits constant.

7 Results

7.1 Simple Logit Analysis

We begin with a simple logit specification to explore the determinants of subprime status. In particular, we model the likelihood of subprime as a function of housing attributes, loan attributes (fixed or variable rate, and yield curve of 30-yr treasure bond), year dummies, the debt-to-income ratio (i.e., income risk), and the cumulative hazard rate (shale risk). We estimate this model separately for two time periods – preceding and following the financial crisis (2003-2008 and 2010-2012). We expect that relative concerns of lenders over various sources of default risk might have changed after the financial crisis; in particular, the discussion in Section 3 suggests that policy makers placed increasing attention on shale risk beginning in 2010. While estimates based on the simple logit are generally insignificant, point estimates do behave as expected. Focusing on columns (1) and (2) of Table 8, the parameters on both forms of risk both increase in magnitude following the financial crisis, with that on shale risk becoming statistically significant. The lender's average willingness to trade-off income risk for shale risk (i.e., the ratio of the shale risk to income risk parameters) rises substantially.

We next differentiate additionally by water source (i.e., groundwater v. piped water). Muehlenbachs et al. (2015) demonstrated that housing markets are particularly prone to capitalize the risks of nearby shale gas development when houses are dependent upon groundwater. We may therefore expect lenders to be particularly aware of shale risk when houses are groundwater dependent, and that this concern may have grown after the financial crisis. We find that this is indeed the case – we recover a large and statistically significant increase in the parameter on shale risk for groundwater dependent properties following the financial crisis, suggesting a large increase in lenders' willingnesses to take on additional income risk to avoid shale risk on these properties. The same is not true of houses that rely on piped water, where the ratio of income risk to shale risk actually goes down (although all piped water house risk parameters are statistically insignificant).

Specification 2 with drill_1k

Specification 3 with minimum distance to producing well

While the simple logit results suggest that lenders may have become increasingly concerned about shale risk following the financial crisis, and that this was particularly true for groundwater dependent houses, the strong functional form restrictions placed on the model by the simple logit may constrain our ability to learn about these relationships. In the following subsection, we relax these constraints by employing a flexible local logit specification.

7.2 Local Logit Estimation

We estimate the flexible local logit specification paying particular attention to the role of water source. In particular, we use estimated preference ratios $\rho(x)$ to illustrate lenders' indifference curves in (shale risk, income risk) space. We begin by taking the distribution of the ratios of

Table 8: Simple Logit Regression Results

	Logit Model period 1: 2001-2007	Logit Model period 2: 2010-2016	Tobit Model 2001-2007	Tobit Model 2010-2016
Cumulative hazard rate	0.0766***	0.107***	0.0413***	0.0817***
	(6.02)	(4.71)	(5.07)	(4.96)
Debt-to-income	0.196***	0.0611**	0.123***	0.0517**
	(17.20)	(2.72)	(16.78)	(3.15)
FHA loans	-3.242***	1.860***	-1.939***	1.254***
	(-49.22)	(42.05)	(-59.03)	(37.04)
VA loans	-5.341***	-3.041***	-2.896***	-1.810***
	(-11.92)	(-8.54)	(-17.58)	(-10.99)
Securitized loans	1.156***	-0.158***	0.745***	-0.222***
	(58.38)	(-4.28)	(58.03)	(-8.11)
Yield of 30 year treasury bond	-0.258***	0.0922	-0.191***	0.00940
	(-6.02)	(1.64)	(-7.04)	(0.24)
Number of bedrooms	0.343***	0.186***	0.208***	0.144***
	(17.74)	(4.61)	(17.01)	(4.96)
Number of bathrooms	-0.294***	-0.0664	-0.181***	-0.0360
	(-11.83)	(-1.27)	(-11.65)	(-0.97)
Living size	-0.000474***	-0.000616***	-0.000294***	-0.000440***
	(-24.56)	(-14.30)	(-24.50)	(-14.53)
Land size	-0.0134***	-0.00174	-0.00827***	-0.000414
	(-9.06)	(-0.55)	(-9.15)	(-0.19)
Age of property	0.00816***	0.00650***	0.00527***	0.00559***
	(15.06)	(6.44)	(15.14)	(7.56)
_cons	-6.367***	-5.140***	-3.162***	-3.581***
	(-13.04)	(-7.73)	(-14.37)	(-8.28)
sigma				
_cons			1.203***	1.482***
			(161.03)	(74.03)
N	131,489	51,198	131,548	51,264
Year FE	Y	Y	Y	Y
City FE	Y	Y	Y	Y

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Simple Logit Regression Results (specification 2)

	Logit Model period 1: 2001-2007	Logit Model period 2: 2010-2016	Tobit Model 2001-2007	Tobit Model 2010-2016
Number of producing wells within 2000m	0.00264	0.00232	0.00158	0.00172
	(1.53)	(1.58)	(1.43)	(1.59)
Debt-to-income	0.196***	0.0645**	0.123***	0.0545***
	(17.23)	(2.88)	(16.84)	(3.32)
FHA loans	-3.252***	1.865***	-1.945***	1.259***
	(-49.38)	(42.19)	(-59.21)	(37.15)
VA loans	-5.356***	-3.039***	-2.907***	-1.812***
	(-11.95)	(-8.53)	(-17.62)	(-10.98)
Securitized loans	1.161***	-0.157***	0.748***	-0.220***
	(58.67)	(-4.25)	(58.29)	(-8.06)
Yield of 30 year treasury bond	-0.257***	0.0899	-0.190***	0.00705
	(-6.00)	(1.60)	(-7.02)	(0.18)
Number of bedrooms	0.343***	0.189***	0.209***	0.147***
	(17.78)	(4.68)	(17.06)	(5.03)
Number of bathrooms	-0.297***	-0.0715	-0.182***	-0.0408
	(-11.94)	(-1.37)	(-11.76)	(-1.10)
Living size	-0.000475***	-0.000620***	-0.000295***	-0.000443***
	(-24.63)	(-14.38)	(-24.55)	(-14.62)
Land size	-0.0132***	-0.00165	-0.00816***	-0.000319
	(-8.89)	(-0.52)	(-9.01)	(-0.15)
Age of property	0.00734***	0.00549***	0.00486***	0.00480***
	(13.79)	(5.43)	(14.16)	(6.49)
_cons	-6.353***	-5.001***	-3.159***	-3.479***
	(-13.02)	(-7.53)	(-14.34)	(-8.04)
sigma				
_cons			1.203***	1.483***
			(161.03)	(74.02)
N	131,489	51,198	131,548	51,264
Year FE	Y	Y	Y	Y
City FE	Y	Y	Y	Y

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Simple Logit Regression Results (specification 3)

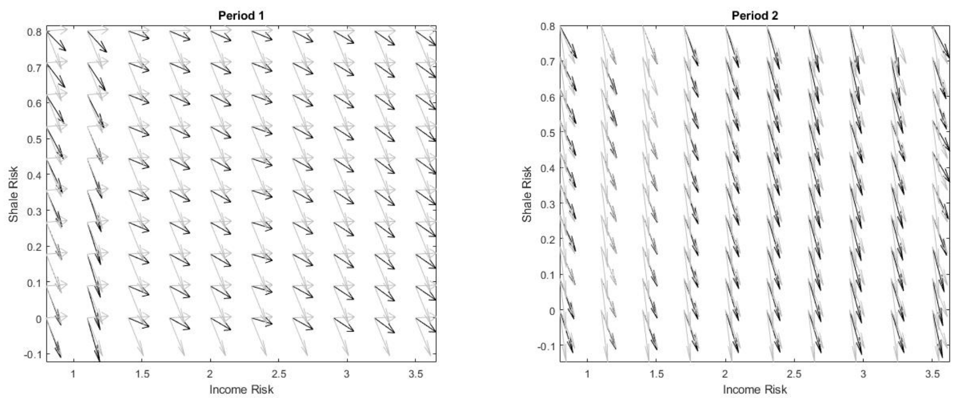
	Logit Model period 1: 2001-2007	Logit Model period 2: 2010-2016	Tobit Model 2001-2007	Tobit Model 2010-2016
Distance to the nearest producing well	0.000000768	-0.0000490	-0.000000134	-0.0000322
	(0.23)	(-1.70)	(-0.06)	(-1.65)
Debt-to-income	0.195***	0.0640**	0.123***	0.0541***
	(17.13)	(2.85)	(16.76)	(3.30)
FHA loans	-3.265***	1.864***	-1.959***	1.258***
	(-49.20)	(42.16)	(-59.20)	(37.12)
VA loans	-5.352***	-3.039***	-2.899***	-1.812***
	(-11.94)	(-8.53)	(-17.63)	(-10.97)
Securitized loans	1.163***	-0.156***	0.751***	-0.220***
	(58.67)	(-4.23)	(58.48)	(-8.04)
Yield of 30 year treasury bond	-0.255***	0.0920	-0.188***	0.00913
	(-5.93)	(1.64)	(-6.88)	(0.23)
Number of bedrooms	0.343***	0.188***	0.210***	0.146***
	(17.76)	(4.66)	(17.10)	(5.00)
Number of bathrooms	-0.298***	-0.0705	-0.184***	-0.0397
	(-11.97)	(-1.35)	(-11.85)	(-1.07)
Living size	-0.000477***	-0.000620***	-0.000296***	-0.000443***
	(-24.66)	(-14.38)	(-24.63)	(-14.61)
Land size	-0.0131***	-0.00152	-0.00823***	-0.000200
	(-8.83)	(-0.48)	(-9.08)	(-0.09)
Age of property	0.00713***	0.00552***	0.00478***	0.00480***
	(13.43)	(5.46)	(14.00)	(6.50)
_cons	-0.852***	-4.911***	-0.416**	-3.419***
	(-3.36)	(-7.41)	(-2.59)	(-7.92)
sigma				
_cons			1.199***	1.483***
			(160.99)	(74.02)
N	87,583	51,198	87,623	51,264
Year FE	Y	Y	Y	Y
City FE	Y	Y	Y	Y

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

debt-to-income (i.e., income risk) and shale risk, as it was defined earlier in the paper. These ratios represent the slopes of lenders' "indifference curves", drawn in the space of the two types of risks. We then calculate the kernal weighted averages of all estimated preference ratios at each point in risk space, where the smoothing parameter is set using Silverman's rule. The subsequent mapping describes the indifference curve map at every point in risk space. In Figure 14 Black arrows are used to describe mean estimates, while the gray arrows describe the 5% and 95% confidence bands calculated from 10 bootstrap iterations.

Figure 14: Indifference Curve Map



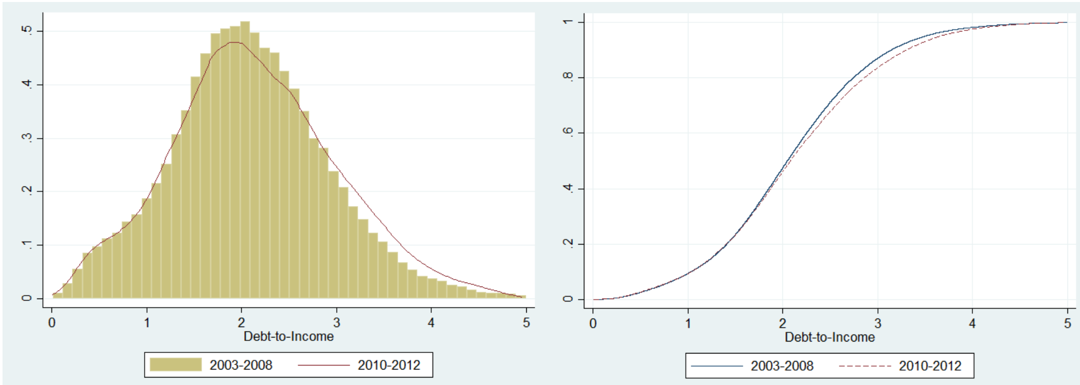
A number of interesting features appear immediately. First, in the period before the financial crisis, the slope of the indifference curves are indistinguishable from zero. Assuming income risk is always considered by lenders to be a "bad," this would imply that shale risk was not taken into consideration by banks in the earlier period. The right panel of figure 14 shows lenders' indifferene curves are now negatively sloped for those houses, and those slopes are statistically significant. The results shows that lenders are willing to take on significantly more income risk to avoid additional shale gas risk in the post-crisis period.

7.3 Robustness Check

Here we address some concerns that may generate alternative interpretations of our results.

One alternative story is that home buyers are likely to overstate their incomes in the period prior to the financial crisis, leading to an under-reporting of debt-to-income ratios , so the debt-to-income ratio in the first period is downward biased in our data set. To address this concern, below is a comparison between the distribution of debt-to-income ratio between the first and second period. Although the distribution in the pre-crisis period does have a bit more mass at the lower values of the debt-to-income ratio, the differences are not large.

Figure 15: Distribution of Debt-to-Income Ratio



8 Valuing Shale Risk

To value the willingness to pay to avoid shale risk, through taking on additional income risk by lending to individuals with higher debt-to-income ratios, we use transaction and foreclosure information from Dallas County. Dallas neighbors Tarrant County and has comparable land and population sizes and population density, yet Dallas does not have active oil and natural gas drilling.[33] Figure 16 describes the rate of foreclosure by county across time, and Figure 17 describes foreclosure rates for homes with well exposure and not for Tarrant and Denton counties. Perhaps interestingly, shale exposed areas of Tarrant county experience the peak foreclosure period in 2011 rather than 2008 like the non-exposed regions including Dallas county. The year 2011 is during a period of volatile natural gas price, which likely affected the streams of royalty incomes for individuals with active oil and natural gas leases.

Our "back-of-the-envelope" calculation combines the estimated relationships between the debt-to-income ratio and rate of foreclosure, as described in the upper panel of Table 11,[34] and between the foreclosure rate and transaction value, as described in the lower panel of Table 11. In particular, we observe that a standard deviation increase in the debt-to-income ratio increases the likelihood that a home is foreclosed in the future by 0.129 in Dallas, and a foreclosed home sells for 25% percent less in Dallas (using the -0.28 coefficient in the lower panel of Table 11), controlling for other house observable characteristics and city and year fixed effects. Using the logarithm of the mean sales value for single family, residential homes located in Dallas County,

[33] In particular, Dallas county is comprised of 909 square miles, 2.5 million people, and a population density of 2,950 individuals per square mile. Tarrant county, on the other hand, is comprised of 902 square miles, 1.8 million people, and a population density of 2,095 individuals per square mile.

[34] Table 11 also includes estimates using data from Tarrant County to demonstrate similarity across counties. However, note that debt-to-income has a smaller effect on the likelihood of foreclosure and foreclosed homes sell for more, on average, than those located in Dallas.

we calculate the willingness to pay to avoid shale risk using the following formula:

$$\begin{aligned}WTP &= \frac{\partial V_0}{\partial foreclosure} P(foreclosure|X) \\ &= exp(11.12) - exp(11.12 - 0.129 * 0.28) = 2,394.9\end{aligned}$$

where V_0 is the mean value of a property sold in Dallas County within the sample period. The estimates from Dallas County suggest that an additional unit of shale risk is valued at $2,394.9, which is roughly 1.4% of the mean sales value among all homes sold in Dallas County during the sample period spanning 2003 to 2012.

In order to scale the willingness to pay, we compare that value to mean value earned on an average mortgage issued in Dallas. We approximate that the average mortgage interest rate is roughly 6.7% and the average loan amount is $151,476. Further, most of the loans in Dallas have 30 year term lengths and lenders earn around $190,513 on the average loan in Dallas. Our willingness to pay comprises roughly 1.2% of the profit earned on an average Dallas mortgage.

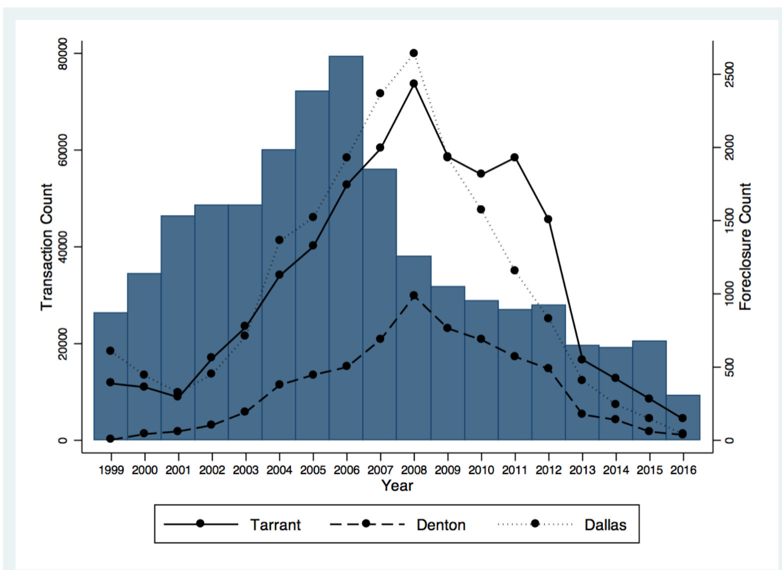


Figure 16: Foreclosures by County

[a] Count of foreclosed sales by year and county (right y-axis) layered on the count of total transactions (left y-axis).

9 Conclusion

This paper explores the housing market impacts of shale gas development. Previous work has done so using data on the capitalization of shale gas activities into housing prices and hedonic theory to give those estimates a welfare interpretation. In this paper, we come at the question

Table 11: Simple Foreclosure Logit & Hedonic

	Tarrant			Dallas	
Dependent Variable: Loan Foreclosed Dummy					
Loan-to-Income	0.23083***	0.19467***	0.19624***	0.12853***	0.11322***
	(0.01035)	(0.01134)	(0.01139)	(0.01002)	(0.01094)
Shale Exposure			-0.00823		
			(0.00767)		
Yield Curve	0.06523**	0.08393***	0.07781***	0.02829	0.05517*
	(0.02841)	(0.02860)	(0.02872)	(0.02909)	(0.02930)
FHA Loan		0.75802***	0.76085***		0.83941***
		(0.01913)	(0.01917)		(0.01924)
VA Loan		0.45506***	0.45263***		0.50779***
		(0.03615)	(0.03635)		(0.04954)
Gov't Secured		-0.27002***	-0.26930***		-0.28448***
		(0.02658)	(0.02668)		(0.02607)
Comm. Secured		-0.02607	-0.02551		-0.01978
		(0.02447)	(0.02457)		(0.02324)
Owner Occupied		-0.31868***	-0.30640***		-0.47901***
		(0.03641)	(0.03688)		(0.03349)
Constant	-2.65700***	-2.75097***	-2.75264***	-5.32625***	-5.06499***
	(0.18941)	(0.19248)	(0.19351)	(0.41727)	(0.41918)
Observations	216,819	216,819	213,986	192,206	191,599
Year FE	x	x	x	x	
City	x	x	x	x	
Dependent Variable: Log Sales Values					
Foreclosure	-0.22386***	-0.21760***	-0.21704***	-0.28039***	-0.26252***
	(0.00221)	(0.00215)	(0.00215)	(0.00255)	(0.00244)
Shale Exposure			-0.00502***		
			(0.00039)		
FHA Loan		-0.18444***	-0.18446***		-0.28621***
		(0.00142)	(0.00143)		(0.00172)
VA Loan		-0.13082***	-0.13113***		-0.19456***
		(0.00220)	(0.00222)		(0.00365)
Gov't Secured		0.02096***	0.02067***		0.02071***
		(0.00212)	(0.00212)		(0.00245)
Comm. Secured		0.05191***	0.05212***		0.05250***
		(0.00197)	(0.00197)		(0.00226)
Owner Occupied		0.13145***	0.13187***		0.25093***
		(0.00257)	(0.00257)		(0.00311)
Beds	-0.05912***	-0.05467***	-0.05507***	-0.09277***	-0.08058***
	(0.00141)	(0.00135)	(0.00136)	(0.00164)	(0.00155)
Baths	0.10867***	0.10132***	0.10150***	0.10123***	0.09249***
	(0.00185)	(0.00179)	(0.00180)	(0.00172)	(0.00162)
Living Size	0.00038***	0.00035***	0.00035***	0.00049***	0.00044***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Land Size	0.00001***	0.00001***	0.00001***	-0.00000***	-0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Age	-0.00214***	-0.00245***	-0.00252***	0.00238***	0.00185***
	(0.00005)	(0.00005)	(0.00005)	(0.00005)	(0.00005)
Constant	11.22294***	11.15283***	11.15097***	11.12650***	10.99325***
	(0.07326)	(0.05717)	(0.05597)	(0.01310)	(0.01359)
Observations	206,601	206,139	204,196	222,065	221,343
R-squared	0.683	0.706	0.707	0.639	0.679
City FE	x	x	x	x	x
Year FE	x	x	x	x	x

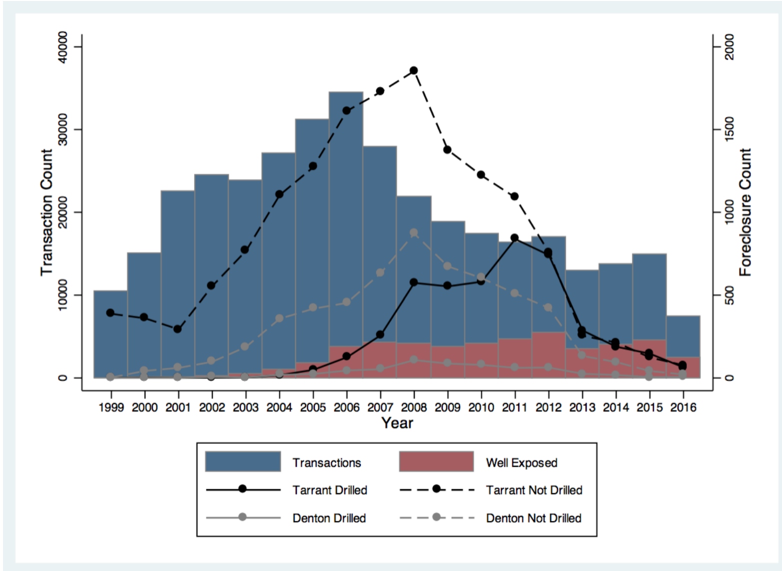


Figure 17: Foreclosures by Shale Exposure

[a] Count of foreclosed sales by year and whether the property is exposed to drilling activity in Tarrant county (right y-axis) and layered on the count of total transactions (blue) and shale exposed transactions (red) each year (right y-axis).

from a different perspective, focusing instead on the pricing decisions of mortgage lenders. Mortgage lenders bear the risk that borrowers will default – a fact highlighted by a housing market bust and financial collapse less than a decade ago. These swings in the housing market coincided with the U.S. shale boom. By 2010, lenders and policy-makers had become much more mindful of these risks and began to recognize the role that shale gas development might play in default.

We estimate lenders relative preferences for two different types of risk – shale and income. We find strong evidence that lenders became more concerned with shale risk by the period following the financial crisis, and that this is particularly true when we differentiate by the source of houses' water supply. We might suspect that shale development would place the residential value of a groundwater-dependent property at greater risk. This indeed appears to be the case, with lenders showing a relative distaste for shale risk on loans to these properties.

These results confirm the conclusions of previous research with respect to the expected house price impacts of shale gas development for groundwater homes (Muehlenbachs et al. 2015). Importantly, they do so based on the decisions of agents (i.e., lenders) who make hundreds (possibly thousands) of lending decisions each year. One might worry that homebuyers in hedonic models might lack expertise – the typical individual may only buy a couple of houses over the course of her lifetime, and the cost of information acquisition is high. The same is not true for professional lenders.

References

Albanesi, Stefania, Giacomo DeGiorgi, and Nosal Jaromir, "Credit Growth and the Financial Crisis: A New Narrative," *NBER Working Paper*, 2017.

Avery, Robert B, Kenneth P Brevoort, and Glenn B Canner, "Higher-Priced Home Lending and the 2005 HMDA Data (Table 8 Revised September 18, 2006)," *Fed. Res. Bull. A123*, 2006, *92.*

Bajari, Patrick and C Lanier Benkard, "Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach," *Journal of political economy*, 2005, *113* (6), 1239–1276.

Bartik, Alexander Wickman, Janet Currie, Michael Greenstone, and Christopher R Knittel, "The local economic and welfare consequences of hydraulic fracturing," 2016.

Black, Harold, Robert L Schweitzer, and Lewis Mandell, "Discrimination in mortgage lending," *The American Economic Review*, 1978, *68* (2), 186–191.

Boslett, Andrew, Todd Guilfoos, and Corey Lang, "Valuation of expectations: A hedonic study of shale gas development and New York's moratorium," *Journal of Environmental Economics and Management*, 2016, *77*, 14–30.

Boxall, Peter C, Wing H Chan, and Melville L McMillan, "The impact of oil and natural gas facilities on rural residential property values: a spatial hedonic analysis," *Resource and energy economics*, 2005, *27* (3), 248–269.

Brandt, Adam R, GA Heath, EA Kort, F O'sullivan, G Pétron, SM Jordaan, P Tans, J Wilcox, AM Gopstein, D Arent et al., "Methane leaks from North American natural gas systems," *Science*, 2014, *343* (6172), 733–735.

Carpenter, David H., "How Fannie Mae and Freddie Mac Typically Handle Requests to Create Oil, Gas, or Mineral Leases on Residential Properties," *Congressional Research Service*, 2011.

Caulton, Dana R, Paul B Shepson, Renee L Santoro, Jed P Sparks, Robert W Howarth, Anthony R Ingraffea, Maria OL Cambaliza, Colm Sweeney, Anna Karion, Kenneth J Davis et al., "Toward a better understanding and quantification of methane emissions from shale gas development," *Proceedings of the National Academy of Sciences*, 2014, *111* (17), 6237–6242.

Cesur, Resul, Erdal Tekin, and Aydogan Ulker, "Air pollution and infant mortality: evidence from the expansion of natural gas infrastructure," *The Economic Journal*, 2017, *127* (600), 330–362.

Cheung, Ron, Daniel Wetherell, and Stephan Whitaker, "Earthquakes and House Prices: Evidence from Oklahoma," *FRB of Cleveland Working Paper No. 16-31*, 2016.

Coase, Ronald Harry, "The problem of social cost," *The journal of Law and Economics*, 1960, *56* (4), 837–877.

Colborn, Theo, Carol Kwiatkowski, Kim Schultz, and Mary Bachran, "Natural gas operations from a public health perspective," *Human and ecological risk assessment: An International Journal*, 2011, *17* (5), 1039–1056.

— , **Kim Schultz, Lucille Herrick, and Carol Kwiatkowski**, "An exploratory study of air quality near natural gas operations," *Human and Ecological Risk Assessment: An International Journal*, 2014, *20* (1), 86–105.

Delgado, Michael S, Todd Guilfoos, and Andrew Boslett, "The cost of hydraulic fracturing: A hedonic analysis," *Work. Pap*, 2014.

Fed, "Home Mortgage Disclosure Act Examination Procedures," 2009. Available at https://www.federalreserve.gov/boarddocs/caletters/2009/0910/09-10_attachment.pdf.

Fontenot, Brian E, Laura R Hunt, Zacariah L Hildenbrand, Doug D Carlton Jr, Hyppolite Oka, Jayme L Walton, Dan Hopkins, Alexandra Osorio, Bryan Bjorndal, Qinhong H Hu et al., "An evaluation of water quality in private drinking water wells near natural gas extraction sites in the Barnett Shale Formation," *Environmental Science & Technology*, 2013, *47* (17), 10032–10040.

Frölich, Markus, "Non-parametric regression for binary dependent variables," *The Econometrics Journal*, 2006, *9* (3), 511–540.

Gerardi, Kristopher, Adam Hale Shapiro, and Paul Willen, "Subprime outcomes: Risky mortgages, homeownership experiences, and foreclosures," 2007.

— , — , **and** — , "Decomposing the foreclosure crisis: House price depreciation versus bad underwriting," *FRB Atlanta Working Paper*, 2009.

— **and Paul Willen**, "Subprime mortgages, foreclosures, and urban neighborhoods," *The BE Journal of Economic Analysis & Policy*, 2009, *9* (3).

Gilje, Erik P, "Does Local Access to Finance Matter? Evidence from U.S. Oil and Natural Gas Shale Booms," *USAEE Working Paper Forthcoming*, 2012.

— , **Elena Loutskina, and Philip E Strahan**, "Exporting liquidity: Branch banking and financial integration," *The Journal of Finance*, 2016.

Gilman, Jessica B, BM Lerner, WC Kuster, and JA De Gouw, "Source signature of volatile organic compounds from oil and natural gas operations in northeastern Colorado," *Environmental science & technology*, 2013, *47* (3), 1297–1305.

Gopalakrishnan, Sathya and H Allen Klaiber, "Is the shale energy boom a bust for nearby residents? Evidence from housing values in Pennsylvania," *American Journal of Agricultural Economics*, 2014, *96* (1), 43–66.

Graham, Jove, Jennifer Irving, Xiaoqin Tang, Stephen Sellers, Joshua Crisp, Daniel Horwitz, Lucija Muehlenbachs, Alan Krupnick, and David Carey, "Increased traffic accident rates associated with shale gas drilling in Pennsylvania," *Accident Analysis & Prevention*, 2015, *74"* (Supplement C), 203 – 209.

Hill, Elaine L, "Shale gas development and infant health: evidence from Pennsylvania," *Charles H. Dyson School of Applied Economics and Management, Cornell University, Working Paper. Available at*, 2013.

— **and Lala Ma**, "Shale Gas Development and Drinking Water Quality," *forthcoming in American Economic Review Papers and Proceedings*, 2017.

Immergluck, Dan and Geoff Smith, "Measuring the effect of subprime lending on neighborhood foreclosures evidence from Chicago," *Urban Affairs Review*, 2005, *40* (3), 362–389.

— **and** — , "The impact of single-family mortgage foreclosures on neighborhood crime," *Housing Studies*, 2006, *21* (6), 851–866.

James, Alexander and Jasmine James, "A canary near a gas well: Gas booms and housing market busts in Colorado," *Work. Pap*, 2014.

Johnsen, Reid, Jacob LaRiviere, and Hendrik Wolff, "Fracking, Coal and Air Quality," *Working Paper*, 2017.

Kau, James B, Donald C Keenan, and Taewon Kim, "Default probabilities for mortgages," *Journal of urban Economics*, 1994, *35* (3), 278–296.

King, Alvin Thomas, *Discrimination in mortgage lending: A study of three cities*, New York University, Graduate School of Business Administration, Salomon Brothers Center for the Study of Financial Institutions, 1981.

Koster, Hans RA and Jos van Ommeren, "A shaky business: Natural gas extraction, earthquakes and house prices," *European Economic Review*, 2015, *80*, 120–139.

LaCour-Little, Michael, "Discrimination in mortgage lending: A critical review of the literature," *Journal of Real Estate Literature*, 1999, *7* (1), 15–50.

Ladd, Helen F, "Evidence on discrimination in credit markets," *Journal of Economic Perspectives*, 1998, *1*, 223–234.

Law, Sedgwick, "Mortgage Lenders Are Becoming Increasingly Concerned With Gas and Oil Leases Associated With Hydraulic Fracturing," *Hydraulic Fracturing Digest*, 2011.

Lin, Zhenguo, Eric Rosenblatt, and Vincent W Yao, "Spillover effects of foreclosures on neighborhood property values," *The Journal of Real Estate Finance and Economics*, 2009, *38* (4), 387–407.

Mason, Charles F, Lucija A Muehlenbachs, and Sheila M Olmstead, "The economics of shale gas development," *Annu. Rev. Resour. Econ.*, 2015, *7* (1), 269–289.

May, Greg, "Gas and Oil Leases: Impact on Residential Lending," *Tompkins Trust Company*, 2011.

McCarthy, G, S Van Zandt, and W Rohe, "The Economic Benefits and Costs of Home Ownership: A Critical Assessment of the Research (Arlington, VA, Research Institute for Housing America)," *Working Paper*, 2002.

Muehlenbachs, Lucija and Alan J Krupnick, "Shale gas development linked to traffic accidents in Pennsylvania," *Common Resources*, 2013.

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

Munnell, Alicia H, Geoffrey MB Tootell, Lynn E Browne, and James McEneaney, "Mortgage lending in Boston: Interpreting HMDA data," *The American Economic Review*, 1996, pp. 25–53.

Newman, Kathe and Elvin K Wyly, "Geographies of mortgage market segmentation: The case of Essex County, New Jersey," *Housing Studies*, 2004, *19* (1), 53–83.

NYT, "Rush to Drill for Natural Gas Creates Conflicts With Mortgages," 2011. Available at http://www.nytimes.com/2011/10/20/us/rush-to-drill-for-gas-creates-mortgage-conflicts.html.

Olmstead, Sheila M, Lucija A Muehlenbachs, Jhih-Shyang Shih, Ziyan Chu, and Alan J Krupnick, "Shale gas development impacts on surface water quality in Pennsylvania," *Proceedings of the National Academy of Sciences*, 2013, *110* (13), 4962–4967.

Plosser, Matthew C, "Bank heterogeneity and capital allocation: Evidence from'fracking'shocks," *FRB of New York Staff Report*, 2014, (693).

Quercia, Roberto G and Michael A Stegman, "Residential mortgage default: a review of the literature," *Journal of Housing Research*, 1992, *3* (2), 341.

_ , _ , and Walter R Davis, "The impact of predatory loan terms on subprime foreclosures: The special case of prepayment penalties and balloon payments," *Housing Policy Debate*, 2007, *18* (2), 311–346.

Roy, Anirban A, Peter J Adams, and Allen L Robinson, "Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas," *Journal of the Air & Waste Management Association*, 2014, *64* (1), 19–37.

Schafer, Robert and Helen F Ladd, *Discrimination in mortgage lending*, MIT Press (MA), 1981.

Shen, Yannan Lily, Chris Cunningham, and Kris Gerardi, "Unexpected Wealth Change and Mortgage Default: Evidence from the Shale Gas Boom," Technical Report 2015.

Simons, Robert, Roberto Quercia, and Ivan Levin, "The value impact of new residential construction and neighborhood disinvestment on residential sales price," *Journal of Real Estate Research*, 2009.

Smith, Adam, *The wealth of nations: A translation into modern English*, Vol. 7, Industrial Systems Research, 2015.

Throupe, Ron, Robert Simons, and Xue Mao, "A review of hydro"fracking" and its potential effects on real estate," *Journal of Real Estate Literature*, 2013, *21* (2), 205–232.

Upton, Gregory and Meagan McCollum, "Local Labor Market Shocks and Residential Mortgage Payments: Evidence from Shale Oil and Gas Booms," *USAEE Working Paper Forthcoming*, 2016.

Urbina, Ian, "Rush to drill for natural gas creates conflicts with mortgages," *The New York Times*, 2011, *10*, 19.

Viscusi, W Kip, "The value of risks to life and health," *Journal of economic literature*, 1993, *31* (4), 1912–1946.

Warner, Nathaniel R, Cidney A Christie, Robert B Jackson, and Avner Vengosh, "Impacts of shale gas wastewater disposal on water quality in western Pennsylvania," *Environmental science & technology*, 2013, *47* (20), 11849–11857.

Yinger, John, "Discrimination in mortgage lending: A literature review," *Mortgage lending, racial discrimination, and federal policy*, 1996, pp. 29–74.

Appendices

A Additional Empirical Analyses

A.1 Year-Specific Effects by Securitization

Table 12 summarizes the sale and loan characteristics across loans that are not securitized (in the data) and commercially and government securitized. Loans securitized by a government entity like Freddie or Farmer Mac and Fannie or Ginnie Mae, are associated with lower income buyers, value homes, and loan-to-income ratios.

Table 12: Summary Statistics by Securitization

	Not Secured		Gov't Secured		Commercially Secured	
	Mean	(Std. Dev.)	Mean	(Std. Dev.)	Mean	(Std. Dev.)
High Interest Loan	0.155	(0.362)	0.043	(0.202)	0.142	(0.349)
Income	101354.7	(92370.11)	84116	(55294.88)	85715.52	(62805.21)
Loan-to-Income	2.115	(0.838)	2.101	(0.787)	2.206	(0.799)
Sale Amount	165756.4	(123645.1)	151275.3	(80178.49)	155358	(93593.59)
Loan Value	181121.9	(164265.4)	154592.1	(83170.38)	166751.1	(108633.4)
Annualized L2I	0.173	(0.069)	0.167	(0.057)	0.175	(0.061)
Foreclosure	0.089	(0.285)	0.074	(0.262)	0.077	(0.267)
Fixed Loan	0.785	(0.411)	0.951	(0.215)	0.855	(0.352)
FHA Loan	0.101	(0.301)	0.281	(0.45)	0.314	(0.464)
Owner Occ.	0.875	(0.331)	0.931	(0.253)	0.934	(0.249)
Obs.	91,311		188,701		300,550	

[a] Summarizes the primary variables used to describe the loan characteristics across loans that are and are not securitized.

A.2 Tarrant County Logit Models - Simple Exposure

In Tarrant county, the relationship between loan-to-income is positive when estimating the logit using all years of data as demonstrated in the first column of Table 13. Breaking the data in periods before 2008 and after 2009, the relationship between loan-to-income is positive in early periods and negative in later periods. After 2009, FHA loans are more likely to have greater interest rates (2.092***), whereas before 2008, the relationship is negative (-3.26***). A similar shift occurs for loans originated for owner-occupied houses. Conversely, a loan that is commercially securitized is less likely to have a high interest rate after 2009 (compared to before 2008), and government securitized are even less likely to be originated with high interest rates post 2009. Figure 18 plot the year-estimates for the relationship between loan-to-income and the interactions with whether the loan is an FHA or Securitizes, and Figure 19 adds the effect of shale risk. Before 2009, the high loan-to-income values among securitized loans were more likely to have high interest rates, and after 2009, the high loan-to-income values among FHA loans were more likely originated with high interest rates.

Stratifying the data by water access, the observed patterns for loan types across periods persist for both district and ground water regions, though the magnitudes are a little different. The primary different is that the relationship between shale risk and a high interest rate is a little significant among groundwater houses, though the magnitude is less after 2009. Figure 19 suggests that shale risk has a positive and significant relationship with high interest rates in 2010 and 2013.

Table 14 captures how the model estimates change with the addition of racial and ethnicity household characteristics and loan characteristics. Moving from column (1) to (2) reveals the importance of controlling for the racial and ethnic composition of borrowers. In Column (3) we control for whether the loan is insured by the Federal Housing Administration or the Veterans Administration. Column (4) controls for securitization, and in particular, we control for whether the loan is securitized by a government or commercial entity. Column (5) controls for whether the loan is originated for an owner-occupied house (in lieu of an investor or secondary home buyer). The last two columns include interactions between loan-to-income values and whether the loan is securitized by a government entity or commercial entity, the house is owner occupied, and finally, whether the house is exposed to shale. Including interactions soaks up the observed effect of loan-to-income independently whereby both forms of securitization are likely to have high interest rates associated with high loan-to-income ratios. Interactions with shale soak up the effect of shale exposure independently, as well.

Table 13: Subprime Logit - Tarrant County

Dependent Variable: Subprime Loan				District Water		Groundwater	
	Full	Before 2009	After 2009	Before 2009	After 2009	Before 2009	After 2009
Shale Risk	0.00890	0.00797	0.01035	-0.02029	0.01446	0.02613*	0.01532*
	(0.00547)	(0.01161)	(0.00728)	(0.02228)	(0.01409)	(0.01495)	(0.00896)
V2I	0.03516***	0.07248***	-0.08415***	-0.03465	-0.19105***	0.14338***	0.01075
	(0.01262)	(0.01679)	(0.02579)	(0.02685)	(0.04058)	(0.02200)	(0.03410)
FHA Loan	-0.55000***	-3.26782***	2.09160***	-3.41661***	2.02102***	-3.16380***	2.09284***
	(0.02406)	(0.07391)	(0.05357)	(0.11581)	(0.08449)	(0.09683)	(0.07066)
VA Loan	-3.92082***	-5.17340***	-2.76600***	-4.51112***	-2.91574***	-6.23394***	-2.70941***
	(0.19058)	(0.44887)	(0.35848)	(0.50372)	(0.58809)	(1.00125)	(0.45309)
Gov't Sec.	-1.04558***	-1.29909***	-2.01126***	-1.12973***	-2.17845***	-1.44928***	-1.89200***
	(0.03097)	(0.04541)	(0.06409)	(0.06945)	(0.09884)	(0.06126)	(0.08644)
Comm. Sec.	0.10268***	0.55416***	-1.74178***	0.64951***	-1.87563***	0.46236***	-1.63638***
	(0.02484)	(0.03094)	(0.06122)	(0.04819)	(0.09332)	(0.04116)	(0.08320)
Owner Occ.	-0.24551***	-0.29383***	1.14094***	-0.17505***	1.38999***	-0.36508***	0.98351***
	(0.03351)	(0.03985)	(0.14986)	(0.06004)	(0.23868)	(0.05460)	(0.19943)
Hispanic	0.84592***	0.79341***	0.80235***	0.76699***	0.98185***	0.81981***	0.61738***
	(0.02321)	(0.03172)	(0.04591)	(0.04864)	(0.07007)	(0.04314)	(0.06272)
Black	1.30950***	1.58969***	0.66769***	1.66176***	0.82186***	1.54720***	0.54313***
	(0.02739)	(0.03636)	(0.06528)	(0.05739)	(0.10288)	(0.04842)	(0.08692)
Asian	-0.28462***	-0.28114***	-0.22017*	-0.17814*	-0.19348	-0.31539***	-0.21228
	(0.04949)	(0.06182)	(0.11615)	(0.10053)	(0.20123)	(0.07924)	(0.14445)
Yield Curve	-0.28462***	-0.21739***	0.21941***	-0.26760***	0.12054	-0.16283**	0.29365***
	(0.03362)	(0.05717)	(0.06270)	(0.09051)	(0.09645)	(0.07522)	(0.08392)
Beds	0.25665***	0.27333***	0.17159***	0.27001***	0.18358***	0.25612***	0.13478**
	(0.02000)	(0.02593)	(0.04435)	(0.04160)	(0.06986)	(0.03399)	(0.05897)
Baths	-0.17966***	-0.22636***	-0.02763	-0.29738***	0.00916	-0.19777***	-0.04606
	(0.02550)	(0.03301)	(0.05629)	(0.05327)	(0.08731)	(0.04314)	(0.07551)
Living Size	-0.00041***	-0.00037***	-0.00057***	-0.00038***	-0.00067***	-0.00029***	-0.00042***
	(0.00002)	(0.00003)	(0.00005)	(0.00004)	(0.00007)	(0.00003)	(0.00006)
Land Size	-0.00429***	-0.00593***	-0.00303	-0.00373	0.00795	-0.00917***	-0.01125**
	(0.00162)	(0.00206)	(0.00371)	(0.00433)	(0.00685)	(0.00254)	(0.00475)
Age	0.00671***	0.00844***	0.00522***	0.00399***	0.00001	0.01752***	0.01630***
	(0.00051)	(0.00069)	(0.00110)	(0.00092)	(0.00149)	(0.00128)	(0.00209)
Constant	0.19631	-0.30055	-2.28994***	1.18883	-13.24518	-0.23164	-13.39840
	(0.22847)	(0.33985)	(0.77425)	(0.86276)	(674.47492)	(0.82153)	(487.88197)
Observations	121,476	56,750	44,386	21,148	16,282	34,577	27,385
Year FE	x	x	x	x	x	x	x
City FE	x	x	x	x	x	x	x

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[a] Each column reports estimates from a logit model of subprime loans. Each specification controls for year and city fixed effect along with house, household, and loan characteristics. The first column includes all transactions occurring between 2003 and 2016 in Tarrant County Texas, and the next two columns estimate the models separately for sales occurring before and after 2009, the rough timing of the financial crisis. The last four columns estimate the models separately for houses that are likely and not likely accessing groundwater. The shale risk increases the likelihood the loan is subprime among groundwater houses.

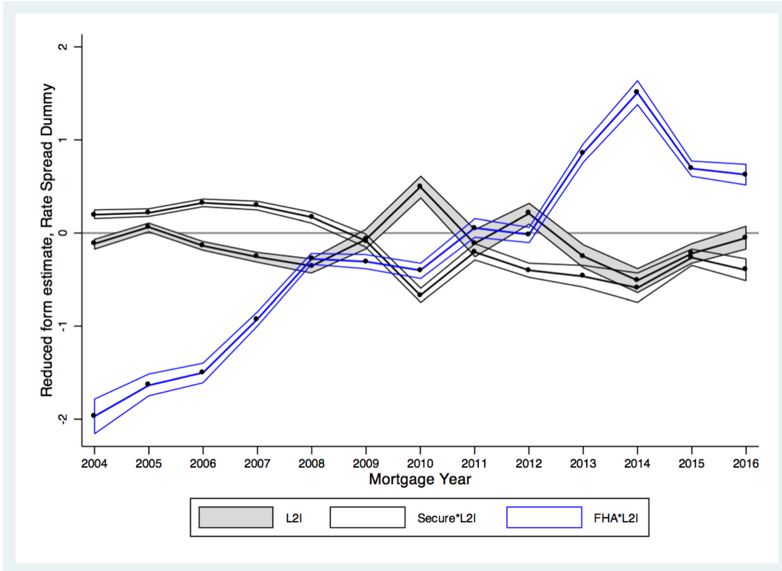


Figure 18: County year-specific effects

[a] Each point is a coefficient estimating the year specific relationship between loan-to-income, L2I*Securitized Loan, and L2I*FHA Loan and the likelihood the loan is subprime. The model is estimated using a logit specification comprised of data from Tarrant, Dallas, and Denton counties, and the model controls for year and city fixed effects along with house, household, and loan characteristics.

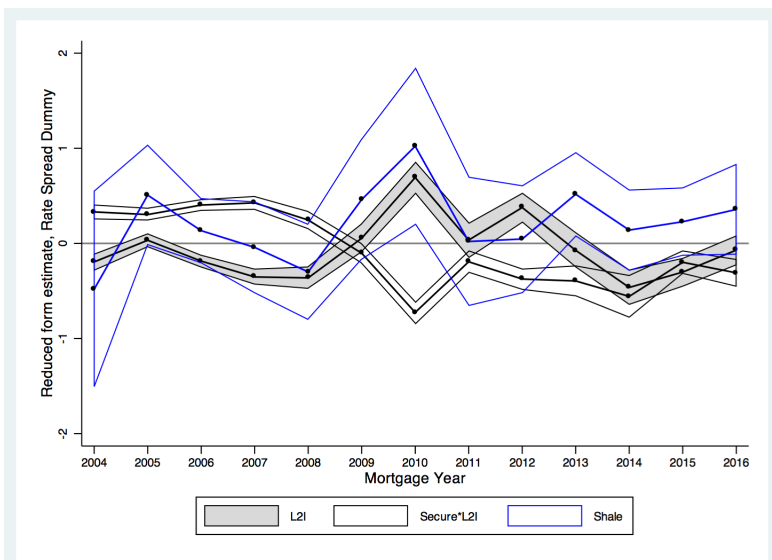


Figure 19: Tarrant county year-specific effects

[a] Each point is a coefficient estimating the year specific relationship between loan-to-income, L2I*Securitized Loan, and shale exposure and the likelihood the loan is subprime. The model is estimated using a logit specification that controls for year and city fixed effects along with house, household, and loan characteristics. The year-specific effects of L2I*FHA Loan are estimated, as well, but not plotted.

Table 14: Subprime Logit Variable Selection - Tarrant County

Dependent Variable: Subprime Loan Indicator							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Shale Risk	0.00868*	0.00981**	0.01541***	0.01181**	0.01081**	0.01107**	-0.00430
	(0.00463)	(0.00498)	(0.00501)	(0.00509)	(0.00510)	(0.00510)	(0.00663)
L2I	0.05583***	-0.04985***	0.01456	-0.01175	0.03252***	-0.05764	-0.06509
	(0.00997)	(0.01110)	(0.01119)	(0.01138)	(0.01224)	(0.04565)	(0.04570)
FHA Loan			-0.54120***	-0.57576***	-0.54553***	-0.55454***	-0.55889***
			(0.02284)	(0.02365)	(0.02392)	(0.02397)	(0.02400)
VA Loan			-3.87966***	-3.91442***	-3.88795***	-3.89959***	-3.90457***
			(0.19030)	(0.19044)	(0.19046)	(0.19047)	(0.19048)
Gov't Sec.				-1.06335***	-1.06335***	-1.73737***	-1.72708***
				(0.03044)	(0.03047)	(0.08062)	(0.08067)
Comm. Sec.				0.10050***	0.09915***	-0.33682***	-0.33179***
				(0.02383)	(0.02385)	(0.06183)	(0.06186)
Owner Occ.					-0.31956***	-0.18222***	-0.17808***
					(0.03249)	(0.06256)	(0.06258)
L2I*Gov't						0.32409***	0.31909***
						(0.03574)	(0.03577)
L2I*Comm.						0.21008***	0.20756***
						(0.02766)	(0.02767)
L2I*Occ.						-0.10399**	-0.10531**
						(0.04276)	(0.04277)
L2I*Shale							0.05005***
							(0.01327)
Yield Curve	-0.30845***	-0.30733***	-0.32267***	-0.29357***	-0.28916***	-0.28746***	-0.28551***
	(0.02971)	(0.03206)	(0.03247)	(0.03274)	(0.03279)	(0.03281)	(0.03282)
Hispanic		0.85134***	0.85624***	0.85769***	0.86334***	0.87042***	0.86961***
		(0.02160)	(0.02187)	(0.02250)	(0.02251)	(0.02254)	(0.02254)
Black		1.25195***	1.36792***	1.34377***	1.33797***	1.34023***	1.33904***
		(0.02504)	(0.02584)	(0.02632)	(0.02636)	(0.02637)	(0.02637)
Asian		-0.17815***	-0.30929***	-0.22970***	-0.26570***	-0.26346***	-0.26193***
		(0.04695)	(0.04705)	(0.04762)	(0.04786)	(0.04789)	(0.04789)
Beds	0.31120***	0.23141***	0.24977***	0.26334***	0.26038***	0.26045***	0.26066***
	(0.01731)	(0.01893)	(0.01911)	(0.01937)	(0.01938)	(0.01940)	(0.01940)
Baths	-0.25433***	-0.16931***	-0.20209***	-0.20318***	-0.20280***	-0.20024***	-0.19665***
	(0.02209)	(0.02387)	(0.02404)	(0.02440)	(0.02441)	(0.02442)	(0.02444)
Living Size	-0.00048***	-0.00040***	-0.00045***	-0.00045***	-0.00043***	-0.00044***	-0.00044***
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Land Size	-0.00722***	-0.00236*	-0.00414***	-0.00316**	-0.00272*	-0.00284**	-0.00288**
	(0.00129)	(0.00137)	(0.00138)	(0.00139)	(0.00139)	(0.00139)	(0.00139)
Age	0.00750***	0.00754***	0.00587***	0.00594***	0.00602***	0.00599***	0.00624***
	(0.00043)	(0.00047)	(0.00048)	(0.00048)	(0.00048)	(0.00049)	(0.00049)
Constant	-0.04147	-0.19294	0.02789	0.03843	0.18609	0.44844**	0.43819**
	(0.16136)	(0.17496)	(0.17705)	(0.17875)	(0.17939)	(0.18897)	(0.18899)
Obs.	141,451	123,507	123,507	123,507	123,392	123,392	123,392
Year FE	x	x	x	x	x	x	x

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[a] Each column reports estimates from a logit model of subprime loans. Each specification controls for year and city fixed effect, and each column adds additional control variables. The second column adds controls for the race and ethnicity of the primary loan signatory and the coefficient for debt-to-income becomes negative, and the coefficient flips sign with the addition of a control for Black households. Controlling for loan type and lien status flips the sign of the estimate again.

A.3 Dallas County Logit Models - Simple Exposure

Table 15 reports estimates for the subprime logit models using data from Dallas county where there is no drilling activity. The relationships between loan-to-income and high interest rates are counter-intuitively negative, and including interactions with loan types only matters before 2009. Figure **??** plots year-specific effects, and the only positive coefficient for loan-to-income appears in 2010. Similar to Tarrant county, the effects of loan-to-income interacted with FHA are positive post-subprime crisis while the effects of loan-to-income alone and loan-to-income interacted with securitization are largely negative.

Table **??** estimates models that interact loan-to-income with whether the sale value is at the bottom or top of the price distribution to see whether houses at different transaction levels receive different interest rates. In Tarrant county, the price point does not appear to matter before 2009; however, after 2009, lower sale value houses are more likely to have higher interest rates when borrowers have higher loan-to-income ratios. Conversely, expensive homes are less likely to have high interest rates post-2009. In Dallas, the same pattern persists regardless of sale timing, though the magnitudes differ across pre- and post-2009.

Table 15: Subprime Logit - Dallas County

Dependent Variable:	Subprime Loan Indicator					
	Full	Before 2009	After 2009	Full	Before 2009	After 2009
Loan-to-Income	-0.02556**	0.00429	-0.12614***	-0.05722	-0.02731	-0.14463
	(0.01135)	(0.01442)	(0.02692)	(0.04463)	(0.05401)	(0.17423)
FHA Loan	-0.96574***	-2.91282***	1.88704***	-0.97385***	-2.92643***	1.88684***
	(0.02410)	(0.05091)	(0.06004)	(0.02418)	(0.05098)	(0.06028)
Loan VA	-3.78806***	-5.62431***	-2.28080***	-3.79365***	-5.63156***	-2.28001***
	(0.19925)	(0.58002)	(0.41438)	(0.19926)	(0.58004)	(0.41439)
Owner Occ.	-0.38281***	-0.34331***	-0.03500	-0.30852***	-0.17595**	-0.00991
	(0.03211)	(0.03844)	(0.12368)	(0.06114)	(0.07484)	(0.23020)
Gov't Sec.	-1.03366***	-1.29715***	-1.82906***	-1.32020***	-2.09264***	-1.71799***
	(0.02919)	(0.03958)	(0.06946)	(0.07868)	(0.10959)	(0.18221)
Comm. Sec.	0.16820***	0.38738***	-1.37379***	-0.06629	-0.01066	-1.50290***
	(0.02201)	(0.02638)	(0.06322)	(0.05968)	(0.07219)	(0.16465)
L2I*Occ.				-0.05733	-0.13296**	0.00092
				(0.04287)	(0.05192)	(0.17495)
L2I*Gov't				0.12792***	0.35468***	-0.04646
				(0.03286)	(0.04523)	(0.07261)
L2I*Comm.				0.10378***	0.17551***	0.05191
				(0.02466)	(0.02998)	(0.06395)
Hispanic	1.41659***	1.49141***	1.07708***	1.41979***	1.49887***	1.07630***
	(0.02164)	(0.02715)	(0.05135)	(0.02166)	(0.02720)	(0.05136)
Black	1.75553***	1.99280***	0.83774***	1.75618***	1.99570***	0.83733***
	(0.02391)	(0.02936)	(0.06651)	(0.02392)	(0.02940)	(0.06651)
Asian	-0.24810***	-0.25599***	-0.27417**	-0.25001***	-0.25935***	-0.27608**
	(0.04631)	(0.05755)	(0.11936)	(0.04633)	(0.05761)	(0.11938)
Yield Curve	-0.37817***	-0.27580***	0.15878**	-0.37667***	-0.27303***	0.15807**
	(0.03116)	(0.05016)	(0.06486)	(0.03117)	(0.05020)	(0.06487)
Beds	0.23941***	0.27054***	0.12035***	0.24017***	0.27235***	0.12075***
	(0.01799)	(0.02258)	(0.04518)	(0.01800)	(0.02260)	(0.04519)
Baths	-0.09646***	-0.12032***	-0.04370	-0.09627***	-0.11994***	-0.04406
	(0.01892)	(0.02352)	(0.04848)	(0.01892)	(0.02355)	(0.04849)
Living Size	-0.00059***	-0.00061***	-0.00054***	-0.00060***	-0.00061***	-0.00054***
	(0.00002)	(0.00003)	(0.00006)	(0.00002)	(0.00003)	(0.00006)
Land Size	0.01327***	0.01210***	0.01337***	0.01319***	0.01209***	0.01340***
	(0.00188)	(0.00232)	(0.00488)	(0.00188)	(0.00233)	(0.00488)
Age	-0.00111**	-0.00008	-0.00319***	-0.00118**	-0.00022	-0.00319***
	(0.00047)	(0.00059)	(0.00118)	(0.00047)	(0.00059)	(0.00118)
Constant	0.84746***	0.15080	-0.95984***	0.97472***	0.35392	-0.94135***
	(0.17032)	(0.26502)	(0.27034)	(0.17985)	(0.27471)	(0.32901)
Observations	113,908	59,936	34,998	113,908	59,936	34,998
Year FE	x	x	x	x	x	x
City FE	x	x	x	x	x	x

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[a] Each column reports estimates from a logit model of subprime loans. Each specification controls for year and city fixed effect along with house, household, and loan characteristics. The first column includes all transactions occurring between 2003 and 2016 in Dallas County Texas, and the last two columns estimate the models separately for houses sold before and after 2009, roughly the time of the financial crisis.

A.4 Texas-wide HMDA

The final analysis uses the full set of HMDA loans in Texas to estimate a relationship between the loan-to-incomes and the probability the borrower is given a high interest rate, mimicking the logit models using the house transaction data in Dallas, Denton, and Tarrant. For all scenarios, the estimated relationships are negative, and they estimates increase in magnitude when estimating the model for all data after 2009. However, interacting loan-to-income with loan types, we see that after 2009, both FHA and securitized loans are more likely to have higher interest rates as the loan-to-income values increase rendering the combined estimate a little positive (-0.417 + 0.273 + 0.191).

Shale exposure is measured as the cumulative count of wells in a particular county in a particular year since we only know what year the loan was originated in the HMDA data. After 2009, the relationship between additional drilling and subprime interest rates is more positive, which is consistent with the larger story.

Table 16: Texas-wide Lending

Dependent Variable:	Subprime Interest Rate Indicator				
	Full	Before 2009	After 2009	Before 2009	After 2009
Loan-to-Income	-0.14074***	-0.07967***	-0.20221***	-0.09512***	-0.41774***
	(0.00126)	(0.00197)	(0.00201)	(0.00278)	(0.00302)
Drilling	0.01467***	0.00393**	0.03202***	0.00399**	0.03155***
	(0.00083)	(0.00155)	(0.00253)	(0.00155)	(0.00254)
FHA Loan	0.17207***	-1.28805***	1.07695***	-0.80766***	0.47613***
	(0.00241)	(0.00659)	(0.00385)	(0.02087)	(0.01012)
VA Loan	-1.66194***	-2.26956***	-1.30414***	-2.27979***	-1.23998***
	(0.01172)	(0.02599)	(0.01814)	(0.02599)	(0.01821)
Gov't Secured	-0.80109***	-0.79976***	-1.18910***	-0.79780***	-1.13538***
	(0.00279)	(0.00457)	(0.00448)	(0.00458)	(0.00453)
Comm. Secured	-0.29059***	0.16134***	-1.02929***	0.08027***	-1.41374***
	(0.00241)	(0.00351)	(0.00430)	(0.00723)	(0.00988)
Owner Occupied	0.13514***	0.02419***	0.47208***	0.01638***	0.65138***
	(0.00328)	(0.00452)	(0.00654)	(0.00453)	(0.00687)
L2I*FHA				-0.22860***	0.27338***
				(0.00956)	(0.00428)
L2I*Secured				0.04505***	0.19130***
				(0.00355)	(0.00409)
Hispanic	0.50070***	0.59024***	0.41186***	0.59149***	0.39830***
	(0.00236)	(0.00358)	(0.00384)	(0.00359)	(0.00386)
Black	0.66204***	0.86604***	0.37880***	0.86893***	0.35894***
	(0.00329)	(0.00478)	(0.00585)	(0.00479)	(0.00589)
Asian	-0.26815***	-0.26015***	-0.28585***	-0.26101***	-0.28063***
	(0.00493)	(0.00725)	(0.00869)	(0.00725)	(0.00876)
Constant	-0.38082***	-0.70890***	-0.93008***	-0.68690***	-0.74402***
	(0.02346)	(0.04020)	(0.03532)	(0.04025)	(0.03574)
Observations	3,221,677	1,196,852	1,608,589	1,196,852	1,608,589
Year FE	x	x	x	x	x
County FE	x	x	x	x	x

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$