Resource and Energy Economics 53 (2018) 162–197

Contents lists available at ScienceDirect

Resource and Energy Economics

journal homepage: www.elsevier.com/locate/ree

Local labor market shocks and residential mortgage payments: Evidence from shale oil and gas booms

Meagan McCollum[a], Gregory B. Upton Jr.[b,*]

[a] Department of Finance, University of Tulsa, United States
[b] Center for Energy Studies, Louisiana State University, United States

ARTICLE INFO

Article history:
Received 12 July 2017
Received in revised form 30 April 2018
Accepted 1 May 2018
Available online 16 May 2018

JEL classification:
R23
D14
Q33
Q35

ABSTRACT

Frequently, housing is the largest item on a household's balance sheet, and therefore making monthly mortgage payments is often both the largest regular expenditure as well as a primary savings vehicle for households. However, changes to economic conditions impact household spending and savings decisions. To investigate the dynamics of this relationship, we examine mortgage payment choices of homeowners who purchased property in areas that later experienced a positive shock to local economic conditions via the shale oil and gas boom. We find that borrowers with properties located in areas with shale oil and gas booms experienced a 6% reduction in the probability of missing a mortgage payment over the period 2007–2014. Indexing these results to the size of the boom, we find that one hundred additional rigs (billion dollars of oil and gas produced) are associated with a 3.2% (1.6%) decrease in default. Additionally, we find differential effects on housing markets across geography, time, loan leverage, and credit risk categories.

1. Introduction

After many years of declining crude oil production in the United States, recent technological developments have made the extraction of previously inaccessible energy resources feasible. Specifically, the advent of horizontal drilling and hydraulic fracturing techniques have enabled the exploration and production of oil and gas from shale geological formations, and lead to significant new drilling activity over the past decade. Contemporaneously, widespread declines in residential housing values and sharp increases in mortgage default rates in 2007–2009 were a central component of the Great Recession. Notably, in the midst of the Great Recession, the technological innovations that enabled shale oil and gas extraction provided a catalyst for an economic "boom" to clearly specified local areas where these previously inaccessible resources could now be profitably extracted. This research focuses on how this natural resource boom impacted local residents of areas where these shale oil and gas resources were extracted. Specifically, we examine the impact of shale oil and gas discovery on the mortgage payment decisions of long-term residents of six geographic areas that have the geological formations that allow for shale oil and/or gas extraction, namely: *Bakken*, *Eagle Ford*, *Haynesville*, *Marcellus*, *Niobrara*, and *Utica*.[1]

We estimate the impact of the shale boom on mortgage payment activity of individuals who purchased property in one of these areas prior to the natural resource discovery. Specifically, we examine the impact of the shale boom on the probability of mortgage default during a time period where aggregate default rates nationwide were sharply increasing. For the average

* Corresponding author.
[1] These areas are definitions are based on EIA (2017). The *Permian* basin located in western Texas was not included in this analysis because sufficient data on mortgages were not available.

https://doi.org/10.1016/j.reseneeco.2018.05.001
0928-7655/

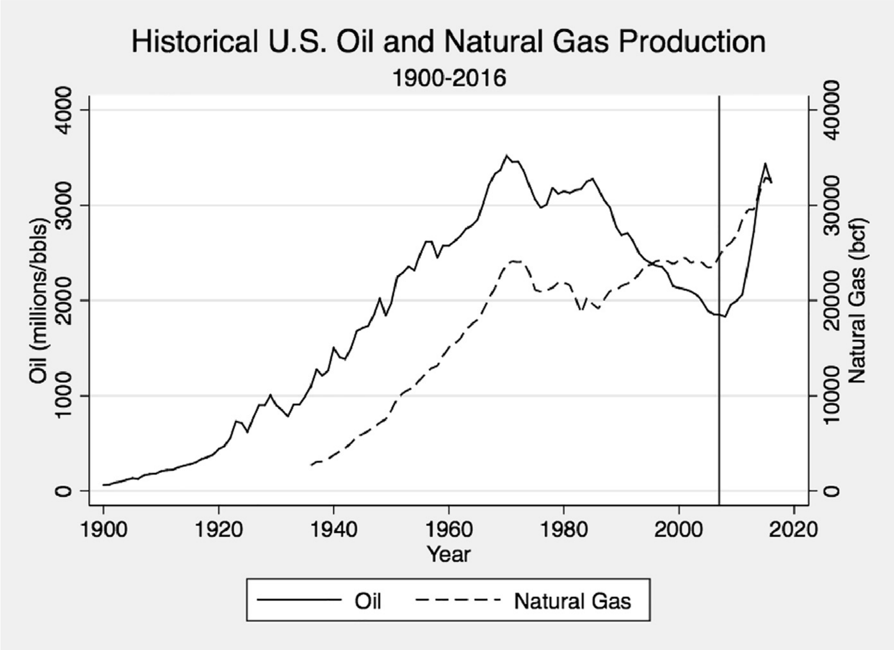


Fig. 1. Historical U.S. crude oil and natural gas production.

homeowner, their house is typically the largest asset on their household balance sheet, typically making up over two-thirds of a household's wealth (Iacoviello, 2011). Additionally, for homeowners with outstanding mortgages, this loan is typically their largest financial obligation.

Using a difference-in-differences framework, we find that borrowers with properties in counties with shale oil and gas resources experience, on average, a 6% reduction in the probability of mortgage default as compared to similar mortgages in non-shale areas after the boom began. This reduction in the probability of default reaches a maximum of approximately 7–9% in 2009, during the peak of the shale boom, and attenuates to approximately a 1–2% difference in default probabilities by the end of 2014. These results are robust to choice of control group, risk categories, alternate definitions of default, inclusion of prepayment as a loan outcome, local population levels, and placebo tests.

1.1. Economic impact of oil and gas booms

Modern crude oil production began in 1859 with Drake Well, five miles south of Titusville, Pennsylvania and began a period of rapid growth and expansion in the oil industry (Yergin, 1999). As people from all income ranges around the country began "pushing back the night" for the first time with inexpensive fuel that could be used for lighting homes, oil became an almost instant necessity. So began the age of oil that quickly spread throughout the world.

For almost a century the U.S. experienced consistent increases in oil production. But in 1970, this age of increasing domestic production reached its end and for the first time in U.S. history production began a period of decline that continued for the next four decades. However, over the last ten years, the oil landscape has changed both suddenly and dramatically as illustrated in Fig. 1. By 2007, after a long period of declining production in the U.S., a technological breakthrough allowed shale oil and gas extraction to become economically viable for the first time in history; the shale boom was underway.[2] Through a combination of horizontal drilling and hydraulic fracturing (informally referred to "fracking") the trend in oil production reversed itself and the U.S. has since experienced increases in production. By the end of 2014, the U.S. was observing crude production similar to the historic levels achieved during "peak oil" of the 1970s (EIA, 2014).[3] There is a growing literature on the economic impact of fossil fuel based shocks to economic activity; and this literature has seen resurgence due to the recent shale boom. Black et al. (2005) examines the impact of the coal boom and subsequent bust in the 1970s and 1980s on local labor markets and finds that in addition to increases in employment in the coal sector, employment increased in non-coal sectors as well. More recently, Allcott and Keniston (2018) utilize historical oil and gas production data in the U.S. since the 1960s and find that booms increase both employment and wages of local workers, and these increases are not just

[2] For the main empirical specifications in this research, the shale boom will begin in 2007 consistent with the time that EIA began tracking shale production (EIA, 2017). We will consider the specific timing of the treatment in an alternative specification.

[3] By about mid-2015, the shale boom was slowing substantially due to the sharp drop in the oil price seen worldwide. This research will consider data until the end of 2014.

restricted to the oil and gas sectors. Feyrer et al. (2017) find that the shale boom specifically induced significant economic shocks to local labor markets; every million dollars of oil and gas extracted is estimated to generate $243,000 in wages, $117,000 in royalty payments, and 2.49 jobs within a 100 mile radius. In total, the authors estimate that the shale boom was associated with 725,000 jobs in aggregate and a 0.5% decrease in the unemployment rate during the Great Recession. Agerton et al. (2017) find that one addition rig results in the creation of 31 jobs immediately and 315 jobs in the long-run. Other studies corroborate the positive impact of the shale boom on local labor markets (Weber, 2012; Marchand and Weber, 2017; Komarek, 2016; Bartik et al., 2017; Upton and Yu, 2018; Decker et al., 2018).

Thus, shocks to both the coal and oil and gas sectors have been found to impact the local economy, not only by creating jobs within the respective fossil fuel sector but also spilling over into other sectors within the economy. This is important, because if these shocks only provide increases in employment and wages in a specific sector, and that sector makes up a relatively small share of the local economy, then it is questionable whether or not the shock will be of interest to economists who want to use these shocks to generally understand how plausibly exogenous overall shocks to employment and wages impact outcomes of interest.

Because these shocks have been shown to have economy wide labor market implications, some studies have used natural resource shocks as an instrument for local labor market conditions. Black et al. (2002) exploit the coal boom and bust in the 1970s–1980s as a shock to the value of labor market participation to test the impact of earnings on Social Security Disability Insurance (DI) payments and participation, finding a negative relationship between earnings growth and growth in the DI program. Acemoglu et al. (2013) take advantage of variation in oil prices interacted with local oil reserves to estimate the impact of rising income on health expenditures.

Additionally, there has been interest in how shale booms impact local financial conditions. Gilje (2012) treats the shale boom as a catalyst to an exogenous increase in local bank deposits and local credit supply and finds that counties with large shale booms also experience a large increase in new business establishments that are reliant on external bank financing. This effect is particularly strong in counties that are dominated by small local banks. Similarly, Gilje et al. (2016) exploit the shale boom to show that bank branch networks continue to play an important role in financial integration by demonstrating that banks with branch exposure to shale booms also increase mortgage lending in non-boom counties.

There are a number of studies that have specifically focused on the impact of oil and gas development on local real estate markets, but this literature has been largely focused on the negative impact of drilling wells in close proximity to residential properties. Boxall et al. (2005) examine the impact of oil and gas facilities on rural residential property values in Alberta, Canada and find that property values are negatively correlated with the number of sour gas wells and flaring oil batteries. Using hedonic pricing models, Muehlenbachs et al. (2015) find that groundwater-dependent homes near oil and gas wells in the Marcellus shale (located in Pennsylvania) experience decreases in housing values, while similar homes that receive water from pipes experience increases in housing values likely associated with lease payments. Gopalakrishnan and Klaiber (2014) estimate a hedonic pricing model for homes near Pittsburgh, PA (also in the Marcellus shale region) and find that homes in close proximity to wells experienced decreases in housing values. While not a study on shale specifically, Boxall et al. (2005) finds that houses located near wells emitting hydrogen sulfide (that smells like rotten eggs) had a negative impact on property values in Alberta, Canada. All of these studies that examine local real estate markets have focused primarily on the negative impact of drilling on homes in close proximity to drilling activity, the “not in my back yard” (NIMBY) mantra that is commonly used in this context. To date there is limited work on the impact of drilling activity on housing markets beyond properties directly impacted by natural resource extraction. One notable exception is work by Shen et al. (2015) in examining the impact of the shale oil and gas boom on mortgage markets in Pennsylvania. The authors examine default probabilities specifically for loans within the state of Pennsylvania and find no evidence that nearby “fracking” triggers mortgage default, but do find evidence that the economic activity associated with the boom can decrease the probability of mortgage default for new mortgages that originate after the boom begins.

There are two plausible channels through which an economic boom may decrease mortgage delinquency rates. The first channel is through increased earnings and employment. A number of studies have used natural resource booms as an instrument for local earnings to test the impact of earnings on a number of outcomes such as disability program participation (Black et al., 2002), health spending (Acemoglu et al., 2013) and labor migration (Vachon, 2015). The second channel is through a plausible housing price increase. Housing prices are known to be pro-cyclical (Leamer, 2008; Davis and Heathcote, 2005). Thus, if housing values in an area are increasing, and a household finds themselves in a situation where they are unable to pay their mortgage and expect this inability to pay to extend beyond the short-term, the household with positive equity will rationally choose to sell the property instead of defaulting on the loan. We extend the literature on the impact of natural resource booms on housing and more broadly, household financial decisions in several ways. First, instead of focusing on changes in real estate values with a hedonic pricing model like much of the previous literature (Boxall et al., 2005; Muehlenbachs et al., 2015; Gopalakrishnan and Klaiber, 2014), we examine individual households’ mortgage payment decisions, specifically focusing on the probability of mortgage default. We hypothesize that probability of mortgage default will decline for households in these areas compared to similar households in other parts of the country with no shale boom.

Second, instead of focusing on a single area that experienced a boom (Shen et al., 2015), we take advantage of the fact that the shale boom is unique in that it impacted multiple areas across the country that happened to be located on specific geological formations. Therefore we are able to analyze the effect of the boom on six clearly defined geographic areas. These shale plays include *Bakken*, *Eagle Ford*, *Haynesville*, *Marcellus*, *Niobrara* and *Utica*. We identify individual mortgages that originated in counties with shale oil and/or gas activity prior to the technology shock that enabled the profitable extraction of these resources and track their payment activity until the end of 2014, or until the mortgage is terminated. Thus, this is the most comprehensive study on the impact of shale on housing markets across different shale plays.

Finally, we employ a large detailed data set with national coverage, BlackBox Logic (BBx), which provides information on over 90% of the privately securitized mortgages in the U.S. which includes origination information on over 20 million unique single-family residential mortgages.[4] By using a nationwide sample of loans we can select a control group of loans that are not in geographic proximity to shale oil and gas extraction to mitigate concerns about spillover effects. We observe information on individual loans at origination as well as detailed monthly payment histories of each mortgage. Therefore, we know if and when a household has missed a mortgage payment as well as the current outstanding balance on the loan. We match individual loan payments to counties that experienced increases in oil and gas activity from 2007 to 2014 compared to similar loans in similar counties across the U.S. that did not experience a shale boom.[5]

2. Data, variables, and summary statistics

2.1. Data

We use loan-level information for properties located in the shale (treated) areas compared to properties in non-shale (control) areas; we observe characteristics about the loan and the borrower at origination as well as the time series of monthly payments until the loan is terminated. We begin to observe monthly payments after the loan is placed into a mortgage backed security.[6] A loan can leave our sample for three reasons: the loan is terminated by the lender for lack of payment (severe default, foreclosure, and/or bankruptcy), the borrower prepays the remaining balance of the mortgage (e.g. lump sum prepayment, refinancing the property, or selling the property before the mortgage is repaid in full), or the original contract period of the loan ends. To avoid the potentially confounding effects of changes to the local real estate market after the beginning of the shale boom, we only include observations from loans that were originated prior to the discovery of shale oil and gas in a given area. Therefore, we are considering changes in mortgage payment activity on loans that were originated before the shale boom began.

Loan-level data used in our analysis is from BlackBox Logic, LLC (BBx).[7] BBx contains information on over 20 million loans and includes over 900 million monthly remittance records as of December 2014. This proprietary mortgage database covers over 90% of non-agency residential securitized loans including prime, alt-a, and subprime loans. This dataset provides both characteristics of each loan at origination as well as monthly payment records.

For this study, only first-lien,[8] single family, owner occupied properties[9] that have fully-amortizing loan terms are included in our sample. Additionally, we restrict our sample to new purchase loans; that is we exclude loans with the stated purpose of refinancing. Furthermore, we include only loans that originated between 2000 and 2006, before the shale boom began.[10] We observe the payment activity of these loans on properties located in the treatment areas until the end of 2014, or until the mortgage is terminated, whichever occurs first, and compare them to the payment choices of loans outside of the treatment areas.

2.2. Control and treated areas

EIA (2017) provides a list of counties that are located within each shale play. We classify counties that are located within the *Bakken*, *Eagle Ford*, *Haynesville*, *Marcellus*, *Niobrara* and *Utica* plays as treated areas. Fig. 2 shows a map of where these shale plays are located.[11] We consider all available fully-amortizing, first-lien, purchase mortgages on single family, owner-occupied properties originated between 2000 and 2006 located in these shale counties as our treatment group. In order to reduce the risk of our results being contaminated by spillover effects, mortgages on properties located in counties that are in states with shale activity, but that themselves do not contain shale oil and/or gas activity were removed from the list

[4] Agency securitized loans (e.g. Freddie Mac or Fannie Mae) and loans held on lenders' balance sheets are not privately securitized and therefore not included in the data.

[5] For purposes of this research, we consider post 2007 as the treatment time period consistent with EIA (2017). We observe mortgages originated as early as 2000 and follow these mortgages through termination (due to default or prepayment) or the end of 2014, whichever comes first. Due to a large drop in the oil price in 2015, the shale boom largely ended in mid-2015. Therefore, studying the time period from 2007 to 2014 is likely the most appropriate definition of the "boom" time period. We will consider the specific timing of the treatment in an alternative specification.

[6] There is, on average, a three month lag between loan origination and securitization.

[7] Detailed BBx data information is available at http://www.bbxlogic.com/data.htm.

[8] We look at only first-lien loans because a first mortgage is, by and large, the biggest loan against the property. Additionally, we are unable to connect the second lien loans we may observe in our dataset to the companion first lien against that property, since all information about the loan is attached to a unique loan identifier, not a property identifier. The exclusion of second lien loans, which are typically much riskier than first lien loans, likely understates the magnitude of our results.

[9] Although we do not know loan purpose with certainty, we exclude investment properties or vacation homes to the best of our ability as these are not the focus of our analysis. Multifamily properties comprise less than 1% of our sample; we exclude these properties to obtain the most homogeneous sample possible.

[10] Prior to 2000, we do not have a sufficient sample size of treated loans to observe.

[11] This map is based on the counties identified in EIA (2017). The Permian region, while included in the map, was excluded from the analysis because there were not a sufficient number of loans in these rural areas for analysis.

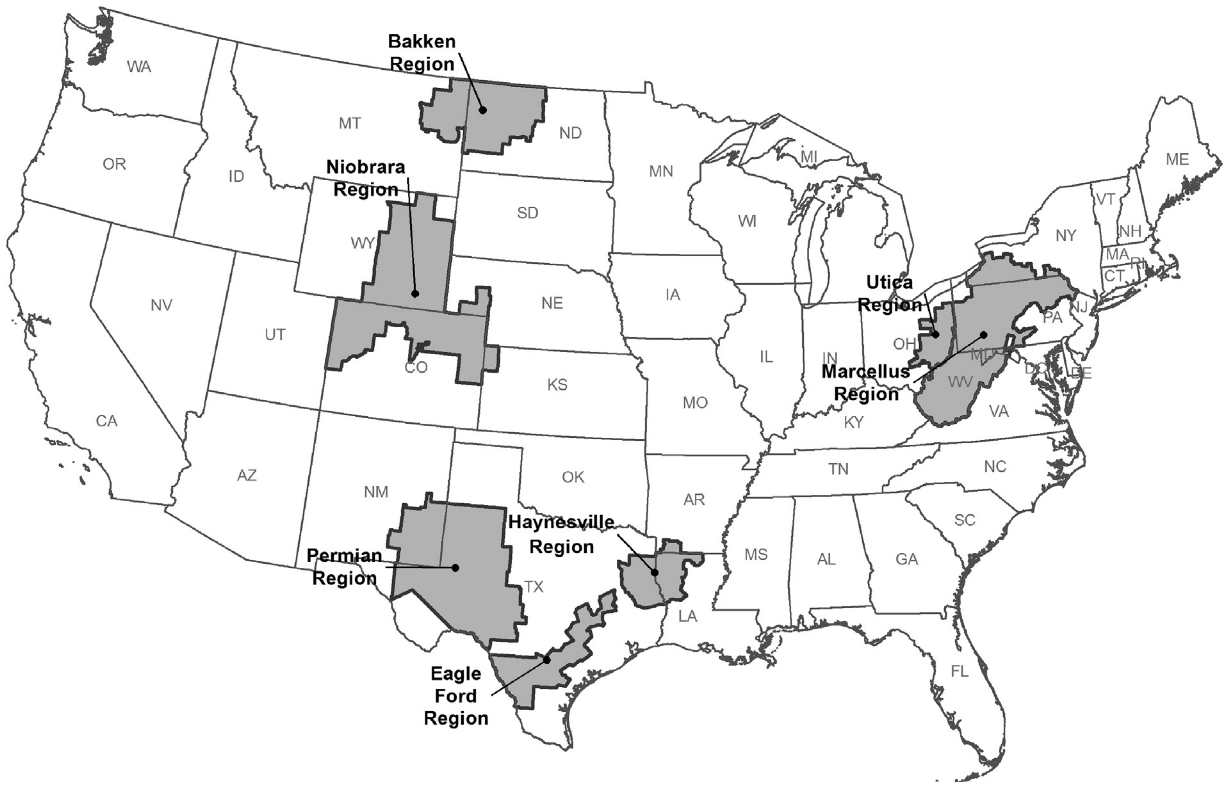


Fig. 2. U.S. shale plays.

of potential control areas. In addition, states that directly border counties with shale activity were also removed from the potential control group.[12] This approach is similar to the selection of treatment and control areas in James and Smith (2017), although we have taken a more conservative approach by eliminating entire states from the potential control group. This conservative approach has both pros and cons. Our approach allows us to rule out, for all intents and purposes, the possibility of spillovers into non-shale counties or cities that are in close proximity and therefore, may be impacted by shale activity. On the other hand, this choice of controls precludes us considering these spatial spillovers, as has been done in other research in the area, namely Feyrer et al. (2017). Given the research question at hand, we opt for a conservative approach to selecting a control group and leave any analysis of spatial spillovers in mortgage payments to future research.

For our main specifications, we employ propensity score matching to identify a control group of loans based on all observable origination characteristics from the population of mortgages in non-shale areas in our sample as well as employment and average earnings in the counties for which these loans are located.[13] In other words, we find a corresponding "control" loan for every loan originated in a shale area that has (a) similar mortgage characteristics and (b) is located in a county that has a similar labor force size and similar average earnings. As a robustness check, we choose 20 control groups randomly sampled from the entire universe of loans that originate in the pre-shale time period.[14] As we will show, estimated treatment effects are robust to different choices of control group, empirical specifications, and placebo tests.

2.3. Variables

Mortgage default is the dependent variable of interest. For robustness, we consider two alternative definitions referred to as "default" and "mild default." Mortgage default results from the borrower not paying the contractually obligated monthly payment in a timely manner. For each loan, we observe a time series of monthly payment records and therefore can observe when an individual borrower misses one or more payments. Missing a single payment could be a reflection of a short term liquidity problem, or even forgetfulness on the part of the borrower. However, missing more than one payment in a row is

[12] After applying these criteria, the potential control group comes from loans in the following twenty-nine states; AL, AZ, CA, CT, DE, FL, GA, HI, ID, IL, IN, IA, ME, MI, MN, MS, MO, NV, NH, NJ, NC, OK, OR, RI, SC, TN, VT, WA, and WI.

[13] County level employment and average wage data is from Census provided Quarterly Workforce Indicators (QWI).

[14] Simply pooling all mortgages nation-wide into a sample was not feasible due to computing constraints, as it includes more than 900 million remittance records.

indicative of a more serious financial problem. We will consider both mortgages that are just one month delinquent as well as mortgages with multiple consecutive missed payments.

First, we create a binary variable for default for a given loan, *i*, observed in time period, *t*, in which $default_{it} = 1$ if the loan is 90 or more days delinquent at the time of observation. If the borrower later makes sufficient back payments, i.e. they "catch up" on their payments, the loan can be reclassified in later periods ($default_{it} = 0$). We would expect loans in areas that experienced a shale boom to have a reduced probability of default as compared to similar mortgages in areas that have no shale discoveries. Next, we construct another binary variable for a less stringent definition of default. Instead of using 90 days delinquency as our threshold, we restrict our definition of default to mortgages that are 30 days or more delinquent, that is the borrower is one or more months behind on their mortgage payments. This is referred to as *mild default*. Additionally, as loans are prepaid for reasons such as moving or refinancing they leave our sample. We do not explicitly model prepayment in our main specifications which could lead to censoring concerns if borrowers in shale areas prepay at greater frequency than non-shale borrowers (perhaps due to increased property values or easier local credit markets). However, we argue that if this is the case, this would bias us against finding results on default. Therefore the baseline results on default we present can be thought of as a lower bound of the magnitude of this effect.

In all specifications, we include several control variables that are standard in the real estate finance literature. First, we construct a dummy variable for the loan term to control for differences in longer and shorter term loans. The most common loan terms are 30 year and 15 year mortgages. We also observe other loan terms, such as 10, 20, and 40 year loans. We create a dummy variable for loans 30 year terms or longer. Previous studies have found that compared to 30 year loans, 15 year loans have a lower probability of default (Quercia and Stegman, 1992) and we expect our results will be consistent with these earlier findings.

We also control for the type of loan; that is, whether the loan a fixed rate mortgage (FRM) or adjustable rate mortgage (ARM). In addition, we include a continuous variable for the current interest rate on each loan. By definition, for FRM loans, the interest rate is determined at origination and remains constant for the entire life of the mortgage. For ARM loans, the interest rate is adjusted during the life of the loan and therefore can vary over time for a specific loan. Typically a ARM borrower receives a rate at origination below that of a comparable fixed rate loan and the interest rate is locked in for a set, relatively short, period. At the end of the period, the interest rate resets according to a predetermined formula that is a function of current market rates. We control for initial leverage of the loan at the time of origination through the initial loan to value ratio (*LTV at origination*). This ratio relates the initial balance of the mortgage to the purchase price of the property. For example, a \$100,000 home purchased with a \$20,000 down payment and therefore a \$80,000 loan would have a LTV ratio of 0.8. All else equal, we would expect loans with high LTV ratios to have an increased risk of default.

We control for differences in borrower credit quality by including a continuous variable for FICO credit score at origination. These scores are a measure of borrower credit risk; specifically, FICO scores give an indication of the likelihood of a negative credit event for a borrower in the next year. Credit scores are reported in the U.S. by the three major credit bureaus, Experian, Equifax and TransUnion and typically range from a very poor score of less than 400 to a strong score of over 850 which corresponds with a very low risk borrower. For ease of interpretation in our regression models, we scale the FICO variable by dividing the borrower's FICO score by 100. We observe the FICO score of each mortgage at origination and expect this variable is negatively associated with mortgage default as has been shown in previous literature (Mester, 1997; Demyanyk and Hemert, 2011).

Finally, we include a series of fixed effects to control for additional unobservable loan heterogeneity. We include fixed effects for the month the loan is observed, the year the loan is observed, the year the loan is originated, and the servicer assigned at origination.[15]

2.3.1. Summary statistics

Summary statistics are presented for each of the six shale plays and their corresponding propensity score matched control groups in Table 1. Overall, our sample consists of 31,954 treated loans across all of the shale plays, corresponding to 1,681,336 monthly payment observations. By design the summary statistics on the origination characteristics for treated and control group are similar.[16] Overall, the average FICO score for our sample is 650 in the treated group and 649 in the control group. Although there is no universally accepted cutoff for subprime borrowers' credit scores, a commonly quoted threshold for a subprime FICO score in this period is 620, making our sample, on average, above the threshold for subprime lending. There is variation in average credit scores across the plays, ranging from an average of 629 for the treated group of loans in Haynesville to an average of 674 for the treated group of loans in Niobrara.

Overall, the loan to value ratio for the sample is about 85% for both the treated and control groups, indicated that on average, borrowers made a 15% down payment at the point of loan origination. Loans on properties in Niobara have the lowest average LTV (82%) while loans on properties in Utica have the highest (88%). Adjustable rate mortgages (ARMs) make

[15] We individually control for all servicers that service 1% or more of the total volume of loans nationwide in our sample. This construction results in the creation of 18 dummy variables, one for each large servicer, accounting for over 70% of the sample. All other smaller servicers are lumped together in a single category.

[16] We conduct our propensity score match on origination characteristics, so the total number of loans in both the control and treated groups are the same, but the number of payment observations is different.

Table 1
Summary statistics: baseline sample.

	All			Bakken			Eagle Ford			Haynesville		
	Sample Average	Std. Dev.	*N*	Sample Average	Std. Dev.	*N*	Sample Average	Std. Dev.	*N*	Sample Average	Std. Dev.	*N*
Origination characteristics												
FICO score												
Treated	650.1	71.0	31,954	636.5	60.8	147	632.4	66.8	1901	629.4	63.3	1171
PMatch control	649.2	68.0	31,954	644.4	69.9	147	630.7	63.3	1901	632.3	61.4	1171
Loan-to-value ratio (LTV) at origination												
Treated	85.6%	10.7%	31,939	86.3%	10.8%	147	85.2%	8.9%	1900	86.9%	10.3%	1171
PMatch control	85.2%	10.7%	31,905	86.1%	10.5%	147	85.9%	11.3%	1898	86.0%	10.6%	1171
Adjustable rate mortgage (ARM)												
Treated	45.8%	49.8%	31,954	47.6%	50.1%	147	32.7%	46.9%	1901	42.1%	49.4%	1171
PMatch control	44.3%	49.7%	31,954	38.8%	48.9%	147	31.6%	46.5%	1901	41.5%	49.3%	1171
30 Year mortgage												
Treated	96.2%	19.2%	31,954	95.9%	19.9%	147	95.4%	21.0%	1901	96.5%	18.4%	1171
PMatch control	96.8%	17.6%	31,954	97.3%	16.3%	147	96.9%	17.2%	1901	98.0%	14.2%	1171
Appraisal value												
Treated	$167,205	$173,594	31,954	$98,280	$56,836	147	$128,770	$115,483	1901	$131,449	$80,998	1171
PMatch control	$179,329	$205,795	31,954	$102,736	$58,462	147	$139,229	$107,619	1901	$132,546	$79,884	1171
County level employment												
Treated	221,155	248,403	31,954	$2,279	$203	147	70,666	35,536	1901	107,703	53,379	1171
PMatch control	207,792	396,429	31,954	17,749	13,906	147	46,040	71,119	1901	81,286	73,588	1171
County level average earnings												
Treated	$2,871	$442	31,954	$2,279	$203	147	$2,275	$175	1901	$2,547	$170	1171
PMatch control	$2,849	$600	31,954	$2,288	$233	147	$2,279	$199	1901	$2569	$238	1171
Time variant characteristics												
Default												
Treated	16.3%	36.9%	1,681,336	6.1%	23.8%	6544	12.6%	33.2%	122,590	24.0%	42.7%	68,657
PMatch control	20.3%	40.2%	1,477,978	16.2%	36.9%	7511	20.3%	40.2%	95,699	22.0%	41.4%	54,126
Mild default												
Treated	27.7%	44.7%	1,681,336	20.2%	40.1%	6544	27.8%	44.8%	122,590	37.9%	48.5%	68,657
PMatch control	31.4%	46.4%	1,477,978	28.1%	45.0%	7511	32.4%	46.8%	95,699	34.9%	47.7%	54,126
Interest rate												
Treated	7.5%	1.86%	1,681,336	8.4%	1.8%	6544	7.81%	1.70%	122,590	7.93%	1.81%	68,657
PMatch control	7.43%	1.96%	1,477,978	8.1%	1.8%	7511	7.80%	1.97%	95,699	7.66%	1.81%	54,126

	Marcellus			Niobrara			Utica		
	Sample Average	Std. Dev.	N	Sample Average	Std. Dev.	N	Sample Average	Std. Dev.	N
Origination characteristics									
FICO score									
Treated	641.9	71.4	13,832	674.0	66.8	10,733	630.0	67.3	4170
PMatch control	641.6	65.7	13,832	671.3	68.1	10,733	631.0	64.3	4170
Loan-to-value ratio (LTV) at origination									
Treated	87.3%	11.0%	13,828	82.1%	10.1%	10,728	88.3%	10.0%	4165
PMatch control	86.5%	10.5%	13,828	82.7	10.5%	10,719	86.7%	10.2%	4165
Adjustable rate mortgage (ARM)									
Treated	41.2%	49.3%	13,832	53.0%	49.9%	10,733	49.3%	50.0%	4170
PMatch control	39.5%	48.9%	13,832	52.1%	49.9%	10,733	47.1%	50.0%	4170
30 Year mortgage									
Treated	95.7%	20.3%	13,832	96.8%	17.7%	10,733	96.5%	18.5%	4170
PMatch control	97.1%	16.8%	13,832	96.2%	19.1%	10,733	97.0%	17.1%	4170
Appraisal value									
Treated	$124,354	$115,992	13,832	$258,115	$233,598	10,733	$105,341	$73,676	4170
PMatch control	$121,799	$85,311	13,832	$294,473	$303,672	10,733	$107,908	$64,081	4170
County level employment									
Treated	325,846	332,220	13,382	171,223	109,558	10,733	109,849	63,185	4170
PMatch control	299,363	571,445	13,382	159,840	128,338	10,733	97,458	107,645	4170
County level earnings									
Treated	$2,799	$353	13,382	$3,192	$440	10,733	$2,665	$188	4170
PMatch control	$2,792	$438	13,382	$3,138	$768	10,733	$2,656	$360	4170
Time variant characteristics									
Default									
Treated	17.2%	37.7%	781,628	11.7%	32.2%	486,224	23.2%	42.2%	215,693
PMatch control	21.4%	41.0%	644,221	18.0%	38.5%	482,896	21.9%	41.3%	193,525
Mild default									
Treated	29.5%	45.6%	781,628	20.0%	40.0%	486,224	35.2%	47.8%	215,693
PMatch control	33.5%	47.2%	644,221	26.7%	44.3%	482,896	34.6%	47.6%	193,525
Interest rate									
Treated	7.86%	1.90%	781,628	6.61%	1.49%	486,224	8.01%	1.83%	215,693
PMatch control	7.72%	1.89%	644,221	6.73%	1.90%	482,896	7.91%	1.86%	193,525

up 45.8% of our total treated population and 44.3% of the propensity score matched control group. The remainder of loans in the sample are fixed rate mortgages (FRMs). There is variation in the ARM portion of the mortgage market across our individual treated areas, ranging from 32.7% in Eagle Ford to 53.0% in Niobrara. In our matched sample, 30 year or longer mortgages are by far the most common loan term (96%), and this is consistent across all shale plays.[17] Although we do not match on the loan's current interest rate when creating our sample (interest rate can vary post-origination for adjustable rate loans), the current interest rate is about 7.4–7.5% in both the control and treated areas.

Our summary statistics for our outcomes of interest show both default and mild default is lower in the treated areas than the control areas. This is true overall as well as in five of the six shale plays. Pooled across all time periods and geographic areas, 16.3% of the treatment group observations have a severe default versus 20.3% for the control group. These differences are similar in magnitude using our mild default measure; 27.7% of the treatment group observations have a mild default versus 31.4% of observations in the control group.

Of course, these summary statistics only provide a snapshot over the entire sample period from 2000 to 2014. Fig. 3 illustrates the mortgage default and mild default rate in shale areas compared to the propensity score matched control groups to compare changes in the two groups before and after the shale boom began. As can be seen, the two groups have relatively similar levels of default and mild default before 2007, when the shale boom began. But after 2007, there is a divergence of default rates for loans in shale areas compared to non-shale areas. In particular, mortgages in treated areas experienced significant decreases in the probability of both mortgage default and mild default relative to control group mortgages.[18]

3. Empirical strategy

3.1. Cohort analysis

As a first empirical test, we will consider two "cohorts" of mortgages; specifically we examine mortgages that were originated in 2006 and those mortgages that were originated in 2005. That is one year and two years, respectively, before the treatment in 2007. For each loan in a shale area, we conduct a propensity score match from other loans that also originated within that year. These propensity scores are based on both loan characteristics at origination and county level labor market conditions in the pre-treatment time period. Specifically, loan level origination characteristics include the appraisal value of the home, LTV ratio, original interest rate, FICO score, whether the loan had a term longer or shorter than 30 years, and whether the mortgage is an ARM or FRM. County level employment and average earnings are observed in these respective years (QWI). Thus, these control mortgages originated in the same year, have similar origination characteristics, and are located in a county with a similar labor market size and average earnings.

We then observe the first occurrence, if ever, for which each mortgage experienced delinquency (or mild delinquency) and test whether mortgages in shale counties are less likely to have gone into delinquency during, or before, a given year. Thus, we test whether the loans in treated areas were less likely to have gone into default (and mild default per same definitions as above). This is illustrated in Eq. (1),

$$d_{i,c} = \alpha + \beta Shale_c \tag{1}$$

In this specification, $d_{i,c}$ is an indicator variable for whether the mortgage i in county c (that is either a shale or non-shale county), has ever entered into default during or before a given year. For all loans in the 2005 cohort (a total of 24,686 loans) we estimate this model for each year 2007–2014. Then, we repeat this analysis for the 28,453 loans that are in the 2006 cohort.[19]

This first empirical specification is appealing for its simplicity, but also because it mitigates concern for selection bias. Given that over time households may move or refinance their mortgages, the sample of pre-shale mortgages available to analyze in the post-shale periods could differ. This selection out of sample could impact results utilizing full sample of payments (as will be show hereafter), but will not impact this more simplistic analysis.

[17] In the universe of loans we observe, 15 and 30 year loans make up over 98% of the sample. Therefore, we use a dummy variable for loan term instead of including a continuous variable for loan term.

[18] Note that these mortgages originated between 2000 and 2006 and therefore these should not be interpreted as representative of all outstanding mortgages in the U.S. Furthermore, because the earliest date that these mortgages could have originated was January of 2000, the default rate at the very beginning of these samples is very close to zero. There are a few key drivers of this low default rate early in the period. In 2000, the only loans that are observed have been originated that year, so the sample size is relatively small. Additionally, in general, it is very uncommon for a mortgage to immediately default, unless the loan was improperly underwritten. In general, absent large price declines that wipe out borrowers' equity, there is a non-linear relationship between loan age and default. Default is low in initial periods, peaks in year 3–5 of the loan, and then declines as the borrower builds equity in the property (von and Furstenberg, 1969; Campbell and Dietrich, 1983). See Appendix Figs. A1–A6 for similar graphs corresponding to each shale area.

[19] For example, if a loan defaulted in 2009, it would be classified as in default for the remainder of the sample period even if the loan later becomes current or drops out of the sample due to foreclosure. Alternatively, if a loan prepaid in 2009, it would be classified as not in default for the remainder of the sample period, instead of exiting the sample at the date of prepayment.

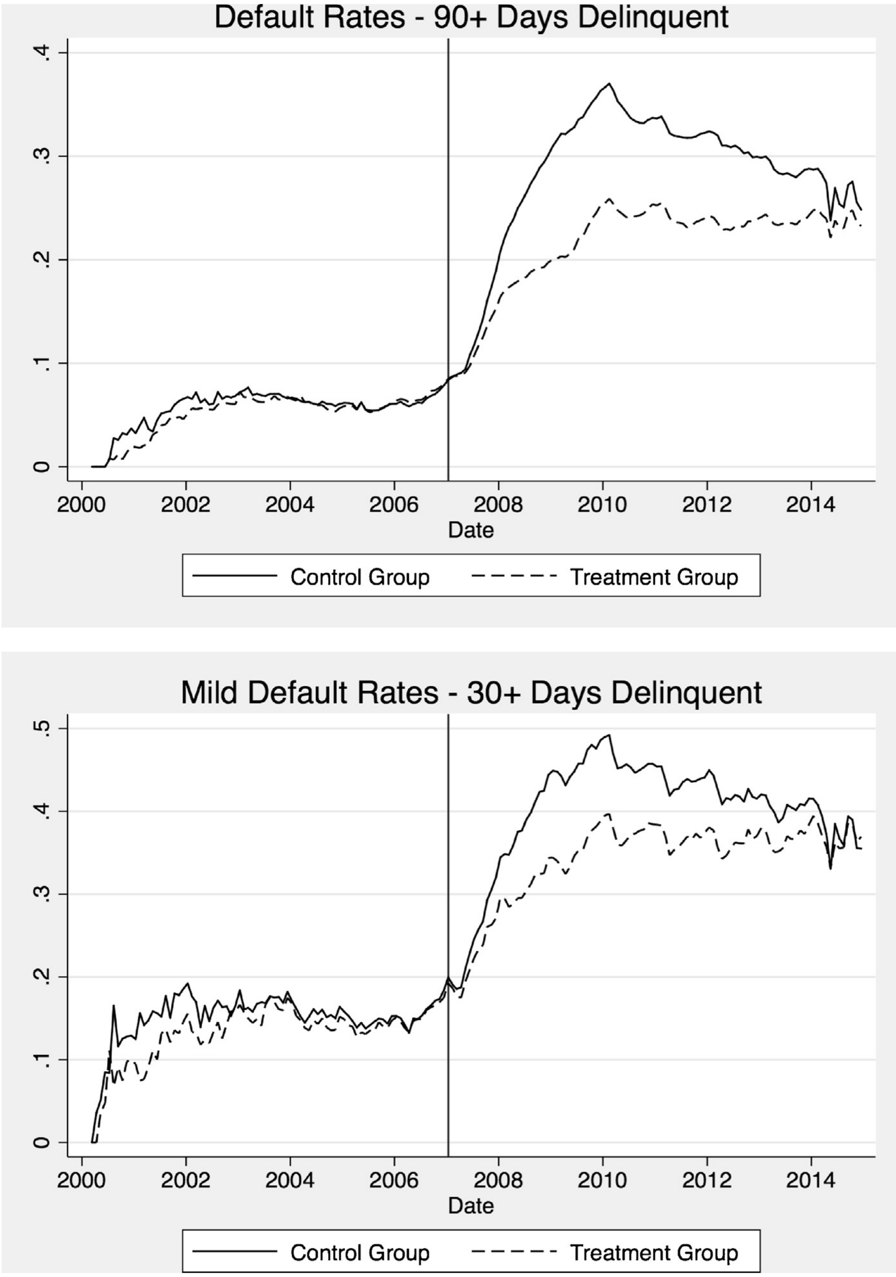


Fig. 3. Comparison of shale and non-shale areas mortgage default and mild default.

3.2. Differences-in-differences

While the first empirical specification is quite conservative in that it mitigates concern for attrition bias in the sample, it does not utilize the full sample of loans available, nor does it utilize the full panel of payment information. Thus, next we will estimate Eq. (2) that illustrates the commonly used difference-in-differences (DD) style estimation strategy used to test for the impact of shale oil and gas on mortgage default utilizing the full panel of data available.

$$d_{i,c,t} = \alpha + \delta(S_{Shale_c} \times Shale_t) + \beta S_{Shale_c} + X'_{i,t}\zeta + \gamma_1 D_s + \gamma_2 D_m + \gamma_3 D_y + \gamma_4 D_o + \varepsilon_{i,c,t} \tag{2}$$

where $d_{i,c,t}$ is the outcome of interest–mortgage default–for individual mortgage i in county c in month t. S_{Shale_c} is an indicator variable corresponding to counties with shale oil and/or gas activity (i.e. the treatment group) and is zero for the mortgages not located in one of the counties with shale activity. $Shale_t$ is an indicator variable that indicates the time periods after shale activity began. All of the shale plays, and therefore counties that EIA defines to have shale activity, saw increases in drilling starting around 2007 and this drilling activity continued until the end of 2014.[20] We use a logistic regression to estimate the change in the probability of default and mild default in each shale play. For ease of interpretation, we present all results as average marginal effects for each variable. For each model, we estimate standard errors clustered at the property zip code level.[21]

The vector X'_{it} contains control variables that are standard in the real estate finance literature, including the term of the mortgage, whether the mortgage has a fixed or adjustable interest rate, the initial interest rate of the mortgage (that is either fixed or varies over time based on the mortgage type), the FICO score of the borrower at origination and the loan to value ratio of the loan at origination. Additionally, all estimations include fixed effects for loan servicer D_s, month of observation D_m, year of observation D_y, and year of origination D_o.

Our primary results are obtained using this DD framework; we estimate the impact of the shale boom on default (and mild default) using our propensity-score matched control group as described in Section 2.1. The estimated δ provides us with the change in default in treated areas relative to non-treated areas during the boom, while controlling for a number of loan specific covariates including the loan term, interest rate, and a number of other origination characteristics.[22] We provide both geographically pooled estimated of this treatment effect as well as conduct separate estimations for each treated shale area.

3.3. Rig counts and value of production

Next, we index results to both rig counts and estimated value of oil and gas production. We do this, as there are two direct channels through which an oil and gas boom can stimulate a local economy. First, economic activity is stimulated because local landowners receive bonus and royalty checks for oil and gas production that occurs beneath their land. A bonus check is given to the landowner at the time that a lease is signed as a lump sum payment. But also, once production begins landowners receive royalty payments that is some share of the value of the oil and gas produced (typically 20–25%).[23] These royalty payments might only continue for a short time if the well is relatively unsuccessful, or can continue for years and even decades as the well continues along its long tail of production. Thus, when local residents receive (sometimes large) payments this can stimulate the local economy through spending.

The second direct channel through which oil and gas operations can stimulate a local economy is through the drilling activities themselves. In the case of the shale plays, the operator typically contracts out a service company to both drill the well and complete the hydraulic fracturing needed to stimulate the well to begin production. These workers will earn income directly, and then will spend some share of these earnings in the local economy.

Thus, policy makers interested in understanding likely implications of an oil and gas boom can scale the size of the shock to two benchmarks; rigs and value of production. Rig counts are a measurement of the current drilling activity, and value of production is an indicator of the amount of royalties going to landowners. It should be noted that these results do not intend to disentangle these two effects from one another, nor is it to describe the channel through which these effects occur, but will allow for scaling these effects to future oil and gas booms and compare results across studies in a standardized fashion.

Rig counts and oil and gas production are provided by EIA's Drilling Productivity Reports. We multiply oil and gas production by the WTI and Henry Hub spot prices, respectively, to obtain estimated value of production.[24] Eq. (3) shows the basic DD specification replacing rig counts with the post-treatment dummy. A corollary analysis is run for value of production.

$$d_{i,c,t} = \alpha + \delta(Rigs_{ct}) + \beta S_{Shale_c} + X'_{i,t}\zeta + \gamma_1 D_s + \gamma_2 D_m + \gamma_3 D_y + \gamma_4 D_o + \varepsilon_{i,c,t} \quad (3)$$

[20] Of course, the exact start time of the boom varies across shale plays. In the initial specification, we include 2007 as the start date for the shale boom, but we later present the year specific estimated treatment effects by shale play (see Fig. 4 and Figs. A7 and A12). We end the analysis at the end of 2014 for two reasons. First, our mortgage data availability only extends through the end of 2014. Additionally, in 2015 global oil prices dropped significantly, and therefore the "bust" plausibly began some time during 2015. Therefore, 2007–2014 is the best general time period that can be considered the "boom" or "treatment" time period.

[21] Our results are robust to our choice of clustering.

[22] In many lending outcomes, time to default is not linear in time. In many types of secured lending, such as mortgages, probability of default is low at origination and increases with loan age. However, as equity builds through loan repayment and/or price appreciation, probability of default declines as loan age increases until it reaches more or less a steady state where only idiosyncratic default occurs at some rate. Of course, the shape of the time to default distribution will depend on loan characteristics and current individual and local level economic conditions. One popular approach to handle this outside of an explicit hazard framework is to include higher order polynomials for loan age to control for this known relationship between loan age and default probability. This approach is used in work such as (Gross and Souleles, 2002; Elul et al., 2010; Elul, 2016), among many others. In each case the authors note that the results using this approach mirror preliminary results found using hazard models. Using this approach of including variables for loan age, loan age^2, ..., loan age^5, we find results consistent with our baseline model. These results for default and mild default over all plays are included in Appendix Table A1.

[23] The landowner for which the actual well is drilled also typically receives a "rental" payment that is the value of renting that land to the company for production. Most landowners, though, receive a bonus and royalty payment even though there is no actual drilling activity physically on their land.

[24] It should be noted that well-head prices, for oil in particular, can differ substantially across the country, especially during this time period of significant price differentials between the mid-continent and the gulf coast (Agerton and Upton, 2017). We use the WTI spot price, which was discounted significantly, in lieu of a gulf coast crude prices, such as Louisiana Light Sweet (LLS), in order to provide for a more conservative point estimate and to remove any judgement calls of assigning well-head prices to specific plays.

3.4. Mortgage prepayment

Next, we consider the possibility that our treatment and control groups also differ in their probability of mortgage payoff. Absent default, households may pay off their mortgages for three primary reasons: the borrower reaches the end of the loan term, the borrower chooses to relocate to a new property, or the borrower chooses to refinance their existing mortgage debt. Given the length of our sample period and distribution of loan terms, we have virtually no loans maturing in our sample. We cannot disentangle the other two motivations for mortgage payoff, given that each loan is only identified at a zip code level we cannot connect the borrower to any new property or financing they may acquire after loan prepayment. Therefore, we classify any observation currently non-delinquent borrower who repays their loan in full as a prepayment.

The simplest way of estimating multiple probabilities is by multinomial logistic regression, which has been employed in several other mortgage payment studies (Archer et al., 1996; Clapp et al., 2001, 2006; Elul et al., 2010; McCollum et al., 2015). Multinomial logit estimates a separate equation for each identified choice as shown in Eq. (4) where l is the identified choice which runs from 1 to J as shown in Eq. (5), i is the ith individual observation, x_{it} is a 1 by k vector containing the ith individual's explanatory variable data in the tth time period, and y_{it} is the ith individual's choice in the tth time period. However, the mortgage data contains information about individuals without differentiating information specific to each choice. Accordingly, one choice is not identified and becomes the base case. The other estimated equations are in terms of the differences between the respective choice and the base case. Let $\gamma^{(l)}$ represent the difference between the k by 1 parameter vector associated with the lth choice $\beta^{(l)}$ and the base case choice parameter values $\beta^{(0)}$.

$$Pr(y_{it} = l) = \frac{\exp(x_{it}\gamma^{(l)})}{1 + \sum_{j=1}^{J} \exp(x_{it}\gamma^{(j)})} \tag{4}$$

$$\gamma^{(l)} = \beta^{(l)} - \beta^{(0)}, \quad l = 1, \ldots, J \tag{5}$$

In this context, the base case is that the mortgage is current; the other cases are delinquent or fully prepaid.

3.5. Risk category

Next, we investigate if the expected reduction in default probability for the treatment group is concentrated in a subset of borrowers. We split our sample by credit scores and initial LTV ratios – two standard non-geographic measures of risk – and repeat our DD estimation. We first divide our full sample over four FICO credit score buckets: less than 620, 621–680, 681–739, and 740 or higher. Although there is no precise industry prescribed cut-off for what is considered to be a subprime or prime loan, these buckets roughly correspond to subprime, near-prime, prime, and super-prime credit categories, respectively. Additionally, we repeat this test for different loan-to-value (LTV) buckets at the time of origination. We estimate our model for each of the following LTV categories: ≤80%, 80–85%, 85–90%, 90–95%, >95%. These results will provide an idea of the type of borrowers that are most sensitive to changes in the probability of default associated with economic booms. For both tests, we estimate the regression model described in Eq. (2) for each risk bucket using the complete list of controls described in the base model.

3.6. Urban and rural

Additionally, even within plays, there is potentially substantial variation in local economic conditions, including job opportunities and housing market conditions. Therefore, one might hypothesize differential impacts for rural and urban areas if urban areas are more susceptible to housing market booms and busts than rural areas (Mota, 2016).[25] If our results are driven by differences in housing market conditions in our treatment and control groups, we would expect to find significant results in urban areas but not in rural areas. To test for this possibility, we divide our sample into population quartiles. Using pre-treatment county level population, we classify loans that are made in counties that are in the bottom quartile of population as *Rural* and loans that are made in counties that in the top quartile of population as *Urban*.

3.7. Robustness

Finally, we investigate the robustness of our main results by conducting a series of tests. First, we test the sensitivity of our results to alternative control groups by repeating the baseline DD estimation 20 times for each shale play, as well as the

[25] Given that our sample contains many rural areas, using popular housing price indices such as Case-Shiller to directly control for pre-treatment differences in local housing market conditions to form our samples in the propensity score match is impractical since these indices focus on major MSAs. However, we do include house price at loan origination as one of our matching variables.

Table 2
Cohort analysis – comparison of probability of mortgage default in shale and non-shale counties.

	(1) 2007	(2) 2008	(3) 2009	(4) 2010	(5) 2011	(6) 2012	(7) 2013
Default: 90+ days delinquent							
2005 cohort	−0.0192***	−0.0408***	−0.0534***	−0.0584***	−0.0614***	−0.0617***	−0.0616***
	(0.00424)	(0.00504)	(0.00539)	(0.00555)	(0.00563)	(0.00568)	(0.00572)
Observations	24,686	24,686	24,686	24,686	24,686	24,686	24,686
2006 cohort	−0.0310***	−0.0667***	−0.0871***	−0.0868***	−0.0842***	−0.0805***	−0.0801***
	(0.00360)	(0.00464)	(0.00511)	(0.00528)	(0.00537)	(0.00543)	(0.00546)
Observations	28,453	28,453	28,453	28,453	28,453	28,453	28,453
Mild default: 30+ days delinquent							
2005 cohort	−0.0174***	−0.0376***	−0.0471***	−0.0510***	−0.0548***	−0.0532***	−0.0515***
	(0.00542)	(0.00604)	(0.00622)	(0.00628)	(0.00630)	(0.00631)	(0.00632)
Observations	24,686	24,686	24,686	24,686	24,686	24,686	24,686
2006 cohort	−0.0322***	−0.0700***	−0.0824***	−0.0759***	−0.0702***	−0.0652***	−0.0615***
	(0.00486)	(0.00561)	(0.00580)	(0.00584)	(0.00585)	(0.00585)	(0.00584)
Observations	28,453	28,453	28,453	28,453	28,453	28,453	28,453

Mortgage default is defined as 90 days behind on mortgage payments. Mild default is defined as 30 days behind on mortgage payments.

overall universe of treated areas, using a randomly selected control group from the nationwide sample of loans in lieu of the propensity score matched sample.[26] All of these results are estimated for both definitions of default.

As a second robustness check, we consider the exact timing of the treatment effect by estimating the average difference in default rates for the treated and control mortgages using Eq. (6):

$$d_{i,c,t} = \alpha + \sum_{t=2003}^{2014} \delta_t (S_{Shale_c} \times D_t) + X'_{i,t}\zeta + \gamma_1 D_s + \gamma_2 D_m + \gamma_3 D_y + \gamma_4 D_o + \varepsilon_{i,c,t} \tag{6}$$

where again $d_{i,c,t}$ is the outcome of interest–mortgage default–for individual mortgage i in county c in month t. S_{Shale_c} is an indicator variable corresponding to counties with shale oil and/or gas activity (i.e. the treatment group) and is zero for the mortgages not located in one of the counties with shale activity and D_t is a dummy variable for each year from 2003 to 2014.

Consistent with all prior regressions, the vector X'_{ist} contains all control variables and all estimations include fixed effects for loan servicer D_s, month of observation D_m, year of observation D_y, and year of origination D_o. In this specification, δ_{2003}, δ_{2004}, . . ., δ_{2014} are the coefficient estimates of interest. Unlike the previous differences-in-difference estimation strategy, these coefficients simply show the estimated difference in default rates in treated and control areas after controlling for all covariates by year. In the years prior to the shale boom, we expect these coefficient estimates to be relatively small in magnitude. Once the shale boom begins, we expect that these coefficient estimates will become larger in magnitude. This will provide evidence of when the boom began (and possibly ended), for both the entire sample and for each shale area separately.

Finally, we employ three falsification tests. First, we implement a placebo treatment on randomly selected loans that are not located in areas impacted by the shale boom from our baseline control groups. Second, we repeat the placebo test, this time using the universe of treated loans and randomly assigning these loans to either the treatment or control group. Third, we perform a temporal placebo test where we only use the pre-treatment data and assign the treatment date to be 2004.

4. Results

Table 2 presents the results of the cohort analysis. For the 2005 cohort (loans originated in the calendar year 2005), loans that are located in shale areas are 1.9% less likely to default by 2007, 4.1% less likely to default by 2008, and by 2013 are about 6.2% less likely to have ever defaulted. Similarly, for the 2006 cohort, mortgages in shale areas were 3.1% less likely to default through the year 2007, about 6.7% less likely to default by 2008 and by 2013 were about 8.0% less likely to have ever defaulted. While specific point estimates differ for default and mild default, the pattern is the same. Each of these 28 estimated marginal effects is statistically significant at $p = .01$. These results show that of loans originated in the two years leading up to the boom, loans in shale counties had lower probabilities of later going into default compared to other mortgages in non-shale counties with similar origination characteristics and located in counties with a similar labor market size and average wages.

Table 3 shows our main results utilizing the full panel of data available with the propensity score matched control group for both default and mild default. For the overall sample, we estimate that the shale boom is associated with a 6.1% decrease

[26] Similar to the propensity score match sample, we only pull loans out of states that have no shale activity, and are not directly adjacent to a county with shale activity. Therefore, the potential control group comes from the following states; AL, AZ, CA, CT, DE, FL, GA, HI, ID, IL, IN, IA, ME, MI, MN, MS, MO, NV, NH, NJ, NC, OK, OR, RI, SC, TN, VT, WA, and WI.

Table 3
Impact of shale on mortgage default.

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Default: 90+ days delinquent							
Treatment effect	−0.0610***	−0.0908**	−0.0410**	−0.0594**	−0.0557***	−0.0655***	−0.0617***
	(0.00557)	(0.0444)	(0.0201)	(0.0239)	(0.00765)	(0.00848)	(0.0142)
Shale area	0.00297	−0.0809*	−0.0436**	0.0352*	−0.0109	0.00762	0.0519***
	(0.00515)	(0.0423)	(0.0190)	(0.0213)	(0.00679)	(0.00961)	(0.0128)
FICO score	−0.0882***	−0.0658***	−0.0934***	−0.107***	−0.0869***	−0.0780***	−0.101***
	(0.00209)	(0.0240)	(0.00846)	(0.0111)	(0.00304)	(0.00339)	(0.00607)
Interest rate	0.0217***	0.00264	0.0133***	0.0248***	0.0217***	0.0213***	0.0266***
	(0.000649)	(0.00651)	(0.00193)	(0.00289)	(0.000955)	(0.00118)	(0.00165)
LTV at origination	0.0282**	0.104	0.0980***	0.0805	−0.00925	0.0817***	−0.0511
	(0.0122)	(0.165)	(0.0367)	(0.0628)	(0.0182)	(0.0201)	(0.0396)
Adjustable rate mortgage	0.0673***	0.0286	0.0489***	0.0841***	0.0670***	0.0565***	0.0927***
	(0.00329)	(0.0408)	(0.0143)	(0.0126)	(0.00467)	(0.00563)	(0.00921)
30 Year mortgage	0.0269***	−0.0282	0.0811***	0.144***	0.0316***	−0.00244	0.00908
	(0.00814)	(0.0871)	(0.0269)	(0.0557)	(0.0120)	(0.0115)	(0.0230)
Observations	3,152,559	13,811	217,809	122,529	1,422,701	967,114	408,256
Mild default: 30+ days delinquent							
Treatment effect	−0.0570***	−0.0930*	−0.0544***	−0.0503**	−0.0436***	−0.0783***	−0.0485***
	(0.00531)	(0.0476)	(0.0161)	(0.0226)	(0.00733)	(0.00776)	(0.0144)
Shale area	−0.00605	−0.0796*	−0.0149	0.0218	−0.0291***	0.0133	0.0301**
	(0.00468)	(0.0440)	(0.0169)	(0.0178)	(0.00615)	(0.00930)	(0.0117)
FICO score	−0.161***	−0.142***	−0.175***	−0.187***	−0.163***	−0.142***	−0.173***
	(0.00226)	(0.0340)	(0.00853)	(0.0119)	(0.00338)	(0.00365)	(0.00632)
Interest rate	0.0231***	0.00408	0.0150***	0.0226***	0.0226***	0.0249***	0.0260***
	(0.000743)	(0.00872)	(0.00243)	(0.00290)	(0.00111)	(0.00141)	(0.00172)
LTV at origination	0.0789***	0.236	0.141***	0.158**	0.0363	0.129***	0.0141
	(0.0147)	(0.208)	(0.0484)	(0.0679)	(0.0223)	(0.0242)	(0.0462)
Adjustable rate mortgage	0.0589***	0.0456	0.0351*	0.0619***	0.0598***	0.0555***	0.0848***
	(0.00386)	(0.0512)	(0.0191)	(0.0134)	(0.00551)	(0.00651)	(0.0100)
30 Year mortgage	0.0399***	−0.0697	0.0862***	0.176***	0.0491***	−0.000657	0.0299
	(0.00871)	(0.0796)	(0.0287)	(0.0541)	(0.0128)	(0.0126)	(0.0261)
Observations	3,152,559	14,045	217,847	122,596	1422,701	967,114	408,256

Mortgage default is defined as 90 days behind on mortgage payments. Mild default is defined as 30 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale (2001–2006). Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages chosen using propensity score match from national sample in non-shale states. Parameters estimated with logistic regression. Marginal effects shown in table.

in the probability of mortgages of default. In addition, each of the six areas experienced a statistically significant decrease in the probability of mortgages default with estimated marginal effects ranging from 4.1% in Eagle Ford to 9.1% in Bakken. Similar results are observed for the less stringent 30+ day definition of default (i.e. mild default). Overall, mortgages in shale areas experience a 5.7% decrease in mild default, with point estimates in individual shale plays ranging from 4.4% in Marcellus to 9.3% in Bakken.

The mortgage specific control variables perform largely as expected. In the 90+ days default specification, borrower creditworthiness, as measured by FICO credit scores at origination, is negatively associated with default. A 100 point increase in FICO score is associated with, on average, a 8.8% reduction in the probability of default. A one point increase in the initial interest rate of the loan is associated with, on average, a 2.2% increase in the probability of default. A 30 year or longer loan is associated with, on average a 2.7% increase in the probability of default. The estimates for initial leverage are positive, but only statistically significant in three of the seven specifications. As compared to fixed rate loans, adjustable rate mortgages on average are 6.7% more likely to default and 30 year mortgages are 2.7% more likely to default. These results are similar for the 30+ days definition of delinquency; in this specification the magnitude of the effects for loan term and credit score is larger and the leverage variable is positive and statistically significant overall.

We next index results to rig counts and value of production. In Table 4, we find results consistent with our baseline differences in differences specification. Across all shale plays, we find that one hundred additional rigs is associated with a 3.2% and 2.7% decrease in default and mild default, respectively. Similarly, a one billion dollar increase in the value of oil and gas produced is associated with a contemporaneous 1.6% and 1.3% decrease in default and mild default, respectively.

Table 5 provides results for the competing outcomes of default and prepayment using a multinomial logistic model across all shale plays. The base case in this specification is that the loan is current; the measure of default is 30+ days of delinquency. We find that in this setting that includes prepayment as an alternative loan outcome that the shale boom is associated with a 5.69% decrease in the probability of 30+ days delinquency, virtually identical to our baseline specification result (5.70%). Interestingly, we find that there is a small, but significant impact on the probability of prepayment of 0.34%. An increase in prepayment probability is another positive economic outcome associated with the shale boom. Potentially

Table 4
Robustness to rig count and value of production.

	(1) Default	(2) Default	(3) Mild default	(4) Mild default
Rig counts	−0.0317***		−0.0268***	
	(0.00513)		(0.00455)	
Value of production		−0.0155***		−0.0128***
		(0.00318)		(0.00288)
Shale area	−0.0277***	−0.0395***	−0.0341***	−0.0435***
	(0.00414)	(0.00362)	(0.00391)	(0.00372)
FICO score	−0.0880***	−0.0876***	−0.161***	−0.161***
	(0.00209)	(0.00209)	(0.00228)	(0.00228)
Interest rate	0.0216***	0.0215***	0.0230***	0.0230***
	(0.000645)	(0.000648)	(0.000741)	(0.000741)
LTV at origination	0.0240**	0.0236*	0.0748***	0.0746***
	(0.0122)	(0.0122)	(0.0147)	(0.0147)
Adjustable rate mortgage	0.0666***	0.0675***	0.0585***	0.0593***
	(0.00327)	(0.00327)	(0.00386)	(0.00386)
30 Year mortgage	0.0268***	0.0270***	0.0399***	0.0399***
	(0.00813)	(0.00815)	(0.00871)	(0.00873)
Observations	3,152,559	3,152,559	3,152,559	3,152,559

Mortgage default is defined as 90 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages chosen using propensity score match from national sample in non-shale states. Parameters estimated with logistic regression.
Value or Production in Millions of USD. Rig counts in hundreds of rigs.

Table 5
Impact of shale on mortgage default and prepayment.

	(1) Default	(2) Prepayment
Treatment effect	−0.0569***	0.0034***
	(0.005309)	(0.0004)
Treatment area	−0.00618	−0.0025***
	(0.0047)	(0.0003)
FICO score	−0.1616***	0.0043***
	(0.0023)	(0.0002)
Interest rate	0.0230***	0.001451***
	(0.00744)	(0.00008)
LTV at origination	0.07911***	−0.00912***
	(0.01476)	(0.00106)
Adjustable rate mortgage	0.0598***	0.0033***
	(0.00386)	(0.00039)
30 Year mortgage	0.0398***	−0.0029***
	(0.0087)	(0.0004)
Observations	3,067,638	3,067,638

Mortgage default is defined as 30+ days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages chosen using propensity score match from national sample in non-shale states. Parameters estimated with multinomial logistic regression; base case is loan is current. Marginal effects shown in table.

borrowers in shale areas are able prepay by refinancing into more attractive loan terms or sell their homes and move to more expensive properties as a result of improvement in local economic conditions relative to the shale boom. However, we cannot disentangle different borrower motivations for loan prepayment using our dataset since we cannot observe the household's mortgage debt after the prepayment, so we cannot determine the channel through which the shale boom may lead to an increase in prepayment.

Table 6 presents results pooled across all geographic areas subset by risk category. We find a significant negative treatment effect for each of our four credit score categories. We find that the largest treatment effect of a 9.3% reduction in default probability in the 681–739 credit score bucket, closely followed by a 8.3% reduction in default probability for the 621–680 credit score bucket. These middle credit buckets, corresponding roughly to the universe of near-prime and prime borrowers in our sample, have treatment effects that are more than double those found in the lowest credit bin (−3.7%) and the highest credit bin (−4.0%). This result is intuitive; those with very high (low) credit scores have a low (high) probability of default, independent of local economic conditions; therefore the change in local economic conditions precipitated by the shale oil and gas boom has a relatively smaller marginal impact on their probability of default than those with average default risks.

Similarly, we find significant treatment effects for each of our LTV categories. However, borrowers with high leverage experience relatively smaller reductions in the their default probabilities (3.6–4.9%) as compared to the 8.3% reduction in

Table 6
Estimated treatment effects by loan risk level.

	Treatment effect	Standard error	*N*
Credit score			
<620	−3.74%***	.009	1,163,889
621–680	−8.38%***	0.010	908,976
681–739	−9.33%***	0.012	650,395
≥740	−4.01%***	.0122	416,586
Loan-to-value ratio			
>95%	−3.95%***	0.009	844,632
90.01–95%	−3.58%**	0.012	318,861
85.01–90%	−3.56%**	0.011	349,278
80.01–85%	−4.93%**	0.012	147,591
≤80%	−8.27%***	0.008	1,773,395

Treatment group pooled from all shale areas. Control group comes from propensity match control loans for which loans fall in same category of risk.

Table 7
Comparison of estimated treatment effects.

	All	Bakken	Eagle Ford	Haynesville	Marcellus	Niabrara	Utica
Default							
Propensity match control group	−6.1%***	−9.1%**	−4.1%**	−5.9%***	−5.6%***	−6.6%***	−6.2%***
Random control group (Min)	−7.9%***	−5.6%	−7.6%***	−4.8%**	−8.8%***	−5.1%***	−8.7%***
Random control group (Mean)	−8.5%	−14.9%	−9.2%	−7.6%	−9.4%	−6.2%	−10.0%
Random control group (Max)	−9.1%***	−19.9%***	−10.9%***	−11.4%***	−10.0%***	−6.8%***	−11.1%***
Mild default							
Propensity match control group	−5.7%***	−9.3%*	−5.4%***	−5.0%**	−4.4%***	−7.8%***	−4.9%***
Random control group (Min)	−7.2%***	−4.9%	−9.9%***	−5.4%***	−7.2%***	−5.7%***	−6.8%***
Random control group (Mean)	−7.7%	−15.6%	−11.3%	−7.4%	−7.9%	−6.5%	−8.1%
Random control group (Max)	−8.2%***	−21.3%***	−13.4%***	−10.0%***	−8.6%***	−7.1%***	−9.5%***

Minimum and maximums in absolute values. Mean is average of treatment effects from 20 random control groups. All values pulled from Tables A2, 3 and A3.

default probability those borrowers who purchased their homes with at least 20% equity (LTV ≤ = 80). These results provide some insight into the relative impact of the shale boom for borrowers of different risk categories, but it is important to note that across the spectrum, of both metrics of borrower level risk, we find that the shale boom had a consistent significant negative impact on the probability of default. [27]

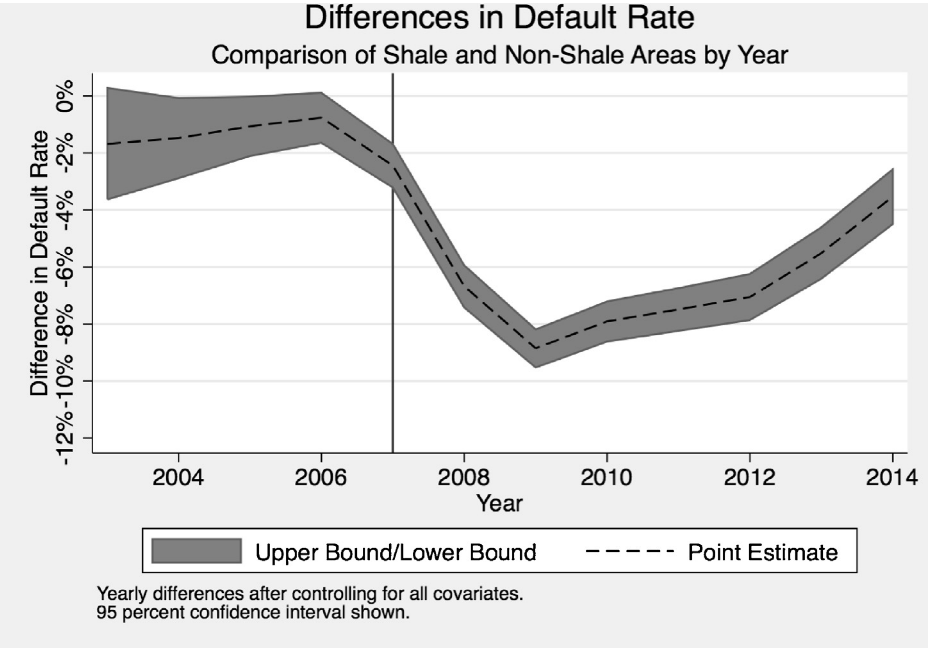
Next, we test the sensitivity of our results to alternative control groups by randomly choosing 20 control groups. The process of generating a random control group is performed 20 times for each of the 6 shale plays as well as 20 times for the overall sample. For each of these iterations, we estimate a treatment effect. These results are summarized and presented alongside the baseline results in Table 7. Using the propensity score matched control group, the estimated treatment effect on mortgage default and mild default, is 6.1% and 5.6% respectively. Both of these estimates are actually conservative relative to the estimated treatment effects simply using the randomly selected control groups. For default, these estimates range from 7.9% to 9.1%, compared to the propensity score match control group specification providing an estimated treatment effect of 6.1%. This pattern is consistently observed across shale plays, with the exception of Niobrara which has very similar average treatment effects using the propensity score match control group and the random control groups.

Appendix Table A2 shows the results for default using 20 random control groups taken from national sample of loans. Out of the 280 estimated treatment effects whose estimates we present in Appendix Table A2, all but one are negative and statistically significant at the p = .05 level and all but four are statistically significant at the p = .01 level. The average treatment effect across all iterations for each play ranges from 6.2% in Niabrara to 14.9% in Bakken with an average treatment effect of 8.5% in the sample pooling all 6 shale plays. Appendix Table A3 presents the same analysis for the mild default variable. These results are similar to the results for the 90+ days of delinquency specification. Overall, these results show that our main results is robust to many different selections of control groups.

For all specifications up to this point, we have simplistically set the treatment time period starting in 2007 and extending until the end of 2014, the most recent date currently available in the data. But more realistically, the shale boom could have started, peaked, and declined in different shale plays at different points. Therefore, next we assess the timing of these effects.

Appendix Tables A4 and A5 show the marginal effects associated with the coefficients $\delta_{2003}, \delta_{2004} \ldots \delta_{2014}$ estimated using Eq. (6) for both default and mild default. From 2003 to 2006, the shale areas have a slightly lower default rate after controlling

[27] We also investigated the extent to which results might differ in urban and rural areas. Table A6 presents results pooled across all geographic areas subset by pre-treatment county level population. We find that the magnitude of the treatment effect on urban areas (9.0%) is approximately twice that of the treatment effect on rural areas (4.4%). These results are consistent with Mota (2016) who also finds that urban areas are more susceptible to housing market booms and busts than rural areas.



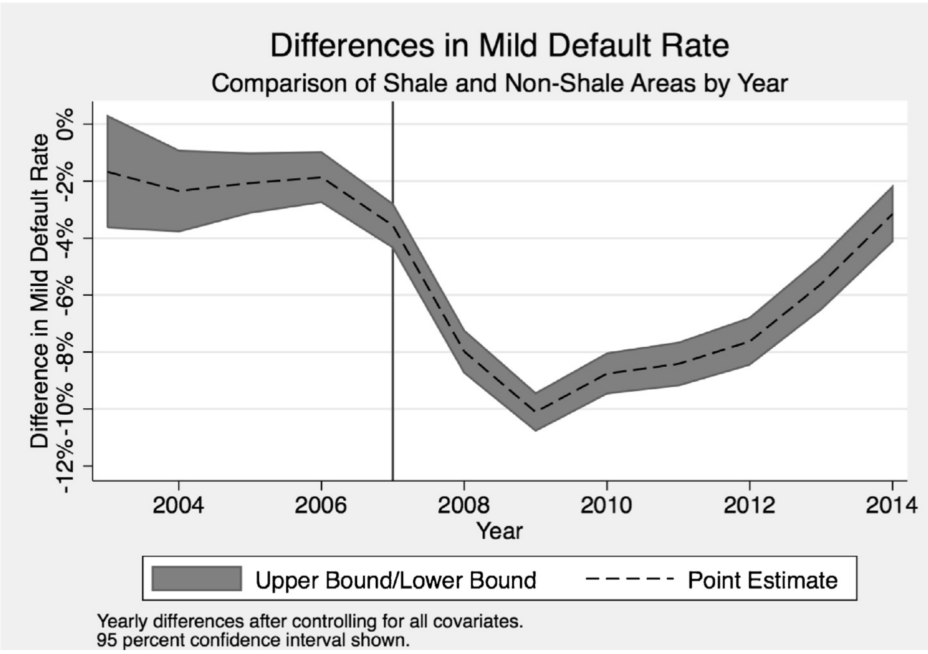


Fig. 4. Comparison of average difference in default rates of mortgages in shale and non-shale areas by year.

for covariates between about 1% and 2%. But starting in 2007, the default rate in the shale areas begins to decline and reaches its peak in 2009, when the shale areas has a default rate almost 9% lower than the control area, implying a treatment effect of about 7–8% at the boom's peak (subtracting out the pre-boom difference between the groups). The default rates steadily begin to converge once again from 2010 to 2014. By 2014, the shale areas have less than a 4% difference in default rates, down from the 9% difference observed at the peak of the book in 2009. A similar pattern is observed for mild default. These coefficient estimates and $p = .95$ confidence intervals are illustrated in Fig. 4.

These results vary by geographic area. For example, in Table A4, Haynesville, Marcellus, and Utica all reached a peak in the magnitude of their respective treatment effects in 2009, and as of the end of sample period in 2014 the treatment effect is no longer statistically significant in any of these plays. On the other hand, Bakken, Eagle Ford and Niabrara still have large negative statistically significant treatment effects through 2014. This is not surprising, given that Haynesville, Marcellus, and Utica are primarily gas plays and gas prices began to fall in 2010, whereas Bakken, Eagle Ford, and Niabrara are primarily oil plays and oil prices did not begin to decline until mid-2015, after the end of our sample period.

Lastly, we perform three sets of placebo tests in Appendix Tables A7–A9. Table A7 shows the results from the first placebo test, in which we randomly assign loans in actually treated areas into either the treatment or control group. We expect to find no consistent effect of this placebo treatment effect on either default or mild default. Across the 14 regressions that span both default definitions and each shale play, eight of the estimated treatment effects are negative while the remaining six are positive. None of these fourteen estimated treatment effects are statistically significant. For the second placebo test, we take all mortgages chosen with the propensity score match that are similar to mortgages in shale areas. We randomly assign half of these non-treated loans a placebo treatment, and keep the other half as the control group for this placebo test. These results, presented in Table A8, provide further evidence that our main findings are robust; out of the 14 regressions, four are statistically significant and with two of the statistically significant treatment effects as positive and two as negative. In total, nine of these placebo treatment effects are positive with the other five negative.

Finally, in Table A9 we present results of a placebo test using only pre-shale data. We assign all loans originated 2001–2003 to the pre-treatment group and all loans originated 2004–2006 to the post-treatment group; out of these 14 regressions 7 are positive and the other 7 are negative. In sum, Tables A7 and A9 provide 42 placebo treatment effects for default and mild default over each shale play separately as well as for the aggregated sample. In total, 20 of these are negative and the other 22 are positive.

5. Conclusion

The technological innovations that enabled oil and gas to be extracted from shale geological formations provides a natural experiment that can be used to test the impact of this shock on a number of outcomes of interest. In this research, we examine the impact of the shale boom on mortgage payment activity of households who resided in these areas before the oil and gas production began. We find that the shale boom lead to, on average, a 6% reduction on the probability of default. We estimate that every one hundred rigs in operation are associated with a 3.2% and 2.7% decrease in default and mild default, respectively. Similarly, a one billion dollar increase in the value of oil and gas produced is associated with a 1.6% and 1.3% decrease in default and mild default, respectively. While point estimates vary across areas, we find evidence of reductions in mortgage default in six shale areas relative to plausible control groups.

These results are largest for near prime and prime borrowers and relatively lower for sub-prime and super-prime borrowers. These positive effects are largest for borrowers with lower (less than 80%) loan to value ratios at origination. Additionally, we find a small, but significant, increase on the probability of prepayment (0.3%).

The divergence in mortgage default rates begins in 2007, peaking in 2009, and reaching near convergence by 2014. Thus, there is a clear pattern showing the booms beginning, peak, and approaching the end that likely ended past our sample time period with the significant drop in global oil prices.[28]

The implications of this research are multifaceted. First, economists are in general interested in how transitory shocks can impact savings behavior, which can have long term implications. We have shown that a labor market shock can impact mortgage default for households who purchased their home in the pre-shock period. While we only identify an example of how a positive shock can decrease default, it is also plausible that a negative shock might have an opposite effect.

Results of this research can have long term implications for not only households that go into default, but also for society at large. The recent financial crisis is a somber reminder of this reality. It is important for policy makers to be able to (a) be aware of how these labor market shocks impact households' mortgage payment activity, (b) have a reasonable idea of the magnitude of these shocks, (c) identify the type of borrowers that are most susceptible to these shocks. This paper provides a causal estimate of local labor market shocks on mortgage payment activity, and is unique in that we examine not just one labor market boom in one area, but instead six different booms in six different parts of the country that had different timings and magnitudes of these shocks. In addition, we identify the types of borrowers with varying sensitivities to these shocks.

There is still substantial room for future research in this area. While we identify a decrease in mortgage default associated with a plausibly exogenous shock, we do not explore the specific channel. It could be that workers with higher earnings are less likely to miss a mortgage payment. It could also be that households located in these areas before the boom begins do not necessarily see increases in earnings, but instead if they do find themselves in a situation where they have to miss mortgage payments, they can quickly sell the property for a gain instead of going into default. Thirdly, it could be that these booms cause decreases in the time it takes for job search. Thus, a homeowner who loses a job might be able to find a job more quickly, thus reducing mortgage default. Teasing out these relative effects would require specific information on individual households in these areas in addition to the individual mortgage payment activity used in this study, and therefore is beyond the scope of this research.

The resource boom studied in this analysis coincided with the Great Recession. Thus, this is a particularly interesting time to study the sensitivity of these housing markets to oil and gas booms. On one hand, our results provide direct evidence that housing markets in shale areas were at least partially insulated from the widespread negative effects of the contemporaneous

[28] Our analysis only extends through the end of 2014.

nationwide housing crisis. However, future research might be interested in whether these sensitivities differ in "normal" times for mortgage markets when aggregate default levels would be lower.

As a final potential extension for future research, we note that we take a conservative approach to defining our treatment and control areas. While this approach is helpful in avoiding spillover effects that might bias treatment effects, it inhibits us from examining the spatial nature of these effects in counties adjacent to areas with significant shale oil and gas activity. We also leave this question to future research.

Understanding how economic conditions impact mortgage default has substantial implications for the U.S. and global economy. Research in this vein can augment policy makers' understanding of how mortgage default behavior is attributable to changes in economic conditions.

Acknowledgements

We thank seminar participants at Louisiana State University, Baruch College CUNY, Sam Houston State University, Southern Economic Association Meeting and AREUEA National Meeting. We also thank Ryan Decker, Briggs Depew, Ozkan Eren, Naci Mocan, Kelley Pace, Christian Raschke and two anonymous referees for their thoughtful feedback. Research support was provided by PSC-CUNY Research Grant 60197-00-48.

Appendix A.

Table A1
Impact of shale on mortgage default: PMatch control group, including age polynomial.

	(1) Default-all	(2) Mild default-all
Treatment effect	−0.0555***	−0.0530***
	(0.00585)	(0.00544)
Treatment area	−0.00314	−0.0107**
	(0.00564)	(0.00492)
FICO score	−0.0929***	−0.166***
	(0.00213)	(0.00228)
Interest rate	0.0193***	0.0211***
	(0.000628)	(0.000742)
LTV at origination	0.0408***	0.0925***
	(0.0126)	(0.0152)
Adjustable rate mortgage	0.0648***	0.0571***
	(0.00329)	(0.00384)
30 Year mortgage	0.0369***	0.0479***
	(0.00974)	(0.0104)
Observations	3,067,638	3,067,638

Mortgage default is defined as 90 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages chosen using propensity score match from national sample in non-shale states. Parameters estimated with logistic regression. Marginal effects shown in table.

Table A2
Estimated treatment for 20 random control groups: default.

Iteration	All	Bakken	Eagle Ford	Haynesville	Marcellus	Niabrara	Utica
1	−.085***	−.135***	−.093***	−.055***	−.093***	−.061***	−.108***
2	−.085***	−.161***	−.010***	−.095***	−.093***	−.058***	−.107***
3	−.087***	−.179***	−.106***	−.059***	−.098***	−.062***	−.094***
4	−.091***	−.156***	−.109***	−.080***	−.097***	−.067***	−.111***
5	−.085***	−.123**	−.076***	−.114***	−.095***	−.060***	−.104***
6	−.084***	−.155***	−.085***	−.069***	−.096***	−.062***	−.088***
7	−.083***	−.180***	−.088***	−.082***	−.093***	−.058***	−.095***
8	−.084***	−.137***	−.010***	−.084***	−.090***	−.061***	−.102***
9	−.085***	−.169***	−.098***	−.071***	−.090***	−.065***	−.096***
10	−.084***	−.056	−.076***	−.081***	−.091***	−.064***	−.104***
11	−.080***	−.187***	−.102***	−.065***	−.091***	−.051***	−.101***
12	−.086***	−.141***	−.090***	−.085***	−.096***	−.061***	−.095***
13	−.087***	−.128***	−.084***	−.077***	−.096***	−.065***	−.101***
14	−.089***	−.143***	−.087***	−.089***	−.010***	−.063***	−.110***
15	−.079***	−.117**	−.096***	−.048**	−.088***	−.059***	−.087***
16	−.085***	−.161***	−.089***	−.077***	−.096***	−.062***	−.090***
17	−.086***	−.191***	−.080***	−.096***	−.092***	−.066***	−.097***
18	−.083***	−.076*	−.105***	−.061***	−.091***	−.060***	−.104***
19	−.086***	−.182***	−.082***	−.060***	−.096***	−.062***	−.108***
20	−.089***	−.199***	−.094***	−.082***	−.095***	−.068***	−.099***
Mean	−.085	−.149	−.092	−.076	−.094	−.062	−.100
Std. Dev.	.003	−.037	.010	.016	.003	.004	.007
Min	−.079	−.056	−.076	−.048	−.088	−.051	−.087
Max	−.091	−.199	−.109	−.114	−.100	−.068	−.111

Mortgage default is defined as 90 days behind on mortgage payments. Controls show in prior results also estimated in all regressions but not shown for purposes of brevity. Statistical significance based on zip code level clustered standard errors. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale (2001–2006). Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages are randomly chosen out of entire population of loans in the US; 20 random control groups are pulled for each regression run. Control group size chosen to match the number of treated loans in each regression. Parameters estimated with logistic regression. Marginal effects shown in table. Minimum and maximums in absolute values.

Table A3
Estimated treatment for 20 random control groups: mild default.

Iteration	All	Bakken	Eagle Ford	Haynesville	Marcellus	Niabrara	Utica
1	−.078***	−.129***	−.110***	−.063***	−.080***	−.065***	−.089***
2	−.077***	−.152***	−.119***	−.095***	−.076***	−.063***	−.084***
3	−.079***	−.213***	−.133***	−.054***	−.084***	−.064***	−.075***
4	−.081***	−.190***	−.123***	−.070***	−.079***	−.071***	−.082***
5	−.078***	−.133***	−.099***	−.100***	−.081***	−.064***	−.082***
6	−.076***	−.149***	−.107***	−.067***	−.081***	−.064***	−.072***
7	−.076***	−.201***	−.118***	−.080***	−.076***	−.063***	−.078***
8	−.075***	−.125***	−.119***	−.081***	−.077***	−.062***	−.078***
9	−.076***	−.181***	−.114***	−.071***	−.077***	−.067***	−.075***
10	−.076***	−.049	−.099***	−.070***	−.078***	−.067***	−.084***
11	−.076***	−.208***	−.134***	−.056***	−.079***	−.057***	−.085***
12	−.078***	−.177***	−.113***	−.096***	−.078***	−.067***	−.076***
13	−.080***	−.154***	−.107***	−.078***	−.082***	−.068***	−.086***
14	−.081***	−.151***	−.106***	−.083***	−.086***	−.064***	−.095***
15	−.072***	−.141***	−.107***	−.064***	−.072***	−.063***	−.073***
16	−.075***	−.143***	−.104***	−.067***	−.076***	−.066***	−.068***
17	−.079***	−.172***	−.102***	−.088***	−.079***	−.066***	−.087***
18	−.075***	−.076*	−.119***	−.055***	−.076***	−.065***	−.081***
19	−.077***	−.178***	−.100***	−.072***	−.077***	−.065***	−.091***
20	−.082***	−.195***	−.117***	−.072***	−.083***	−.071***	−.085***
Mean	−.077	−.156	−.113	−.074	−.079	−.065	−.081
Std. Dev.	.002	.042	.011	.013	.003	.003	.007
Min	−.072	−.049	−.099	−.054	−.072	−.057	−.068
Max	−.082	−.212	−.134	−.100	−.086	−.071	−.095

Mild default is defined as 30 days behind on mortgage payments. Controls show in prior results also estimated in all regressions but not shown for purposes of brevity. Statistical significance based on zip code level clustered standard errors. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale (2001–2006). Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages are randomly chosen out of entire population of loans in the US; 20 random control groups are pulled for each regression run. Control group size chosen to match the number of treated loans in each regression. Parameters estimated with logistic regression. Marginal effects shown in table. Minimum and maximums in absolute values.

Table A4
Difference in default rates in shale and non-shale areas by year.

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Treatment effect 2003	−0.0168*	0.0251	−0.0866**	−0.0714	−0.0416***	0.0108	0.0473*
	(0.00998)	(0.0695)	(0.0375)	(0.0488)	(0.0147)	(0.0162)	(0.0263)
Treatment effect 2004	−0.0148**	−0.0324	−0.0469	−0.0207	−0.0361***	0.0139	0.0157
	(0.00720)	(0.0607)	(0.0298)	(0.0449)	(0.0107)	(0.0125)	(0.0184)
Treatment effect 2005	−0.0107**	−0.0821	−0.0606**	−0.0300	−0.0266***	0.0202**	0.0112
	(0.00535)	(0.0666)	(0.0239)	(0.0269)	(0.00761)	(0.00960)	(0.0134)
Treatment effect 2006	−0.00761*	−0.169***	−0.0722***	−0.00206	−0.0320***	0.0226***	0.0280***
	(0.00448)	(0.0472)	(0.0161)	(0.0193)	(0.00645)	(0.00751)	(0.0106)
Treatment effect 2007	−0.0245***	−0.221***	−0.0786***	0.0194	−0.0507***	−0.00326	0.0134
	(0.00387)	(0.0407)	(0.0106)	(0.0143)	(0.00511)	(0.00644)	(0.00861)
Treatment effect 2008	−0.0668***	−0.230***	−0.0955***	0.00454	−0.0863***	−0.0585***	−0.0166**
	(0.00378)	(0.0361)	(0.0123)	(0.0135)	(0.00433)	(0.00610)	(0.00775)
Treatment effect 2009	−0.0885***	−0.197***	−0.0986***	−0.0430***	−0.0925***	−0.0868***	−0.0520***
	(0.00338)	(0.0390)	(0.0100)	(0.0154)	(0.00465)	(0.00462)	(0.00834)
Treatment effect 2010	−0.0791***	−0.219***	−0.0913***	−0.0306*	−0.0798***	−0.0781***	−0.0471***
	(0.00357)	(0.0408)	(0.0106)	(0.0170)	(0.00558)	(0.00445)	(0.00967)
Treatment effect 2011	−0.0749***	−0.196***	−0.111***	−0.0169	−0.0698***	−0.0788***	−0.0371***
	(0.00381)	(0.0497)	(0.0106)	(0.0195)	(0.00570)	(0.00491)	(0.0106)
Treatment effect 2012	−0.0705***	−0.199***	−0.120***	−0.0184	−0.0496***	−0.0941***	−0.0283**
	(0.00415)	(0.0517)	(0.0127)	(0.0190)	(0.00601)	(0.00573)	(0.0117)
Treatment effect 2013	−0.0551***	−0.158***	−0.134***	0.0128	−0.0262***	−0.0953***	−0.00544
	(0.00460)	(0.0531)	(0.0135)	(0.0199)	(0.00624)	(0.00630)	(0.0129)
Treatment effect 2014	−0.0353***	−0.150***	−0.106***	0.0296	−0.000527	−0.0832***	−0.00352
	(0.00487)	(0.0542)	(0.0139)	(0.0190)	(0.00663)	(0.00694)	(0.0137)
FICO score	−0.0896***	−0.108***	−0.0928***	−0.0964***	−0.0915***	−0.0836***	−0.0993***
	(0.00163)	(0.0209)	(0.00636)	(0.00813)	(0.00254)	(0.00241)	(0.00476)
Interest rate	0.0228***	0.00390	0.0168***	0.0238***	0.0213***	0.0257***	0.0267***
	(0.000526)	(0.00642)	(0.00161)	(0.00241)	(0.000805)	(0.000861)	(0.00134)
LTV at origination	0.0161	0.0680	0.0293	0.00747	−0.00252	0.0649***	−0.0467
	(0.00997)	(0.140)	(0.0336)	(0.0459)	(0.0152)	(0.0163)	(0.0312)
Adjustable rate mortgage	0.0677***	0.0329	0.0535***	0.0910***	0.0665***	0.0571***	0.0931***
	(0.00278)	(0.0341)	(0.0138)	(0.0127)	(0.00406)	(0.00439)	(0.00780)
30 Year mortgage	0.0191***	−0.0457	0.0109	0.0895**	0.0184*	0.00973	0.0366*
	(0.00627)	(0.0611)	(0.0229)	(0.0394)	(0.00964)	(0.00984)	(0.0196)
Observations	4,818,286	19,704	283,063	185,835	1,965,975	1,743,385	620,104

Mortgage default is defined as 90 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages chosen using propensity score match from national sample in non-shale states. Parameters estimated with logistic regression. Marginal effects shown in table.

Table A5
Difference in mild default rates in shale and non-shale areas by year.

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Treatment effect 2003	−0.0167**	−0.0386	−0.0308	−0.0707*	−0.0490***	0.0148	0.0305
	(0.00818)	(0.0745)	(0.0296)	(0.0399)	(0.0124)	(0.0134)	(0.0232)
Treatment effect 2004	−0.0234***	−0.111*	−0.0159	−0.00960	−0.0540***	0.0202**	−0.0181
	(0.00626)	(0.0613)	(0.0265)	(0.0300)	(0.00911)	(0.0102)	(0.0165)
Treatment effect 2005	−0.0207***	0.00267	−0.0502***	−0.0221	−0.0457***	0.0142*	−0.000199
	(0.00462)	(0.0606)	(0.0176)	(0.0204)	(0.00667)	(0.00815)	(0.0119)
Treatment effect 2006	−0.0186***	−0.173***	−0.0532***	−0.0130	−0.0470***	0.0117*	0.0128
	(0.00411)	(0.0434)	(0.0145)	(0.0166)	(0.00579)	(0.00709)	(0.00989)
Treatment effect 2007	−0.0356***	−0.231***	−0.0664***	0.00706	−0.0645***	−0.0119*	−0.00265
	(0.00391)	(0.0416)	(0.0128)	(0.0138)	(0.00517)	(0.00662)	(0.00874)
Treatment effect 2008	−0.0798***	−0.249***	−0.0974***	−0.00434	−0.0936***	−0.0735***	−0.0446***
	(0.00414)	(0.0408)	(0.0135)	(0.0168)	(0.00497)	(0.00703)	(0.00876)
Treatment effect 2009	−0.101***	−0.221***	−0.119***	−0.0483**	−0.102***	−0.103***	−0.0604***
	(0.00402)	(0.0483)	(0.0145)	(0.0203)	(0.00578)	(0.00558)	(0.0103)
Treatment effect 2010	−0.0876***	−0.186***	−0.111***	−0.0191	−0.0850***	−0.0906***	−0.0534***
	(0.00415)	(0.0489)	(0.0157)	(0.0199)	(0.00652)	(0.00535)	(0.0116)
Treatment effect 2011	−0.0842***	−0.129**	−0.130***	−0.0197	−0.0758***	−0.0917***	−0.0488***
	(0.00439)	(0.0649)	(0.0152)	(0.0231)	(0.00684)	(0.00586)	(0.0124)

Table A5 (*Continued*)

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Treatment effect 2012	−0.0763***	−0.176***	−0.128***	−0.000408	−0.0522***	−0.107***	−0.0281**
	(0.00477)	(0.0629)	(0.0163)	(0.0235)	(0.00731)	(0.00651)	(0.0132)
Treatment effect 2013	−0.0561***	−0.102	−0.135***	0.0368	−0.0246***	−0.0973***	−0.0102
	(0.00515)	(0.0636)	(0.0177)	(0.0239)	(0.00741)	(0.00707)	(0.0145)
Treatment effect 2014	−0.0315***	−0.0277	−0.0931***	0.0500**	0.0128*	−0.0885***	−0.01000
	(0.00530)	(0.0689)	(0.0171)	(0.0232)	(0.00759)	(0.00778)	(0.0149)
FICO score	−0.161***	−0.192***	−0.181***	−0.173***	−0.167***	−0.144***	−0.168***
	(0.00175)	(0.0205)	(0.00622)	(0.00943)	(0.00279)	(0.00274)	(0.00497)
Interest rate	0.0252***	0.00620	0.0191***	0.0237***	0.0226***	0.0307***	0.0268***
	(0.000598)	(0.00809)	(0.00196)	(0.00251)	(0.000939)	(0.000972)	(0.00144)
LTV at origination	0.0656***	0.0529	0.0693	0.0533	0.0422**	0.114***	0.00993
	(0.0117)	(0.149)	(0.0434)	(0.0507)	(0.0178)	(0.0199)	(0.0356)
Adjustable rate mortgage	0.0605***	0.0815*	0.0424**	0.0872***	0.0595***	0.0558***	0.0846***
	(0.00315)	(0.0443)	(0.0175)	(0.0133)	(0.00466)	(0.00503)	(0.00841)
30 Year mortgage	0.0317***	−0.0955	0.0266	0.131***	0.0347***	0.0184*	0.0447**
	(0.00690)	(0.0785)	(0.0244)	(0.0432)	(0.0107)	(0.0107)	(0.0219)
Observations	4,818,286	19,704	283,063	186,012	1,965,975	1,743,385	620,104

Mortgage default is defined as 30 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages chosen using propensity score match from national sample in non-shale states. Parameters estimated with logistic regression. Marginal effects shown in table.

Table A6
Impact of shale on mortgage default by pre-treatment county population.

	(1) <100,000 (rural)	(2) 100,000–250,000	(3) >250,000 (urban)
Default: 90+ days delinquent			
Treatment effect	−0.0437***	−0.0525***	−0.0900***
	(0.00764)	(0.0106)	(0.0114)
Treatment area	−0.00351	0.00859	0.000563
	(0.00668)	(0.00925)	(0.0110)
FICO score	−0.0840***	−0.0953***	−0.0897***
	(0.00307)	(0.00391)	(0.00381)
Interest rate	0.0165***	0.0283***	0.0242***
	(0.000907)	(0.00132)	(0.00111)
LTV at origination	0.0252	0.0462*	0.0273
	(0.0177)	(0.0252)	(0.0221)
Adjustable rate mortgage	0.0609***	0.0613***	0.0726***
	(0.00500)	(0.00641)	(0.00578)
30 Year mortgage	0.0253**	0.0230	0.0350**
	(0.0119)	(0.0152)	(0.0156)
Observations	1,328,712	813,212	1,010,635
Mild default: 30+ days delinquent			
Treatment effect	−0.0389***	−0.0557***	−0.0759***
	(0.00750)	(0.0103)	(0.0102)
Treatment Area	−0.00859	0.00293	−0.0181*
	(0.00634)	(0.00847)	(0.00994)
FICO score	−0.163***	−0.161***	−0.162***
	(0.00353)	(0.00424)	(0.00414)
Interest rate	0.0186***	0.0311***	0.0237***
	(0.00115)	(0.00142)	(0.00123)
LTV at origination	0.0607***	0.122***	0.0780***
	(0.0217)	(0.0295)	(0.0260)
Adjustable rate mortgage	0.0546***	0.0528***	0.0639***
	(0.00617)	(0.00685)	(0.00679)
30 Year mortgage	0.0376***	0.0327**	0.0525***
	(0.0133)	(0.0161)	(0.0162)
Observations	1,328,712	813,212	1,010,635

Mortgage default is defined as 90 days behind on mortgage payments. Mild default is defined as 30 days behind on mortgage payments. Zip code level clustered standard errors shown. Geographic classification is based on pre-treatment county-level population. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale (2001–2006). Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Control mortgages chosen using propensity score match from national sample in non-shale states. Parameters estimated with logistic regression. Marginal effects shown in table.

Table A7
Placebo test – using loans in treated areas.

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Default: 90+ days delinquent							
Placebo treatment effect	−0.00657	0.00291	0.0132	−0.00431	−0.000860	−0.0187	−0.0191
	(0.00709)	(0.0841)	(0.0260)	(0.0362)	(0.0103)	(0.0129)	(0.0190)
Placebo treatment area	0.00646	0.0144	−0.0259	−0.00780	0.00709	0.0141	0.0149
	(0.00648)	(0.0768)	(0.0232)	(0.0328)	(0.00922)	(0.0119)	(0.0169)
FICO score	−0.0813***	−0.0839**	−0.0953***	−0.106***	−0.0782***	−0.0784***	−0.0856***
	(0.00294)	(0.0386)	(0.0117)	(0.0184)	(0.00441)	(0.00455)	(0.00870)
Interest rate	0.0232***	0.00249	0.0164***	0.0226***	0.0250***	0.0230***	0.0235***
	(0.000953)	(0.0110)	(0.00346)	(0.00433)	(0.00139)	(0.00158)	(0.00233)
LTV at origination	0.0336*	0.462*	0.0967	0.0738	−0.0129	0.116***	−0.0448
	(0.0182)	(0.254)	(0.0631)	(0.100)	(0.0279)	(0.0300)	(0.0528)
Adjustable rate mortgage	0.0814***	0.00162	0.0763***	0.0863***	0.0883***	0.0709***	0.0916***
	(0.00485)	(0.0664)	(0.0188)	(0.0257)	(0.00705)	(0.00839)	(0.0133)
30 Year mortgage	0.0342**	−0.100	0.141**	0.224***	0.0333	−0.00791	0.0239
	(0.0146)	(0.0730)	(0.0571)	(0.0732)	(0.0216)	(0.0178)	(0.0387)
Observations	1,473,462	7,268	95,387	53,957	641,797	481,267	193,086
Mild default: 30+ days delinquent							
Placebo treatment effect	−0.000789	0.0563	0.00915	0.0465	−0.00500	0.00359	−0.0125
	(0.00675)	(0.0820)	(0.0193)	(0.0363)	(0.00926)	(0.0142)	(0.0153)
Placebo treatment area	0.00201	0.0166	−0.0327**	−0.00565	0.00152	0.00741	0.00825
	(0.00568)	(0.0522)	(0.0160)	(0.0376)	(0.00790)	(0.0103)	(0.0150)
FICO score	−0.173***	−0.197***	−0.183***	−0.199***	−0.176***	−0.145***	−0.190***
	(0.00295)	(0.0244)	(0.0120)	(0.0146)	(0.00451)	(0.00509)	(0.00791)
Interest rate	0.0212***	0.00636	0.0143***	0.0218***	0.0185***	0.0258***	0.0281***
	(0.00101)	(0.0105)	(0.00275)	(0.00344)	(0.00154)	(0.00199)	(0.00239)
LTV at origination	0.0677***	−0.0606	0.165***	0.136	0.0301	0.106***	0.0125
	(0.0209)	(0.122)	(0.0529)	(0.103)	(0.0314)	(0.0381)	(0.0674)
Adjustable rate mortgage	0.00158	0.0401	0.0326***	0.0135	−0.00501	0.00871	−0.00959
	(0.00354)	(0.102)	(0.00778)	(0.0171)	(0.00483)	(0.00620)	(0.00957)
30 Year mortgage	0.0401***	−0.117	0.0612*	0.175**	0.0506***	0.00474	0.0123
	(0.0105)	(0.0917)	(0.0359)	(0.0742)	(0.0154)	(0.0173)	(0.0319)
Observations	1,679,097	6,417	122,452	68,591	780,904	485,437	215,170

Mortgage default is defined as 90 days behind on mortgage payments. Mild default is defined as 30 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Half of mortgages in shale areas randomly assigned to "treatment" group; the other half of mortgages in shale areas assigned to "control" group. Parameters estimated with logistic regression. Marginal effects shown in table.

Table A8
Placebo test – using loans in control areas.

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Default: 90+ days delinquent							
Placebo treatment effect	0.0130**	0.0183	0.00645	0.0382	0.00593	0.0253***	0.00422
	(0.00605)	(0.0560)	(0.0165)	(0.0267)	(0.00934)	(0.00906)	(0.0185)
Placebo treatment area	−0.00883	0.0127	−0.00249	−0.0382*	0.000462	−0.0221***	−0.00577
	(0.00569)	(0.0291)	(0.0199)	(0.0216)	(0.00833)	(0.00777)	(0.0188)
FICO score	−0.102***	−0.0993***	−0.0982***	−0.121***	−0.103***	−0.0828***	−0.128***
	(0.00279)	(0.0232)	(0.0125)	(0.0149)	(0.00406)	(0.00434)	(0.00782)
Interest rate	0.0194***	0.00264	0.0113***	0.0271***	0.0182***	0.0196***	0.0284***
	(0.000881)	(0.0111)	(0.00185)	(0.00378)	(0.00131)	(0.00167)	(0.00240)
LTV at origination	0.0154	0.00201	0.123***	0.0220	−0.00646	0.0522*	−0.0835
	(0.0163)	(0.131)	(0.0309)	(0.0925)	(0.0240)	(0.0290)	(0.0581)
Adjustable rate mortgage	0.00842***	0.0673	0.0316***	0.0219	0.00495	0.0141***	−0.0102
	(0.00300)	(0.0465)	(0.00583)	(0.0145)	(0.00431)	(0.00433)	(0.00850)
30 Year mortgage	0.0318***	−0.0356	0.0448*	0.148**	0.0390***	0.00749	−0.00398
	(0.00935)	(0.0844)	(0.0237)	(0.0636)	(0.0138)	(0.0159)	(0.0278)
Observations	1,679,097	6,252	122,422	68,572	780,904	485,437	214,870

Table A8 (*Continued*)

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Mild default: 30+ days delinquent							
Placebo treatment effect	−0.00165	−0.118***	−0.0427**	−0.0198	0.00472	−0.00505	0.0167
	(0.00595)	(0.0394)	(0.0186)	(0.0286)	(0.00899)	(0.00904)	(0.0182)
Placebo treatment area	0.00139	0.0738	0.0436***	0.0119	−0.00540	0.00343	−0.00851
	(0.00490)	(0.0464)	(0.0160)	(0.0264)	(0.00770)	(0.00709)	(0.0150)
FICO score	−0.173***	−0.189***	−0.183***	−0.199***	−0.176***	−0.145***	−0.190***
	(0.00295)	(0.0266)	(0.0120)	(0.0148)	(0.00450)	(0.00504)	(0.00789)
Interest rate	0.0212***	0.0101	0.0142***	0.0215***	0.0185***	0.0258***	0.0281***
	(0.00101)	(0.0101)	(0.00277)	(0.00339)	(0.00154)	(0.00200)	(0.00241)
LTV at origination	0.0677***	0.0971	0.172***	0.137	0.0302	0.106***	0.0126
	(0.0209)	(0.108)	(0.0532)	(0.0998)	(0.0315)	(0.0380)	(0.0675)
Adjustable rate mortgage	0.00158	0.0487	0.0324***	0.0139	−0.00502	0.00864	−0.00975
	(0.00354)	(0.106)	(0.00797)	(0.0164)	(0.00482)	(0.00615)	(0.00949)
30 Year mortgage	0.0401***	−0.0785	0.0584*	0.170**	0.0506***	0.00479	0.0114
	(0.0105)	(0.0865)	(0.0350)	(0.0759)	(0.0154)	(0.0170)	(0.0320)
Observations	1,679,097	6,417	122,452	68,591	780,904	485,437	215,170

Mortgage default is defined as 90 days behind on mortgage payments. Mild default is defined as 30 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Half of mortgages chosen using propensity match that are not in shale areas are randomly assigned to "treatment" group; the other half of mortgages in non-shale areas assigned to "control" group. Parameters estimated with logistic regression. Marginal effects shown in table.

Table A9
Placebo test – using pre-2006 data.

	(1) All	(2) Bakken	(3) Eagle Ford	(4) Haynesville	(5) Marcellus	(6) Niabrara	(7) Utica
Default: 90+ Days Delinquent							
Treatment effect	0.00249	−0.159**	−0.0363*	0.0261	−0.00714	0.0146	0.0429***
	(0.00524)	(0.0652)	(0.0196)	(0.0246)	(0.00641)	(0.00908)	(0.0134)
Post 2004	0.107***	0.309***	0.0830***	0.134***	0.118***	0.0835***	0.112***
	(0.00562)	(0.0975)	(0.0208)	(0.0300)	(0.00767)	(0.0102)	(0.0159)
FicoScoreOriginationCalc	−0.0529***	−0.183***	−0.0702***	−0.0757***	−0.0547***	−0.0417***	−0.0484***
	(0.00359)	(0.0552)	(0.0133)	(0.0256)	(0.00539)	(0.00556)	(0.00989)
Interest rate	0.0212***	−0.0401**	0.0180***	0.0178***	0.0197***	0.0215***	0.0309***
	(0.00130)	(0.0166)	(0.00485)	(0.00555)	(0.00178)	(0.00222)	(0.00359)
OrigLTVRatioCalc	0.0890***	0.419	0.209***	0.218	0.0597**	0.117***	0.0685
	(0.0207)	(0.263)	(0.0718)	(0.134)	(0.0287)	(0.0343)	(0.0669)
ARMInd	−0.00499	−0.238***	0.0212	−0.000419	−0.00473	−0.0147	0.000386
	(0.00606)	(0.0722)	(0.0213)	(0.0374)	(0.00828)	(0.0110)	(0.0193)
morth30year	0.0237**	−0.303***	0.0792**	0.0357	0.0259*	−0.00221	0.0248
	(0.0101)	(0.0587)	(0.0347)	(0.0477)	(0.0149)	(0.0139)	(0.0296)
Observations	363,532	1,431	24,205	12,276	184,330	86,925	52,906
default: 30+ days delinquent							
Treatment effect	−0.00776	−0.184***	−0.0238	0.0345	−0.0278***	0.0247**	0.0160
	(0.00660)	(0.0592)	(0.0269)	(0.0291)	(0.00837)	(0.0124)	(0.0171)
Post 2004	0.150***	0.432***	0.142***	0.189***	0.166***	0.0919***	0.179***
	(0.00660)	(0.0732)	(0.0232)	(0.0351)	(0.00893)	(0.0125)	(0.0182)
FicoScoreOriginationCalc	−0.120***	−0.235***	−0.146***	−0.149***	−0.125***	−0.103***	−0.105***
	(0.00449)	(0.0870)	(0.0161)	(0.0277)	(0.00682)	(0.00709)	(0.0119)
Interest rate	0.0337***	−0.0568**	0.0303***	0.0333***	0.0323***	0.0366***	0.0444***
	(0.00167)	(0.0227)	(0.00661)	(0.00712)	(0.00235)	(0.00302)	(0.00453)
OrigLTVRatioCalc	0.131***	0.324	0.154	0.429***	0.104***	0.127***	0.119
	(0.0262)	(0.477)	(0.0979)	(0.156)	(0.0359)	(0.0425)	(0.0883)
ARMInd	−0.0195**	−0.186*	0.00253	−0.0550	−0.0156	−0.0291*	−0.0176
	(0.00790)	(0.0985)	(0.0366)	(0.0524)	(0.0107)	(0.0150)	(0.0234)
morth30year	0.0308**	−0.287***	0.0978**	0.0932	0.0211	0.0134	0.0528
	(0.0125)	(0.0827)	(0.0491)	(0.0752)	(0.0177)	(0.0200)	(0.0337)
Observations	363,532	1,651	24,367	12,408	184,826	87,099	52,906

Mortgage default is defined as 90 days behind on mortgage payments. Zip code level clustered standard errors shown. Yearly, monthly, origination year, and loan servicer fixed effects estimated in all regressions, but output not shown in table. All mortgages originated pre-shale. Treatment time period post 2007. Treated areas include mortgages in areas in counties with shale production. Only payment records from 2001 to 2006 included. Placebo pre-treatment period is 2001–2003, post-treatment period is 2004–2006. Parameters estimated with logistic regression.

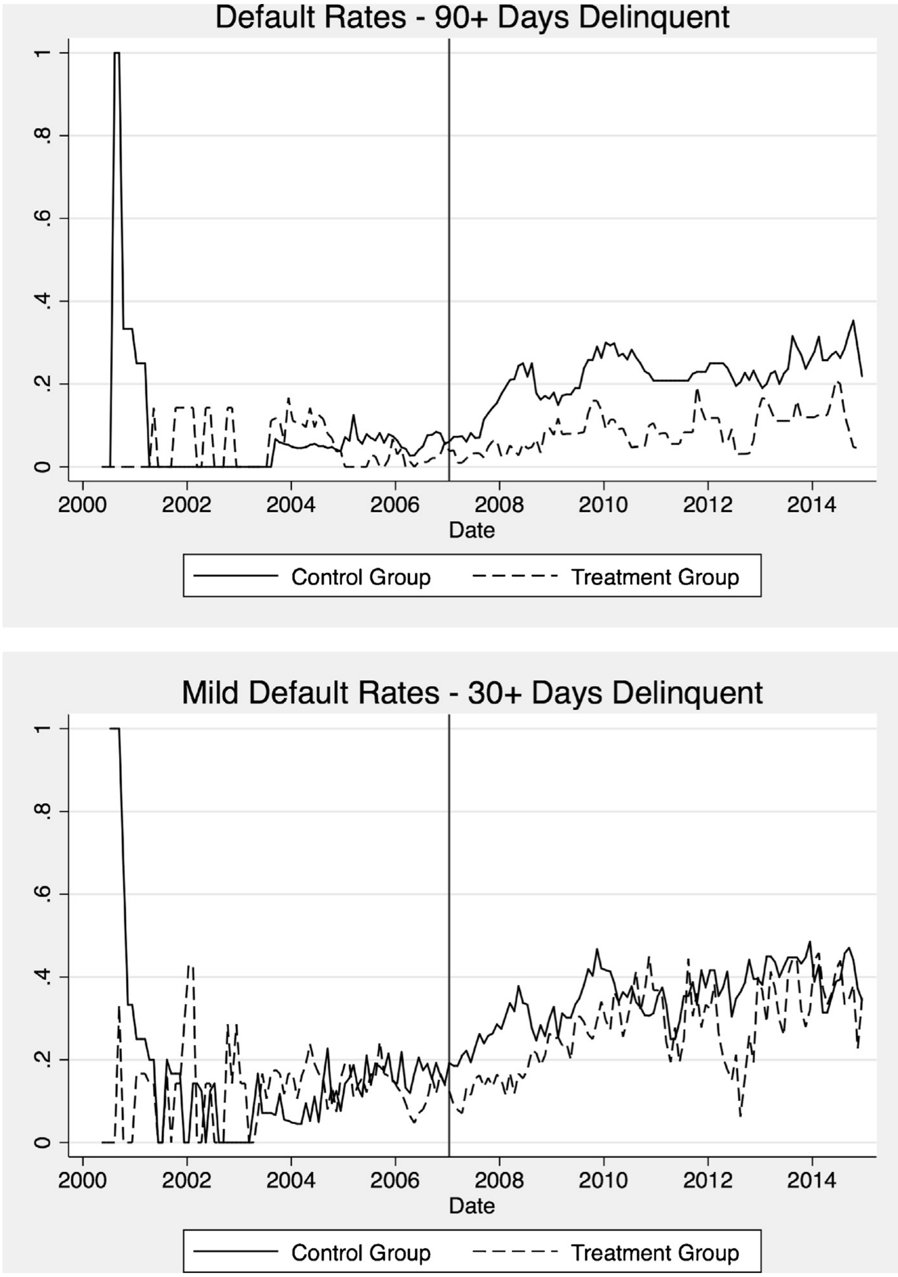
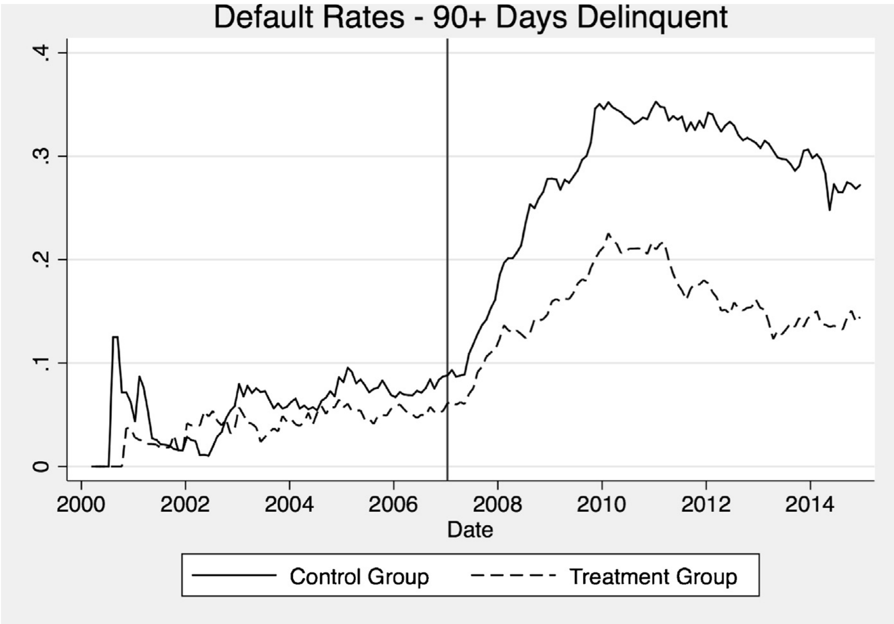


Fig. A1. Bakken: comparison of shale and non-shale areas mortgage default and mild default.



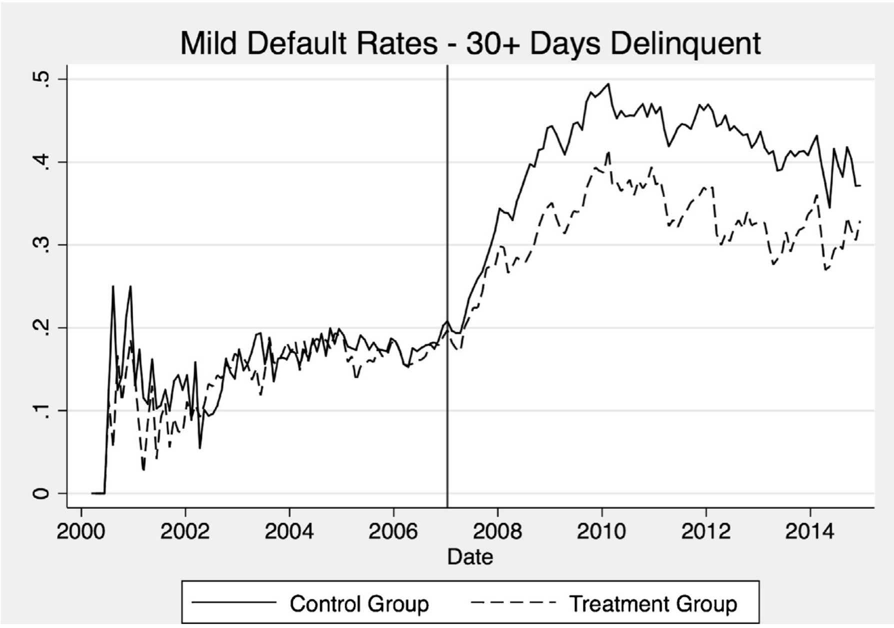


Fig. A2. Eagle Ford: comparison of shale and non-shale areas mortgage default and mild default.

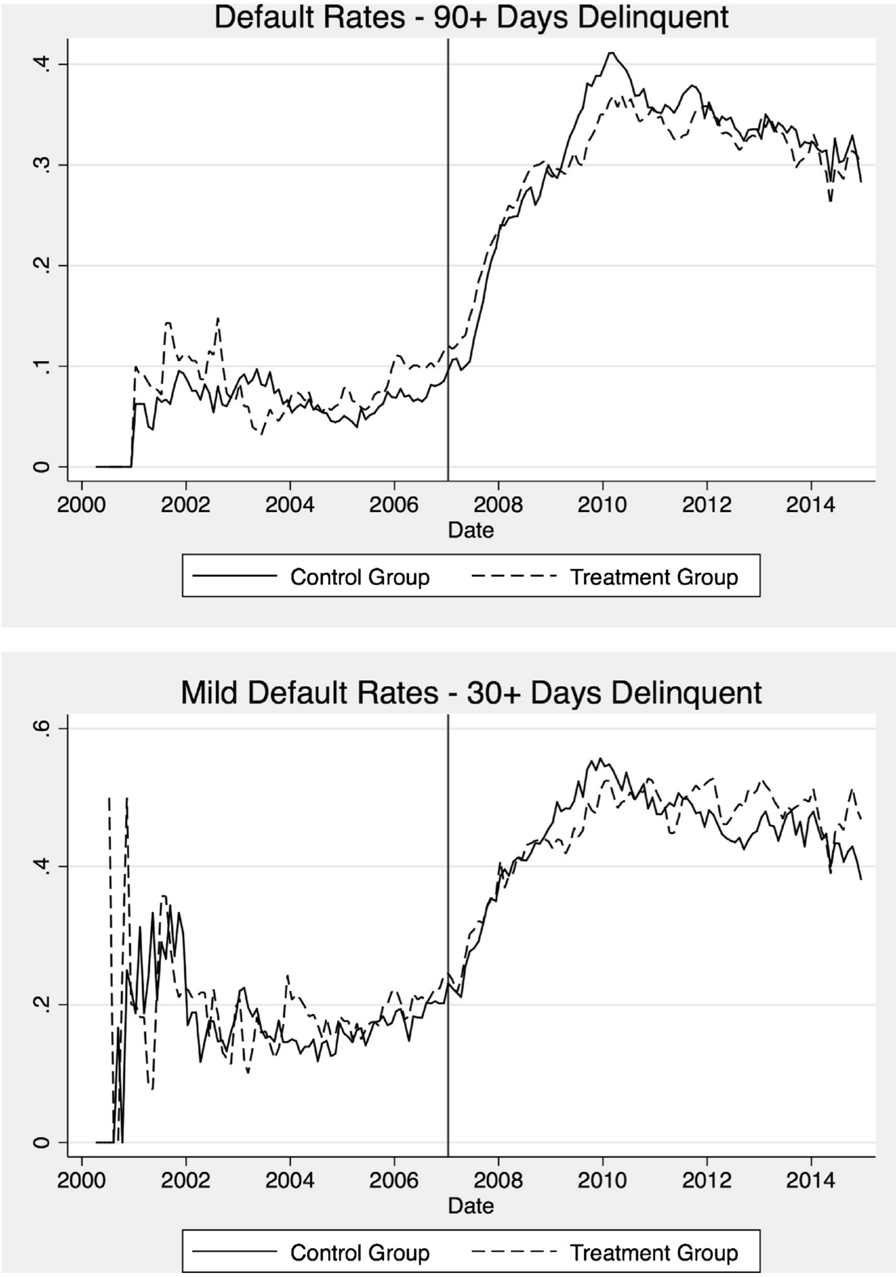


Fig. A3. Haynesville: comparison of shale and non-shale areas mortgage default and mild default.

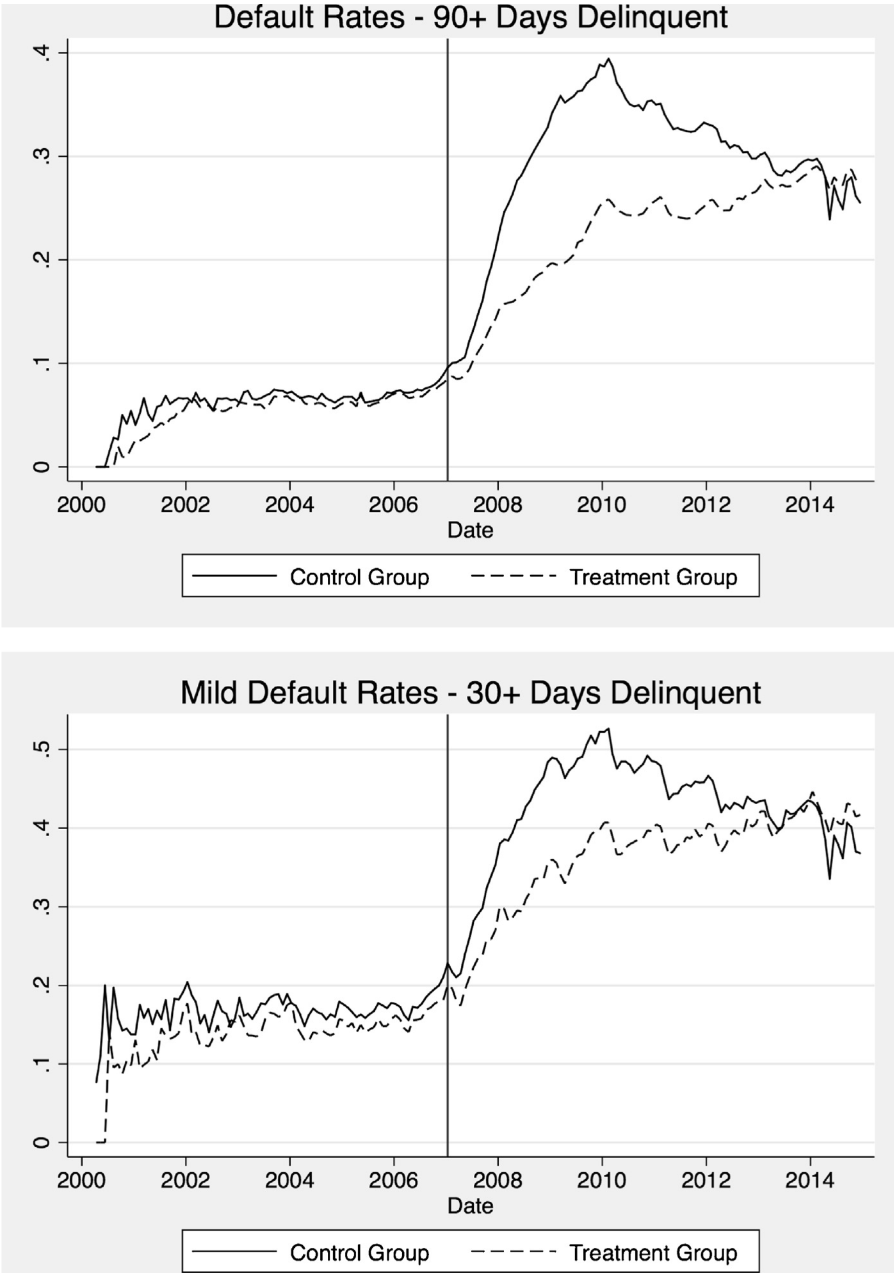


Fig. A4. Marcellus: comparison of shale and non-shale areas mortgage default and mild default.

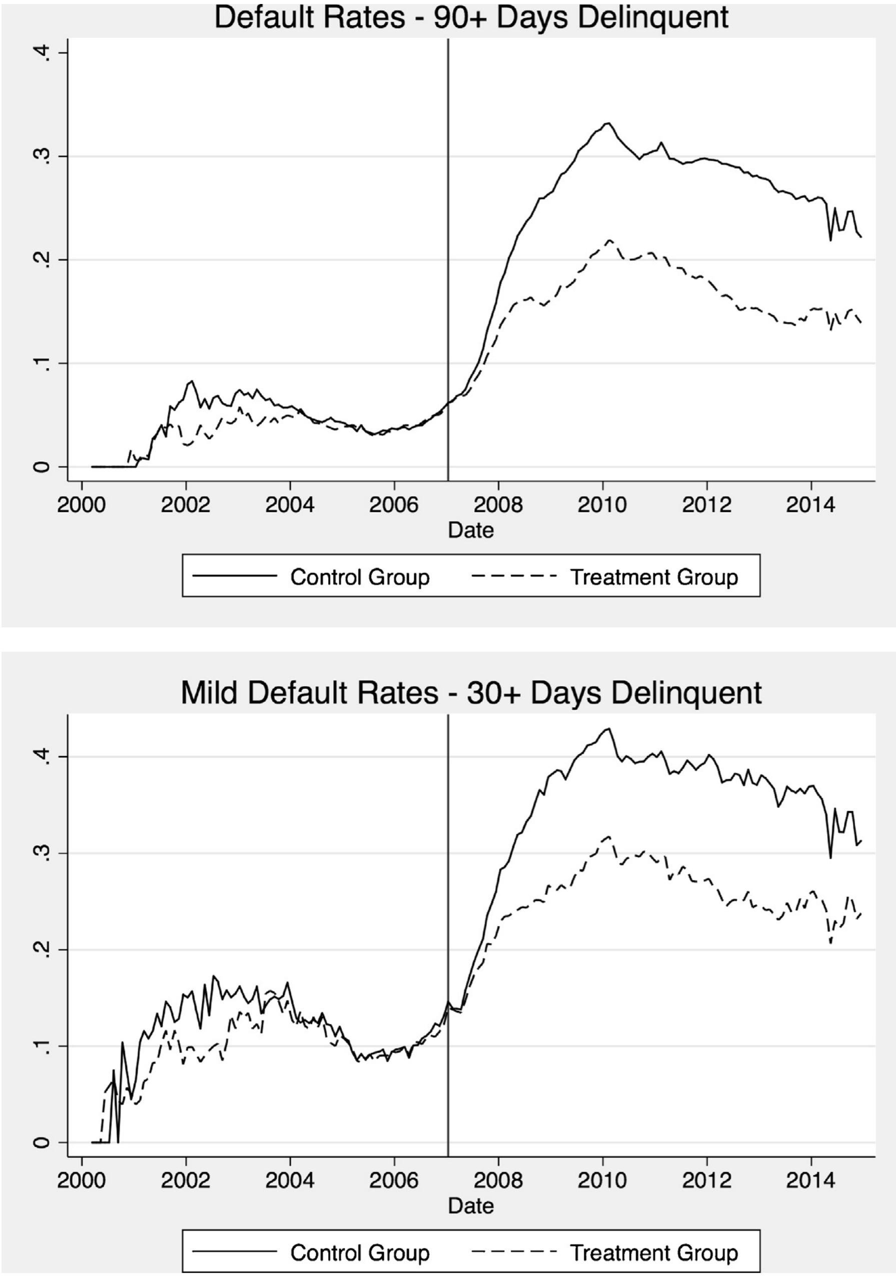


Fig. A5. Niobrara: comparison of shale and non-shale areas mortgage default and mild default.

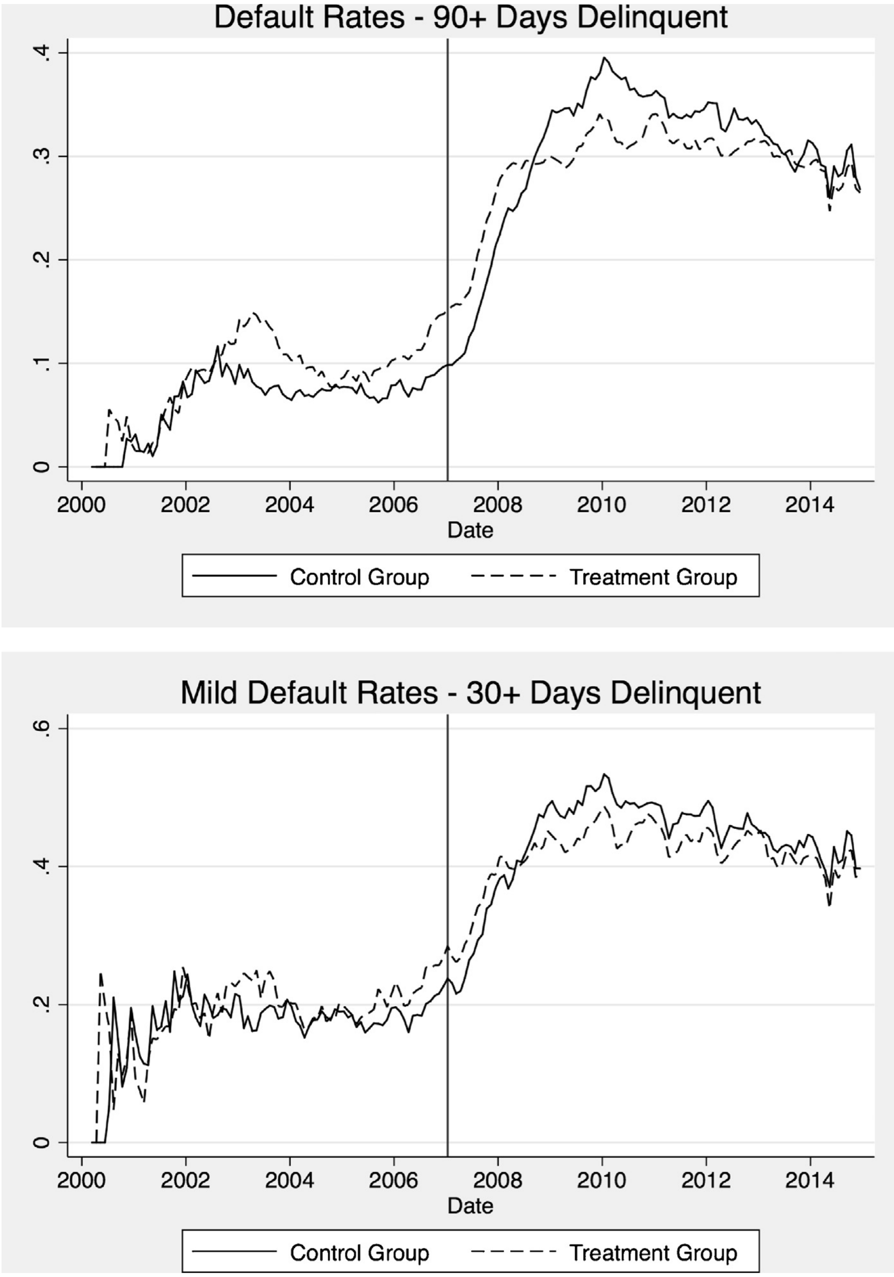


Fig. A6. Utica: comparison of shale and non-shale areas mortgage default and mild default.

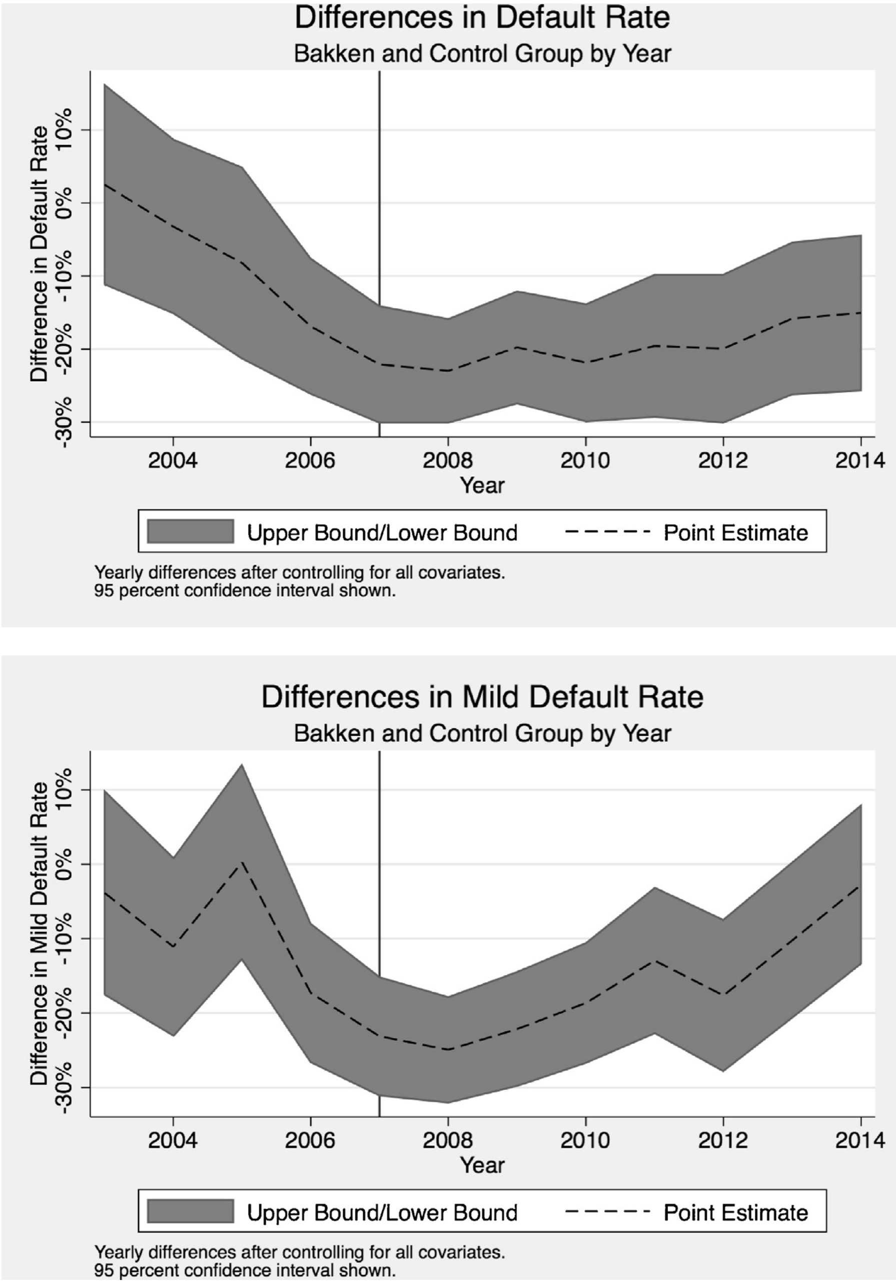
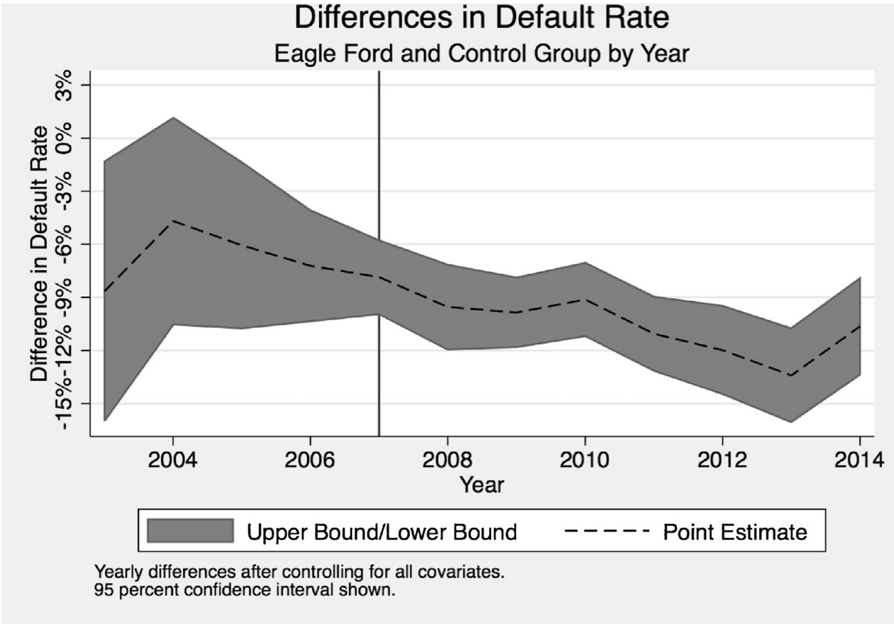


Fig. A7. Bakken: comparison of average difference in default rates of mortgages in shale and non-shale areas by year.



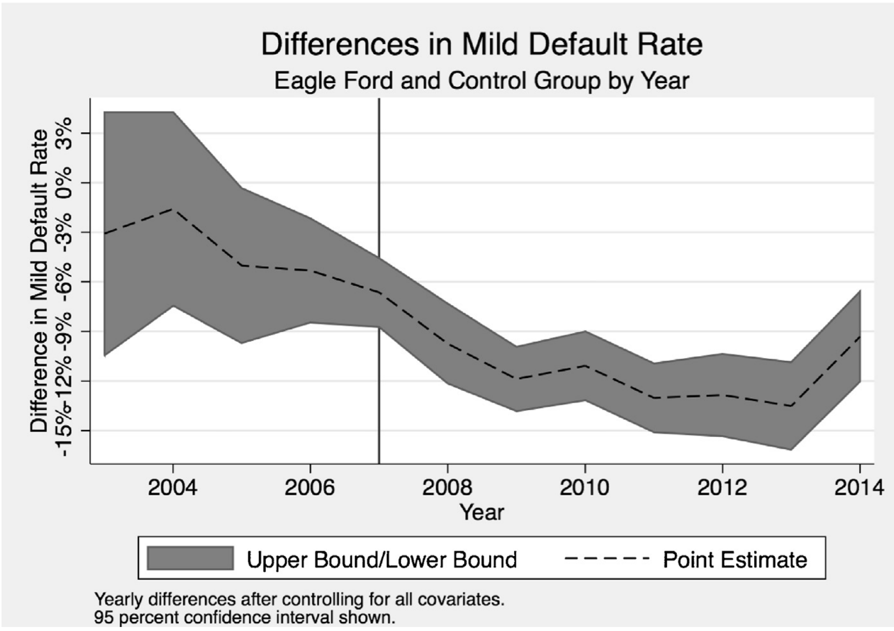
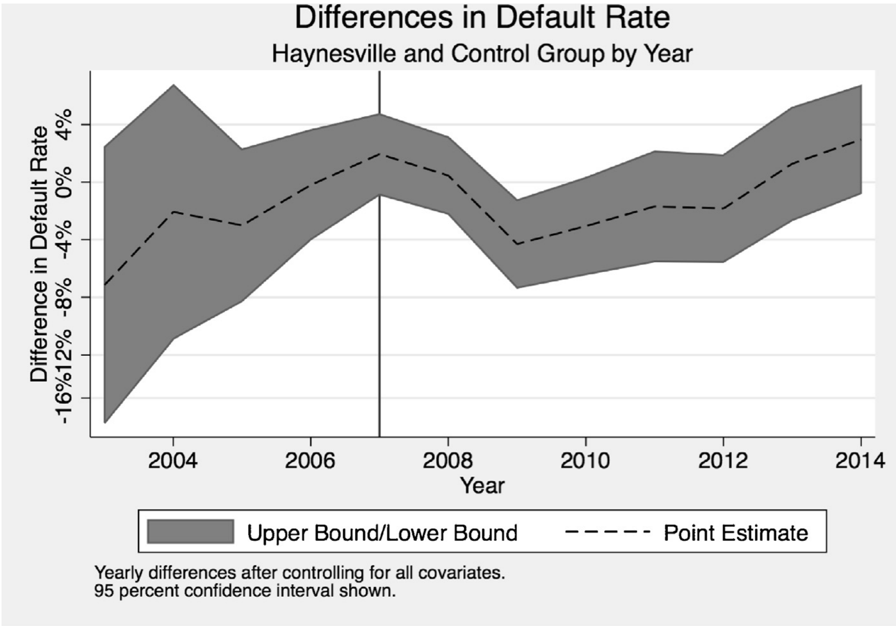


Fig. A8. Eagle Ford: comparison of average difference in default rates of mortgages in shale and non-shale areas by year.



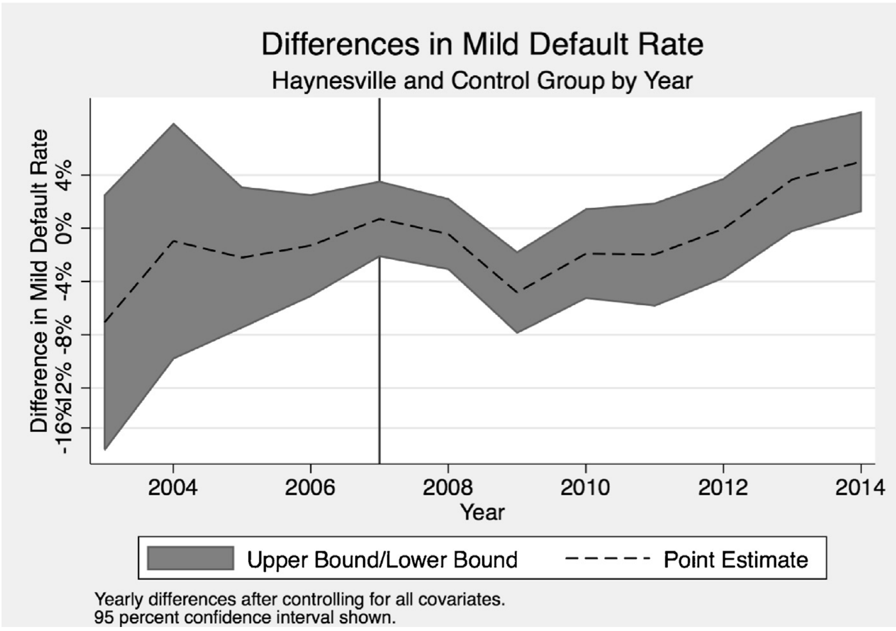
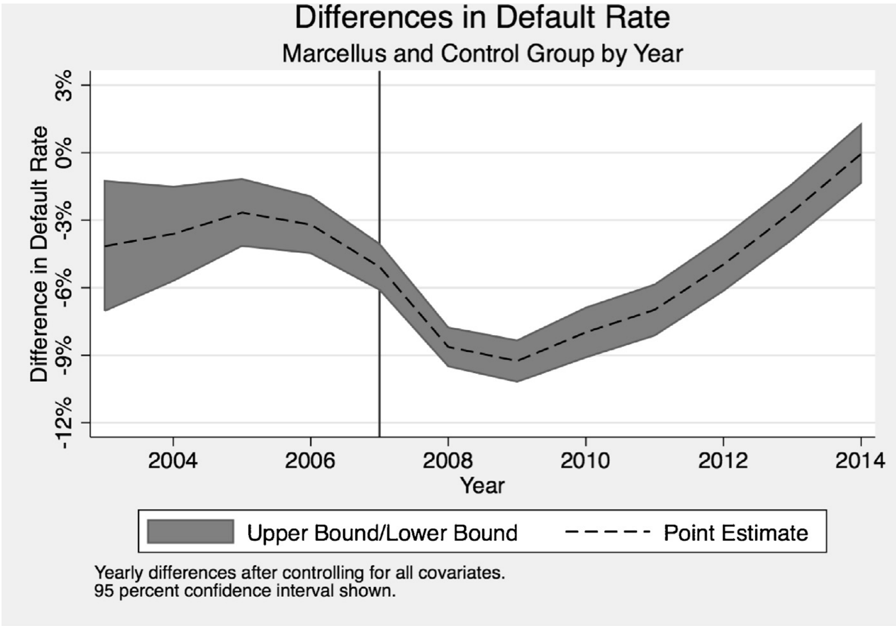


Fig. A9. Haynesville: comparison of average difference in default rates of mortgages in shale and non-shale areas by year.



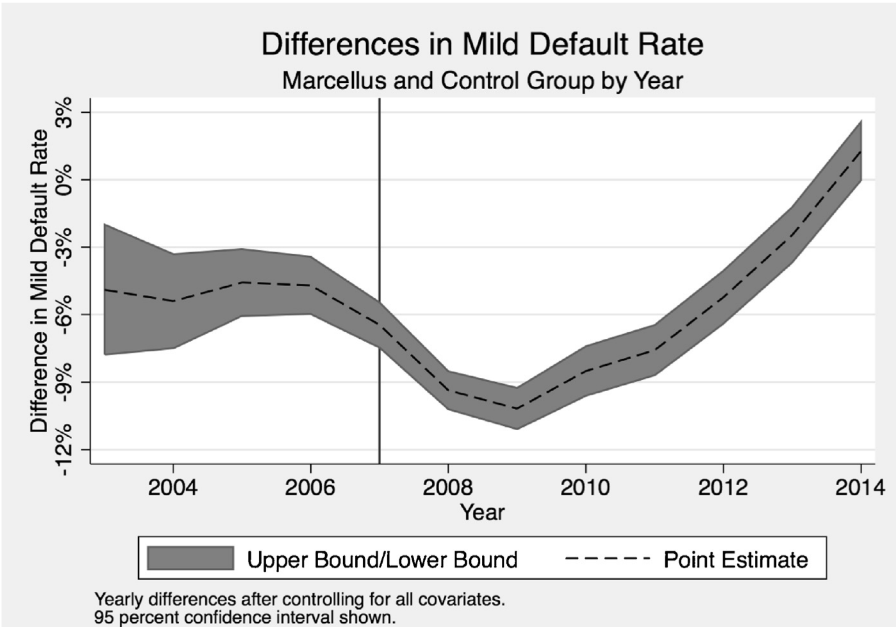


Fig. A10. Marcellus: comparison of average difference in default rates of mortgages in shale and non-shale areas by year.

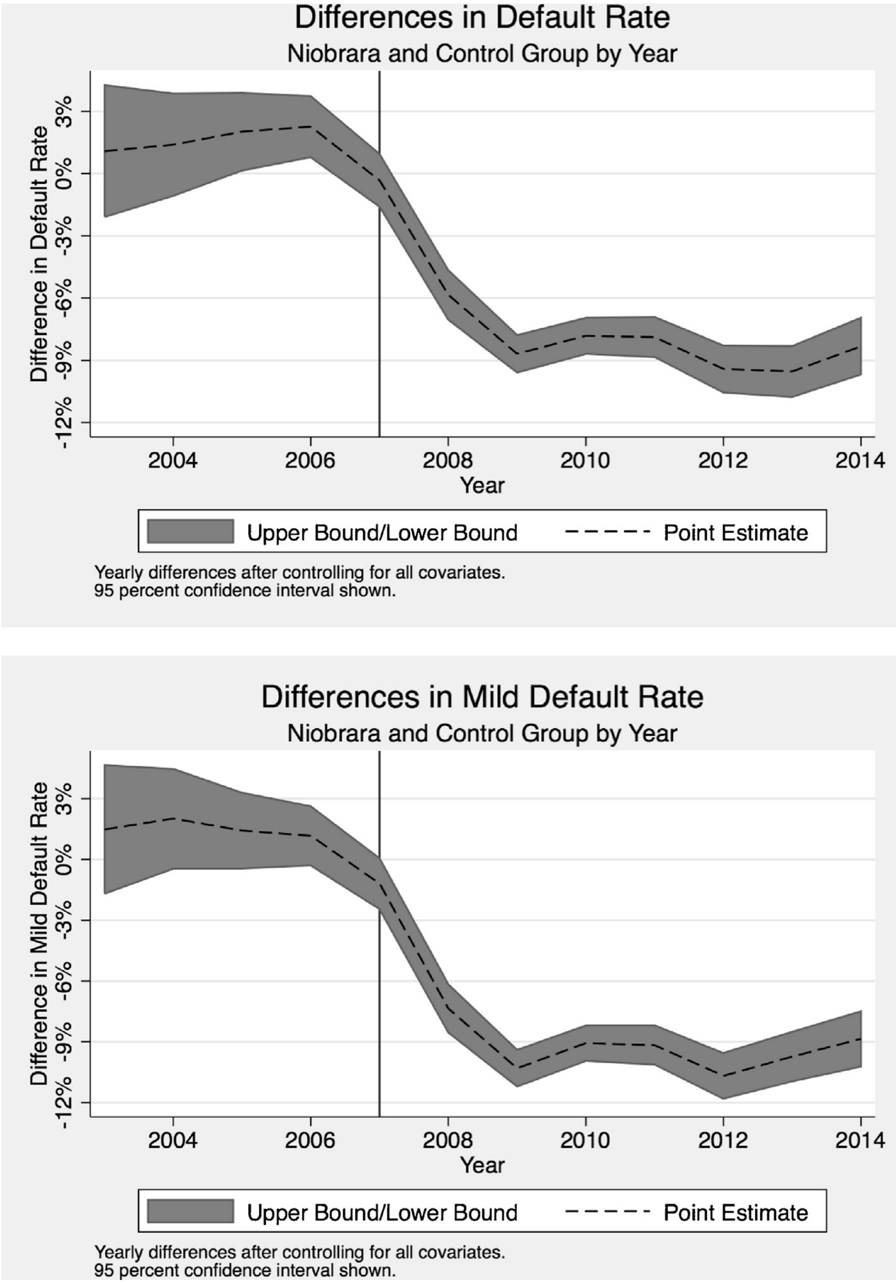


Fig. A11. Niborara: comparison of average difference in default rates of mortgages in shale and non-shale areas by year.

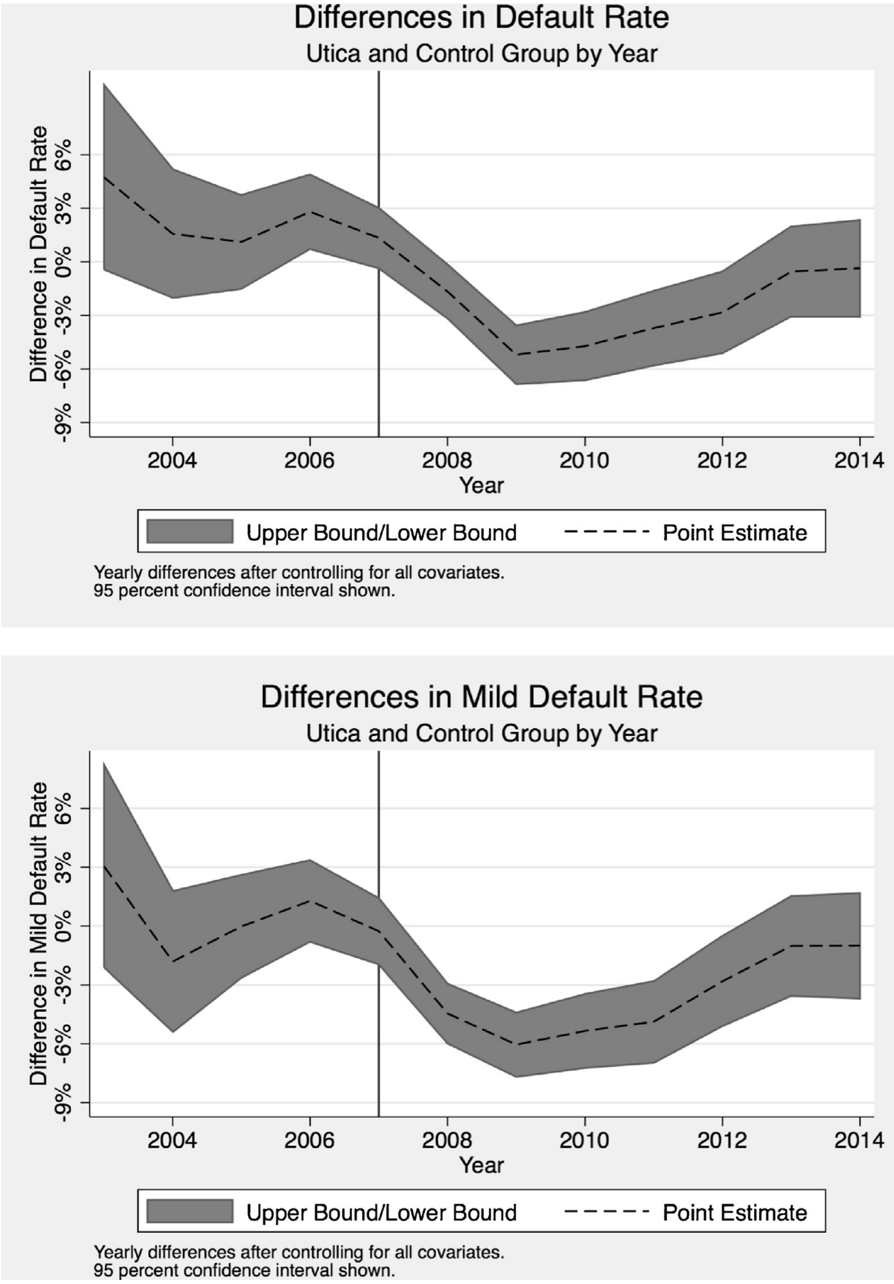


Fig. A12. Utica: comparison of average difference in default rates of mortgages in shale and non-shale areas by year.

References

Acemoglu, D., Finkelstein, A., Notowidigdo, M.J., 2013. Income and health spending: evidence from oil price shocks. Rev. Econ. Stat. 95 (4), 1079–1095.

Agerton, M., Upton, G.B., 2017. Decomposing Crude Price Differentials: Domestic Shipping Constraint or the Crude Oil Export Ban. USAEE Working Paper No. 17-307.

Agerton, M., Hartley, P.R., Medlock, K.B., Temzelides, T., 2017. Employment impacts of upstream oil and gas investment in the United States. Energy Econ. 62, 171–180.

Allcott, H., Keniston, D., 2018. Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America. Rev. Econ. Stud.

Archer, W., Ling, D., McGill, G., 1996. The effect of income and collateral constraints on residential mortgage terminations. Reg. Sci. Urban Econ. 26, 235–261.

Bartik, A.W., Currie, J., Greenstone, M., Knittel, C.R., 2017. The Local Economic and Welfare Consequences of Hydraulic Fracturing. NBER Working Paper 23060.

Black, D., Daniel, K., Sanders, S., 2002. The impact of economic conditions on participation in disability programs: evidence from the coal boom and bust. Am. Econ. Rev. 92 (1), 27–50.

Black, D., McKinnish, T., Sanders, S., 2005. The economic impact of the coal boom and bust. Econ. J. 115 (503), 449–476.

Boxall, P.C., Chan, W.H., McMillan, M.L., 2005. The impact of oil and natural gas facilities on rural residential property values: a spacial hedonic analysis. Resour. Energy Econ. 27, 248–269.
Campbell, T.S., Dietrich, J.K., 1983. The determinants of default on insured conventional residential mortgage loans. J. Finance 38 (5), 1569–1581.
Clapp, J., Goldberg, G., Harding, J., LaCour-Little, M., 2001. Movers and shuckers: interdependent prepayment decisions. Real Estate Econ. 29, 411–450.
Clapp, J., Deng, Y., An, X., 2006. Unobserved heterogeneity in models of competing mortgage termination risk. Real Estate Econ. 34, 243–274.
Davis, M.A., Heathcote, J., 2005. Housing and the business cycle. Int. Econ. Rev. 46 (3).
Decker, R.A., McCollum, M.N., Upton, G.B., 2018. Firm Dynamics and Local Economic Shocks: Evidence from the Shale Oil and Gas Boom. Working Paper.
Demyanyk, Y., Hemert, O.V., 2011 June. Understanding the subprime mortgage crisis. Rev. Financ. Stud., 1848–1880.
EIA, 2014. Total Petroleum and Other Liquids Production. Technical Report. U.S. Energy Information Administration.
EIA, 2017. Drilling Productivity Report: For Key Tight Oil and Shale Gas Regions. Technical Report. U.S. Energy Information Administration.
Elul, R., 2016. Securitization and mortgage default. J. Financ. Serv. Res. 49 (2), 281–309.
Elul, R., Souleles, N.S., Chomsisengphet, S., Glennon, D., Hunt, R., 2010 May. What “triggers” mortgage default? Am. Econ. Rev. 100 (2), 490–494.
Feyrer, J., Mansur, E.T., Sacerdote, B., 2017. Geographic dispersion of economic shocks: evidence from the fracking revolution. Am. Econ. Rev. 107 (4), 1313–1334.
Gilje, E., 2012. Does local access to finance matter? Evidence from U.S. oil and natural gas shale booms. Manag. Sci.
Gilje, E., Loutskina, E., Strahan, P.E., 2016. Exporting liquidity: branch banking and financial integration. J. Finance 71 (3).
Gopalakrishnan, S., Klaiber, H.A., 2014. Is the shale energy boom a bust for nearby residents? Evidence from housing values in Pennsylvania. Am. J. Agric. Econ. 96 (1), 43–66.
Gross, D., Souleles, N., 2002. An empirical analysis of personal bankruptcy and delinquency. Rev. Financ. Stud. 15 (1), 319–347.
Iacoviello, M., 2011. Housing Wealth and Consumption. International Encyclopedia of Housing and Home.
James, A., Smith, B., 2017. There will be blood: crime rates in shale-rich U.S. counties. J. Environ. Econ. Manag. 84, 125–152.
Komarek, T.M., 2016. Labor market dynamics and the unconventional natural gas boom: evidence from the Marcellus region. Resour. Energy Econ. 45, 1–17.
Leamer, E.E., 2008. Housing is the Business Cycle. NBER Working Paper 13428.
Marchand, J., Weber, J., 2017. Local labor markets and natural resources: a synthesis of the literature. J. Econ. Surv.
McCollum, M.N., Lee, H., Pace, R.K., 2015. Deleveraging and mortgage curtailment. J. Bank. Finance 60 (C), 60–75.
Mester, L.J., 1997. What's the point of credit scoring? Bus. Rev.
Mota, N., 2016. Rural Mortgage Lending Over the Last Decade. Technical Report. Fannie Mae.
Muehlenbachs, L., Spiller, E., Timmins, C., 2015. The housing market impacts of shale gas development. Am. Econ. Rev. 105 (12), 3633–3659.
Quercia, R.G., Stegman, M.A., 1992. Residential mortgage default: a review of the literature. J. Hous. Res. 3 (2), 341–379.
Shen, Y., Cunningham, C., Geraldi, K., 2015. Unexpected Wealth Change and Mortgage Default: Evidence from the Shale Gas Boom. Working Paper.
Upton, G.B., Yu, H., 2018. Local Labor Market Shocks and Employment and Earnings Differentials: Evidence from Shale Oil and Gas Booms. USAEE Working Paper.
Vachon, M.C., 2015. The Impact of Local Labor Market Conditions on Migration: Evidence from the Bakken Oil Boom. Working Paper.
von, G.M., Furstenberg, 1969. Default risk on FHA-insured home mortgages as a function of the terms of financing: a quantitative analysis. J. Finance 24, 459–477.
Weber, J.G., 2012. The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming. Energy Econ. 34 (5), 1580–1588.
Yergin, D., 1999. The Prize. The Epic Quest for Oil, Money and Power. Free Press.