FEDERAL RESERVE BANK *of* ATLANTA **WORKING PAPER SERIES**

The Double Trigger for Mortgage Default: Evidence from the Fracking Boom

Chris Cunningham, Kristopher Gerardi, and Lily Shen

Working Paper 2017-4a
March 2017 (Revised October 2020)

Abstract: This paper exploits a natural experiment afforded by the fracking boom in Pennsylvania to shed light on the determinants of mortgage default. Looking only at mortgages originated before fracking became viable and using the underlying geology as a supply shifter, we find that mortgages on homes exposed to shale drilling experience a significant reduction in default risk. This effect is more than four times greater for borrowers who are underwater on their loans. Additional evidence shows that fracking activity does not raise house prices but significantly increases household income through higher royalty payments, wages, and salaries. Furthermore, we find that fracking directly leads to employment increases in the drilling/mining and construction sectors at the county level, and it reduces income from unemployment benefits at the ZIP code level. Finally, in addition to reducing mortgage default risk, we show that fracking lowers credit card delinquencies. These results are most consistent with the "double-trigger" theory of mortgage default, where underwater borrowers subject to an adverse income shock are much more likely to lose their homes to foreclosure.

JEL classification: Q51, R11, G21

Key words: mortgage default, hydraulic fracking, house prices, shale gas

https://doi.org/10.29338/wp2017-4a

For helpful comments and discussions, the authors thank seminar participants at the 2016 Energy and Commodity Finance Conference, 2016 Southern Finance Association Conference, 2016 Financial Management Association Conference, seminar attendees at Clemson University/Finance, University Catholique de Louvain/IRES Macro Lunch, the London School of Economics/SERC, and Maastricht University/Finance. The authors are especially grateful for comments and suggestions from Brent Ambrose, Henry Overman, Christian Hilber, and Piet Eichholtz. The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Chris Cunningham, Federal Reserve Bank of Atlanta, 1000 Peachtree St. NE, Atlanta, GA, 30309, chris.cunningham@atl.frb.org; Kris Gerardi, Federal Reserve Bank of Atlanta, 1000 Peachtree St. NE, Atlanta, GA 30309, kristopher.gerardi@atl.frb.org; or Lily Shen, yannans@clemson.edu.

1 Introduction

Central to the financial crisis that precipitated the Great Recession of the late 2000s was the dramatic rise in mortgage defaults and the decline in the value of credit instruments linked to those mortgages. The rise in defaults during that period has been well-documented in the literature, but the precise determinants of those defaults and the factors that led to millions of borrowers losing their homes to foreclosure remain unclear. Broadly speaking, there are two main theories of mortgage default. The first, which is often referred to as "strategic or "ruthless default, holds that a borrower will choose to stop making payments when the value of her property falls sufficiently below the outstanding balance of her mortgage.[1] Importantly, borrowers who strategically default have the income and/or wealth to continue making mortgage payments, but decide not to do so. An alternative theory, termed the "double-trigger" hypothesis, holds that borrowers generally wish to stay in their home, and will continue to make payments if able to, so that a second shock to the household's income is necessary to generate default, perhaps because of a moral aversion to default. Thus, the theory predicts that in addition to negative equity, default requires the underwater borrower to experience a sufficiently negative liquidity or income shock that prevents her from continuing to pay the mortgage. Due in large part to data limitations, there is surprisingly little empirical evidence on the importance of strategic considerations versus ability-to-pay issues in explaining mortgage default behavior.[2]

In this paper we exploit a natural experiment afforded by the fracking boom in Pennsylvania in the late-2000s to shed light on the importance of these two theories of mortgage default. The fracking boom provides a plausibly exogenous shock to the primary determi-

[1]This need not correspond exactly to the underwater threshold when the value of the property falls just below the balance of the mortgage. As Kau, Keenan, and Kim (1994) originally argued in an option-theoretic model, the possibility of future price recovery means that the default threshold is likely to be somewhat below the point at which equity turns negative.

[2]One notable exception is Gerardi et al. (2017), which documents evidence of both strategic and ability-to-pay motivations driving mortgage default. However, the study is based on a very small sample of mortgages from the Panel Study of Income Dynamics (PSID). Additional studies that have attempted to identify strategic default include Bradley et al. (2015) and Guiso et al. (2013).

nants of mortgage default, including employment and household income (including royalties) and home values. Using detailed, administrative data on mortgages originated in Pennsylvania in the period before the fracking boom began, (2004–2006), and data on both fracking permits and drilling dates, we first test for the impact of fracking activity on mortgage credit risk. We then employ two instrumental variables specifications that interact the thickness and depth of the Marcellus formation with a state-wide count of permits issued or simply a dummy variable for whether fracking was viable. We find that increased shale gas permits and drilling significantly decreased mortgage default rates in Pennsylvania during the 2007–2012 period. Specifically, residing in a ZIP Code with any fracking activity reduces the probability of severe delinquency by 0.29 percentage points, on average, which corresponds to approximately 120% of the average monthly delinquency rate in our sample of Pennsylvania mortgages (0.24%). Results that do not account for drilling endogeneity are only one-third as large, suggesting that governmental restrictions on drilling at the local level, or the decisions taken by energy companies in choosing where to drill may be biasing down the effect of fracking on land markets shown in the existing literature.

These results are quite robust, as the reduction in mortgage default associated with fracking holds for several measures of fracking activity and across various sub-samples of interest. In addition, the findings pass falsification test that creates a placebo fracking boom during the 2001 recession. We also rule out potential selection effects due to the fracking boom generating heterogeneity in mobility and prepayment behavior.

Having established a significant, negative, causal link between fracking and mortgage default, the paper turns to an analysis of the potential underlying mechanisms.

Relying only on the OLS specification, we show that the negative effect of fracking on default is approximately four times greater for borrowers that we estimate to be in negative equity positions, compared to those that have positive or zero home equity. This result is consistent with the double-trigger theory of default, which holds that default is primarily caused by a combination of a house price decline that results in negative equity and a liquidity shock that makes it difficult, if not impossible, to continue making mortgage payments. Specifically, the result suggests that increased fracking activity ameliorates the adverse employment and/or income shocks that increase the likelihood of the liquidity-based trigger.

The paper then turns to a direct analysis of the impact of fracking activity on house prices, household income, and employment. Using our instrumental variables strategy, we document that the fracking boom does not raise house prices at the county-level. If anything, we find evidence of a negative, albeit statistically insignificant, effect of fracking on house prices, which is inconsistent with the idea that fracking is reducing strategic default motives. At the same time, we show that fracking activity significantly increases employment growth in industries that are directly linked to the shale boom, including the drilling/mining and construction sectors. We further show, using the , that fracking activity is associated with higher overall per-capita income levels but lower levels of income from unemployment benefits. Specifically, ZIP Codes with fracking activity have higher per-capita wage/salary income and higher royalty income, but also less unemployment insurance benefits. Our findings strongly suggest that the fracking boom reduced mortgage default in Pennsylvania primarily by raising income levels and/or avoiding employment disruptions that were affecting much of the country and largely validates the double trigger hypothesis.

Finally, we present evidence that the fracking boom lowered the default rates associated with non-mortgage debt, which lends further support to the double-trigger theory. Using a novel dataset of credit bureau accounts matched to administrative mortgage data, we show that increased fracking activity significantly reduces the propensity of borrowers to default on their credit card balances as well as their first and second mortgage balances. This finding is consistent with fracking ameliorating mortgage default by mitigating negative liquidity/income shocks. If fracking instead reduced strategic default through an expectations channel by which borrowers refrain from walking away from their homes in anticipation of future house price increases and/or future royalty income, we would not expect to find any fracking effects on the default rates of non-mortgage debt.

In addition to providing empirical evidence on the underlying sources of mortgage default, this paper makes an important contribution to the policy debate over the costs and benefits of fracking. Since the last quarter of 2007, over 80,000 fracking wells have been permitted and drilled in populated neighborhoods throughout the United States. According to a 2013 Wall Street Journal article, “More than 15.3 million Americans—roughly 1 out of every 20 people living in the U.S.—now live within a mile of a fracking well.”[3] A number of state, local and international governments have recently implemented outright bans on fracking due to concerns about negative environmental and health impacts.[4] The fracking boom has raised especially important concerns among residential mortgage market participants. Mortgage lenders, borrowers, and policymakers are concerned that real or perceived negative effects from fracking could adversely affect property values and increase mortgage defaults. For example, Fannie Mae and Freddie Mac, the two Government Sponsored Enterprises (GSEs) that insure a large fraction of U.S. mortgages have purchase rules that exclude properties close to mineral wellheads. Anecdotal evidence suggests that these restrictions may be binding as some homeowners have been denied access to mortgage credit due to fracking on or near their properties, while other homeowners located near fracking wells are unable to obtain house insurance.[5] In addition, allowing drilling on one’s property without notifying the mortgage holder could constitute a technical default, perhaps precipitating foreclosure.[6] Finally, the GSE rule or other related barriers to mortgage credit could reduce housing demand and property values pushing additional mortgage borrowers into default. Thus, determining the net effect of fracking on mortgage credit risk is an important, policy-

[3]Gold, Russell and Tom McGinty, “Energy Boom Puts Wells in America’s Backyards” The Wall Street Journal, October 25, 2013.

[4]Associated Press. ”Mexican president-elect vows to end use of fracking”, U.S. News and World Report, July 31, 2018; Kaplan, Thomas, “Citing Health Risks, Cuomo Bans Fracking in New York State”, The New York Times, December 17, 2014; Carroll, Rory “Santa Cruz becomes first California county to ban fracking”, Reuters, May 20, 2014; Arenschield, Lauren, “Fracking industry suing over drilling bans”, The Columbus Dispatch, November 21, 2014.

[5]See for example http://marcelluseffect.blogspot.com/2013/08/ny-landowners-denied-homeowners.html and http://grist.org/climate-energy/fracking-boom-could-lead-to-housing-bust/.

[6]https://www.nytimes.com/2011/10/20/us/rush-to-drill-for-gas-creates-mortgage-conflicts.html

relevant question, and this paper provides the first, rigorous empirical evidence on the issue.

A recent literature has developed that uses the housing market to test for the economic impact of fracking. These papers focus on housing transactions and use hedonic methods to estimate how homeowners value development of shale gas near them. The presence or number of shale wells becomes another neighborhood (dis)amenity that can be quantified via a Rosen (1974) derived hedonic model. The evidence thus far has been mixed. Some studies have found negative effects of fracking on home values (Gopalakrishnan and Klaiber, 2014, James and James, 2014) while a few have found positive effects (Weber et al., 2014, Boslett et al., 2016).[7] In a recent influential paper Muehlenbachs et al. (2015) find that proximity to fracking wells lowers the sale prices of homes with well water, but modestly increases the prices of homes with piped water. Focusing on default complements existing hedonic methods, especially since the fracking boom occurred in the middle of the Great Recession when house prices were declining and homeowners may have abandoned homes near wells rather than selling below their purchase prices.

The paper is organized as follows. Section 2 provides a brief overview of the recent history of hydraulic fracturing in Pennsylvania.[8] Section 3 presents the econometric framework. Section 4 discusses the fracking and mortgage data used in the analysis. Section 5 presents our baseline results on the causal effect of fracking activity on mortgage default. Section 6 summarises several extensive robustness checks. Section 7 investigates the causal mechanisms that may link fracking activity with lower mortgage default. Section 8 provides a brief conclusion.

[7] Gilje (2019) also finds evidence that increases in household wealth from fracking royalty and lease payments positively affects business formation, especially in industries more reliant on external finance.

[8] In the Internet Appendix we provide a more detailed discussion of the fracking process as well as some of its environmental risks and economic benefits.

2 Background on Fracking in Pennsylvania

While the boom in shale gas extraction occurred in many areas of the country, Pennsylvania serves as an especially nice laboratory for exploring the effects of fracking on mortgage markets. Unlike many of the other fracking booms that occurred in largely rural areas, the boom in Pensylvania took place in or near urban centers. In our sample, which excludes Philadelphia and Pittsburgh, 80 percent of loans are within a county that is in a designated Metropolitan Statistical Area (MSA), and approximately 11 percent of these homes have a permitted fracking well in their ZIP Code during the sample period. In addition, Pennsylvania's economy was not initially focused on the oil and gas industries. Conventional gas and oil drilling were only viable in a small and largely distinct area of the state in the early 2000s.[9] More importantly, the fracking boom, which was driven by technological improvements in shale extraction techniques, was largely unanticipated by borrowers or lenders in Pennsylvania in the mid-2000s.

In addition, Pennsylvania's geological formations are well mapped because of its history of coal extraction and nineteenth century oil boom, so there is little uncertainty about where the shale is located.[10] Finally, and uniquely to Pennsylvania, an earlier state court decision, *Dunham and Shortt v. Kirkpatrick, 1882*, stipulated that any sales of "mineral rights" does not include oil and gas. The ruling, now referred to as "Dunhams Rule" and re-affirmed in *Butler v. Charles Powers Estate, 2013*, means that, unlike in other fracking states, there is a high probability that current homeowners still retain ownership of the natural gas beneath their land.

The Marcellus shale boom arose because of rapid technological advances in the fracking process. While fracking had long been used to extend the life of conventional oil wells, it was only in the early 2000s that it was first applied to shale rock. The technological innovation

[9] Approximately, 15 percent of our sample had some pre-existing exposure to oil and gas drilling and our results are robust to excluding these properties from the analysis as we will show in Table A.2 of the Internet Appendix.

[10] This also mitigates the concern that the precision of the maps might be endogenous to the preference for drilling on the surface.

at the heart of the shale fracking boom in Pennsylvania was horizontal drilling.[11] Horizontal drilling, in turn, arose from the rapid improvements in computer-based 3D modeling of geological formations and engineering advances in drill bit navigation that allowed the drills to reach those formations.

The emergence of horizontal drilling techniques made the Marcellus Formation newly viable as a source for oil and natural gas in the late 2000s. The first exploratory well in the Marcellus Formation was drilled in 2002, but economically valuable amounts of gas were not recovered until 2007. By 2010 just under 1,400 wells were drilled. While the Marcellus Formation is currently ranked as the largest proven wet gas reserve in the U.S. it was not within the top 100 gas plays as recently as 2008 (Agency, 2010, Agency, 2015). The resulting shale gas boom offers a compelling natural experiment. A known resource went from being effectively worthless to quite valuable in a very short period of time and generated an extraordinary amount of economic activity in the middle of worst national recession since the great depression.

3 Econometric Framework

Investigating the impact of fracking on mortgage performance is challenging because in late 2007 the Pennsylvania mortgage market was subject to both the emergence of fracking as well as the bursting of the housing bubble, which led to the subprime mortgage foreclosure crisis and subsequent financial crisis and recession.[12] Both of these events likely had an impact on the credit risk of outstanding mortgages as well as the credit risk associated with new loans originated in the post-2007 period.[13] For this reason we choose to focus on

[11] Horizontal drilling made extraction in the Marcellus Formation feasible for three primary reasons. First, it allowed a single vertical shaft to turn and follow the relatively thin shale formation over greater distances. Second, it allowed a single well pad to access a larger area, reducing the cost of ground leases and site preparation. Third, because of the asymmetric loads on the shale rock, it tends to fracture vertically so a horizontal well can intersect multiple fissures.

[12] We limit the analysis to Pennsylvania, in part, to avoid introducing state-level variation in mortgage market regulations that can affect the timing of default as documented in Pence (2006).

[13] The literature has documented how the financial crisis caused a significant tightening of mortgage underwriting standards. For example, the Urban Institute has estimated that tight lending standards resulted in

mortgages originated (and underwritten) before 2007 so that our analysis is not contaminated by selection effects related to the emergence of fracking or the decline of the housing market. Thus, our empirical analysis is focused on quantifying the impact of fracking activity on the default risk of *outstanding* loans.

We employ a hazard framework in most of our empirical analysis, where we relate the monthly hazard of first serious delinquency at the individual loan level to fracking activity in the ZIP Code. Our baseline specification is a linear probability model, where the unit of analysis is a loan-month. The dependent variable is zero until the mortgage becomes 90-days delinquent at which point it takes a value of one. A loan enters the sample in the month of origination and exits the sample after the first month it becomes 90 days delinquent, is prepaid voluntarily, or is right-censored at the end of 2012. We do not jointly model prepayment and default in the baseline model, but we do show that the results are robust to doing so in the Internet Appendix.[14] We use a measure of delinquency instead of foreclosure in order to isolate a decision margin that is under the purview of the borrower. The decision to foreclose is made by the mortgage servicer, which may choose to delay initiating foreclosure proceedings for a number of reasons (Piskorski et al., 2010). Our benchmark specification is given by the following equation:

$$Prob(Delinq_{it} = 1) = \alpha + \theta frack_{zt} + dur'_{it}\beta_1 + X'_{it}\beta_2 + \eta_c + \delta_t + \varepsilon_{it}, \tag{1}$$

where i indexes the individual mortgage, z indexes the ZIP Code in which the loan is originated, t indexes the year-month (in calendar time), and c corresponds to the county in which the underlying property is located. The term dur_{it} measures the number of months since the mortgage was originated and enters as a second-order polynomial. X_{it} is a vector of mortgage-level control variables, which we describe in detail below. The variable η_c corresponds to a full set of county fixed effects and δ_t is a full set of year-month fixed effects.

approximately 4 million fewer mortgages in the 2009–2013 period (http://www.urban.org/urban-wire/four-million-mortgage-loans-missing-2009-2013-due-tight-credit-standards).

[14]See Section A.2.3 and Table A.6 in the Internet Appendix.

The term $frack_{zt}$ refers to a measure of fracking activity in ZIP Code z in month t. We are focused on determining the sign and magnitude of the coefficient θ. As fracking could potentially raise or lower the propensity to default, our null hypothesis is that it has no effect: $H_0 : \theta = 0$.

We consider multiple measures of fracking activity and also instrument for fracking activity to address potential endogeneity concerns, which we discuss in more detail below.[15] Our covariate set, X_{it}, includes detailed mortgage and borrower characteristics at the time of origination, which are typically used by underwriters. These include the balance of the loan at origination, the mortgage interest rate, loan-to-value (LTV) ratio (at origination), the ratio of the borrower's monthly payment to monthly income (debt-to-income ratio), and the borrower's FICO score. We also include dummy variables that indicate whether a loan has a fixed or adjustable interest rate, is a refinance or purchase loan, is a jumbo or conforming loan, has a 30-year term or a different maturity length, has income and/or assets that are less than fully documented, has private mortgage insurance, has a LTV ratio of exactly 80%[16], contains a prepayment penalty, has an interest-only payment, and whether the loan is characterized by a balloon payment at the end of its term. We also control for whether the mortgage was ultimately retained in the lender's portfolio, pooled into a private-label, mortgage-backed security (MBS), or packaged into a GSE (Fannie Mae or Freddie Mac) security with a credit guarantee, which is the omitted category in all regressions.

3.1 Identification

The decision to sell mineral rights and/or sell or lease surface land by landowners and the decision by local government to permit drilling may be endogenous. For example, a struggling community may be more willing to accept drilling in the hope of attracting additional

[15] To address potential spatial and serial correlation, we employ two-way clustering of the standard errors by county and year. Clustering by county accounts for possible spatial correlation within counties over time, while clustering by year accounts for potential correlation across counties within a calendar year.

[16] Mortgages with a LTV ratio of exactly 80% often had subordinate liens (piggy-back loans) in the pre-crisis period. Since the McDash dataset does not contain any information on subordinate liens, we use this variable as a proxy.

employment or tax revenue, whereas a wealthier community may put more weight on health or environmental concerns or current home prices and choose to block drilling. As a result, looking across communities we could observe that areas with extensive fracking activity also have higher rates of serious mortgage delinquency without one causing the other. Blohm et al. (2012) find that 32 percent of the Marcellus Formation is inaccessible because of regulation or current land use.[17] Even if local governments fail to act, engaged citizens may motivate regional, state, or federal agencies to restrict drilling. In addition, house price declines could themselves encourage fracking activity by making it cheaper for firms to acquire land for drilling, storage and pipeline easements.

While the fracking technology shock is exogenous, the location of the wells themselves may not be. To address this endogeneity issue, we employ an instrumental variables strategy using the geologic properties of the underlying Marcellus Formation. According to Wrightstone (2009) there are a number of geologic factors that determine the productivity of a fracking well. We focus on two factors: thickness and depth.[18] All else equal, greater shale thickness and greater shale depth leads to higher well production. Figures 1 and 2 display contour maps of the thickness and depth of the Marcellus Formation with the location of wells superimposed on top. The strong correlation between the location of wells and shale thickness and depth is apparent from the maps.

In addition to these geologic factors, which are time-invariant, we also incorporate the timing of the initial increase in drilling activity into our instrument set. As we discussed above, fracking did not become a viable means of extracting oil and gas from the Marcellus Formation until the late 2000s with the emergence of horizontal drilling techniques. Therefore, we interact the geologic determinants of fracking productivity in our instrument set with the total number of annual horizontal fracking well permits issued in the state, which

[17] By 2012, seven local governments in Pennsylvania had banned fracking and six had imposed restrictions on where fracking could occur (Blohm et al., 2012, Table A2). This authority was stripped from local governments by the state in 2012.

[18] In addition to thickness and depth, Wrightstone (2009) lists the following factors as important productivity determinants: maturity, gas content, areal extent, structural complexity, lateral continuity, pressure gradient, and natural fracking.

captures the aggregate growth of the industry. As a robustness check, we also use a second IV approach that interacts the geologic variables with a simple dummy variable that takes the value of one in the period after horizontal drilling first became viable in Pennsylvania (post-2008). In our preferred specification, we include the geologic variables by themselves in our set of exogenous covariates so that we can control for any time-invariant, unobservable factors that may be correlated with the geologic properties of the Marcellus Formation. This would account for any unobservable factor that generates higher mortgage default in ZIP Codes with greater shale thickness and depth in the pre-fracking period. Thus, our instrument set is given by:

$$Frack_{zt} = \{Thickness_z \times Horizontal_t \ , \ Depth_z \times Horizontal_t\} \tag{2}$$

where $Horizontal_t$ is either the total number of annual horizontal fracking well permits issued in Pennsylvania, or a dummy variable corresponding to the period in which horizontal drilling first became viable (post-2008).[19]

4 Data

In this section we discuss the two primary sources of data used in the analysis. We begin with a brief description of our fracking data followed by a discussion of our mortgage data. As mentioned above, we focus on loans originated in the 2004–2006 period in order to avoid potential selection bias stemming from changes in underwriting standards that took place due to the onset of the financial crisis and the fracking boom that both began in the late-2000s. In addition, we focus on mortgage performance through 2012 in order to isolate the effects of the fracking boom. The impact on mortgage and housing markets from the

[19] We also experimented with quadratic expressions for thickness and depth, as well as non-linear effects around minimum thickness and a specification that included a triple interaction variable, $Thickness_z \times Depth_z \times Horizontal_t$ to capture the possibility that areas with both greater shale thickness and depth are especially attractive. However, these more sophisticated specifications were no more predictive of fracking activity.

subsequent slowdown in fracking activity in response to the recent large global oil and gas price decline is an interesting topic in its own right, but we will not include it in this paper.

4.1 Fracking Data

The data pertaining to gas exploration activity were collected from the Pennsylvania Department of Environmental Protection (PA DEP), which provides monthly reports on permitting, drilling, and compliance activities. The PA DEP monthly report provides the universe of permits for both conventional and fracking wells in Pennsylvania from 1975 to 2017. The information provided in the dataset includes the unique well identification code (API), the exact latitude and longitude of each well, the type of well (conventional versus fracking), and the issuance date of the drilling permit. We use the information on latitude and longitude to compute the ZIP Code in which each well is located (via the Geographic Information System software ArcGIS). We then match the data on permits and wells to the mortgage data described below at the ZIP Code-month aggregation level.

Figure 3 displays annual counts of fracking permits and drilled wells in Pennsylvania from 2001–2017. Panel A shows counts for horizontal wells, while Panel B displays counts for vertical and directional wells. The first observation to note from the figure is that horizontal drilling has been much more popular than vertical drilling over the sample period. Another important observation to note is the different time-series dynamics for the two types of fracking wells. Vertical drilling first began on a very small scale in the early 2000s, ramped up a bit in the 2007–2009 period, and then significantly dissipated thereafter. In contrast, horizontal drilling emerged later, scaled up very quickly to well counts that were almost an order of magnitude larger than vertical and directional drilling, and persisted at very high levels through the end of the sample period. These observations are consistent with our contention in Section 2 that the horizontal drilling technique was the major innovation that made shale gas extraction in the Marcellus Formation economically viable.

A final notable observation from Figure 3 is the difference in the levels of permits and drilled wells. In a given year, permit levels are always significantly larger than the number of drilled wells, which reflects two facts. First, there is typically a lag between the time that the permit is approved and the time that the well is actually drilled. Second, not all permits evolve into successfully drilled wells.[20]

Table 1 shows the status of all fracking permits issued in the state as of the end of our data (January 2017). In total, the PA DEP granted 18,549 fracking permits over the course of this period. Approximately 53 percent of the permits were still active at the end of the period, meaning that wells had been drilled and were either producing or expected to produce in the near future. About 8 percent of permits issued had turned into wells, but by the end of the period had been either abandoned, plugged, or temporarily shut down. Finally, almost 39 percent of permits expired prior to the commencement of drilling.[21]

We focus on permits in the bulk of our analysis. However, since a significant fraction of fracking permits expire without drilling ever taking place, and the permits that do not expire often take several months to become active wells, measures of fracking activity based on permits and actual drilled wells could diverge significantly. Thus, in robustness tests, we also consider measures of fracking activity based on drilled wells.

Table 2 displays a set of basic sample summary statistics for the various fracking measures that we use in our empirical analysis. The first four variables correspond to measures of permits. "Any Fracking" is a dummy variable that takes a value of zero until the first fracking permit is issued in a ZIP Code, and then one for the remainder of the sample period. We will focus on this variable in most of our empirical analysis. "Active Permits" measures the number of permits in a ZIP Code that are active in a given month t. "Cumulative Permits Issued" measures the total number of permits (both active and expired) that have ever been issued in a ZIP Code through the current month t. "Newly Permitted Wells" measures the

[20]Figure A.1 in the Internet Appendix displays the distribution of the number of months between permitting and drilling for the sample of wells that are eventually drilled in our data. The figure shows that the lag is typically quite short. About two-thirds of wells are drilled within three months of the permit being approved.

[21]A fracking permit expires in 12 months in PA. However, those expired permits can be easily renewed as long as the underlying lease is still active.

number of new permits issued in each month. The final two variables correspond to measures of actual drilling activity. “Cumulative Drilled Wells” is the total number of wells that have ever been drilled in a ZIP Code through the current month t, and “Newly Drilled Wells” is the number of new wells drilled in a ZIP Code each month.

4.2 Mortgage Data

The mortgage performance data used in the analysis were obtained from McDash Analytics. The McDash dataset covers between 60 and 80 percent of the U.S. mortgage market (depending on the year), and contains detailed information on the characteristics and performance of both purchase mortgages and refinance mortgages. It includes mortgages from all segments of the U.S. mortgage market: non-agency, privately securitized loans (PLS); loans purchased and securitized by the GSEs; and loans held in lenders' portfolios. The McDash dataset is constructed using information from mortgage servicers, financial institutions that are responsible for collecting mortgage payments from borrowers. Each loan is tracked at a monthly frequency from the month of origination until it is either paid off voluntarily or involuntarily via the foreclosure process. The monthly performance data include detailed information about the mortgage status, including the number of payments that the borrower is behind, the month in which the servicer begins foreclosure proceedings, and the date of foreclosure completion. We follow the convention in the literature and define borrowers who are at least 90 days behind on their mortgage payments as being in default. We focus on a delinquency measure, which is under the borrower's control, rather than a foreclosure measure, which is under the servicer's purview, in order to mitigate any bias that might come from the dramatic changes in servicer incentives and state-level foreclosure timelines that took place during the financial crisis and post-crisis periods.[22]

Our primary focus is on a sample of loans originated in the mid-2000s for single-family

[22]A number of influential papers have worked with administrative loan performance data either from a single servicer as in Jiang et al. (2014) or, like us, have relied on commercial datasets of loan performance Anderson and Dokko (2016).

housing before the Pennsylvania fracking boom. Specifically, we consider mortgages originated between 2004 and 2006 (inclusive). The finest geographic information contained in the McDash dataset is the ZIP Code corresponding to each mortgaged property.[23] We further restrict our sample to loans that are not located in the Philadelphia or Pittsburgh metropolitan areas. Many fracking wells in the state are located in rural neighborhoods, with no wells located in the southeastern part of Pennsylvania where Philadelphia is located, and very few wells located in the area immediately surrounding Pittsburgh. In some of our robustness specifications in the Internet Appendix we also limit the sample to ZIP Codes that have not experienced conventional oil or gas drilling, as well as only ZIP Codes on the Marcellus Formation.

Table 3 presents summary statistics for many of the mortgage variables used in the analysis. Column (1) presents means and standard deviations for our baseline specification that includes all McDash mortgages on properties in Pennsylvania but excluding counties in the Philadelphia and Pittsburgh MSAs. Note that 12.6 percent of these mortgages would ultimately become 90+ days delinquent. In columns (2) and (3) we stratify the sample by ZIP Codes that never had fracking activity ("never fracked" ZIP Codes) and those that did have fracking activity at some point during our sample period ("ever fracked" ZIP Codes). First, we note that less than 10 percent of all mortgages were exposed to fracking within their ZIP Codes, highlighting the rural nature of the fracking industry. Generally, when we compare "ever fracked" and "never fracked" mortgage characteristics we observe that loans, which would eventually be subject to fracking, had, on average, slightly worse credit scores, slightly higher interest rates, were slightly more leveraged at origination, and were more likely to have refinanced their previous mortgage. While these difference are not particularly large, with the exception of mortgage amount, it is certainly consistent with places or people with more vulnerable mortgages being more amenable to fracking.[24]

[23]We use a ZIP Code to county crosswalk provided by the Census Bureau when we include in our analysis variables measured at the county-level, such as house prices and employment flows.

[24]However, we should note that borrowers in fracking ZIP Codes were somewhat less likely to use low-documentation loans or adjustable-rate mortgages.

5 Baseline Results

In Table 4 we present a subset of coefficient estimates for our baseline hazard model of default, in which we regress our outcome variable, an indicator for the first month that a mortgage becomes 90 days delinquent, on our fracking measure. We focus our analysis on the "Any Fracking" dummy indicator, which is zero until the month that the first fracking permit is granted in the ZIP Code and then takes a value of one for the rest of the sample period.[25] In the table, we gradually add control variables and, in the last two specifications, instrument for fracking activity.

In column (1) we present the estimate of θ, when we control only for the duration of the mortgage which we specify as a quadratic polynomial. The coefficient estimate, $\hat{\theta}$, is -0.026 and is not statistically significant. In column (2) we include the full set of underwriting variables associated with mortgage i, described in detail in Section 3. We also include dummy variables for the year of origination to control for potential changes in underwriting standards over time. The coefficient estimates associated with the underwriting variables are largely consistent with what previous studies in the default literature have found. For example, borrowers with worse credit scores are more likely to default, and loans with low documentation as well as mortgages that end up in private-label securities are more likely to default. Controlling for mortgage characteristics significantly increases the coefficient estimate of fracking (in absolute value) to -0.079, and is different from zero at the 5 percent level, which suggests that places with observably riskier loans were more likely to allow fracking activity.

In column (3) we include calendar year fixed effects and in column (4) we include county fixed effects. The inclusion of these fixed effects does not appreciably change the fracking coefficient estimate. In column (5) we substitute the calendar year fixed effects with a full set of year-month dummies to absorb all inter-temporal variation, which does not materially affect $\hat{\theta}$. In column (6) we substitute ZIP Code fixed effects for county effects, so that the

[25] In the Internet Appendix (Section A.2.1) we discuss results for the alternative fracking measures.

effect of fracking activity on default is estimated using only time-series variation within ZIP Codes. The inclusion of ZIP Code fixed effects slightly increases the (absolute) magnitude of $\hat{\theta}$ to -0.093 and is again consistent with an interpretation that communities with riskier mortgages were also more likely to allow fracking.

To motivate the instrumental variable regressions, we first show estimates from a reduced-form regression of mortgage default on ZIP Code shale depth and thickness interacted with year dummies. The coefficient estimates, which are displayed in Figure 4, show how the correlation between mortgage default and the two geologic variables evolves over time during our sample period. The depth coefficients (top panel) are actually positive in the pre-fracking period before 2008, and then turn significantly negative after fracking becomes viable in 2009. The thickness coefficients show a similar, but slightly less pronounced pattern, as most are not statistically different from zero. In general, the patterns in Figure 4 suggest that the probability of default was higher in ZIP Codes characterized by high shale thickness and depth before the fracking boom commenced, and then declined during the fracking boom in the post-2008 period.

In columns (7) and (8) of Table 4 we display the results of our instrumental variables specifications. In column (7) we instrument for the "Any Fracking" dummy using the geologic variables shale depth and thickness interacted with the total number of annual fracking permits granted in the entire Pennsylvania Marcellus Formation. As we can clearly see from Figure 3, the running count of the aggregate number of permits increases dramatically in the 2008–2011 period and captures the adoption of the technological innovation of horizontal drilling in Pennsylvania. In addition, the aggregate number of permits across the entire state should be exogenous with respect to drilling activity in a particular ZIP Code and year-month.[26] In column (8) we instrument for fracking using the geologic variables shale depth and thickness interacted with a simple post-2008 indicator variable, $Post2008$.[27]

[26] We thank an anonymous referee for suggesting this identification strategy.

[27] We also tried an alternative specification where we assume an even more conservative cutoff choice of 2006, as there were effectively zero horizontal permits and drilling activity before 2007. The results are not sensitive to this alternative cutoff choice.

For the IV specifications, we revert to county fixed effects, as we are unable to obtain convergence in the estimator with ZIP Code fixed effects. Instrumenting for fracking increases the estimated magnitude of θ by a factor of 4, from -0.074 (column (5)) to -0.289 (column (7)). Both IV estimates are statistically significant at conventional levels and are quite similar in magnitude. The difference in magnitudes between the IV and LPM estimates is consistent with endogeneity bias stemming from selection on the location of fracking wells.[28]

In Table 5 we display the first stage IV results. The coefficients associated with the instruments are all positive and statistically significant, implying that fracking activity is greater in ZIP Codes with higher shale thickness and depth in the post-2008 period. The Kleibergen-Paap Wald F statistics for both specifications are both above the typical threshold of 10 for weak instruments. While shale thickness and depth are strong predictors of fracking activity in the post-2008 period, Table 5 shows that they are not statistically significant determinants of fracking before 2008. In addition, the coefficient estimates on $Thickness$ and $Depth$ in the second stage (unreported) are not statistically different from zero at standard cut-offs giving us further confidence that the shale variables are not spuriously correlated with other determinants of mortgage default.

Thus far, we have established that fracking activity has a negative effect on mortgage default risk, but an equally important issue that needs to be addressed is whether the effect is economically important. The estimate from the most saturated LPM specification (column (6) of Table 4), is -0.093, which implies that mortgage borrowers in ZIP Codes that have at some point experienced fracking activity are 0.093 percentage points less likely to default in a given month compared to borrowers in ZIP Codes without fracking.[29] While this is a seemingly small effect, the average (unconditional) monthly default rate during our sample period is 0.24 percentage points. Thus, the estimated impact on fracking is 38 percent of

[28]An alternative explanation for why instrumenting yields larger effects is measurement error, which typically biases estimates towards zero.

[29]Note, that in all tables we convert the units of our dependent variable, the default indicator variable, to basis points by multiplying by 100.

the average default rate.[30] The IV estimates, as discussed above, are quite a bit larger. The estimates from both IV specifications (-0.289 and -0.297) are greater than 100 percent of the monthly average default rate in our sample.

6 Robustness

In this section we briefly discuss additional empirical evidence that supports our interpretation that fracking activity causally reduces mortgage credit risk in our sample of Pennsylvania mortgages. Due to space considerations we display all regression estimates in the Internet Appendix.

6.1 Alternative Fracking Measures and Subsample Analysis

In addition to the "Any Fracking" indicator variable, we estimate our IV specifications with the alternative fracking measures described above in Section 4.1 and displayed in Table 3. These measures include variables that measure the flow and stock of permits and drilled wells. The results are displayed in Table A.1 and discussed in Section A.2.1 in the Internet Appendix, and are consistent with the baseline results reported in Table 4.

We also restrict the sample to various sub-populations of mortgages to see whether our results are sensitive to alternative control groups for our fracking treatment variable and to rule out potential confounders. First we exclude mortgages in ZIP codes that had any existing conventional wells in case the households residing in those areas may have anticipated an energy boom (if not the fracking boom) that coincided with the on set of the great recession. We also try excluding loans in metro areas, in case the fracking is simply capturing the relative performance of rural areas in the downturn. We also limit the sample to homes on or within 25 miles of the shale formation (regardless of depth). Finally, we restrict the

[30] We have also estimated a Cox proportional hazard model and obtained a hazard ratio of 0.75 associated with the "Any Fracking" variable, which implies that the likelihood of mortgage default is approximately 25 percent less in fracking ZIP Codes compared to non-fracking ZIP Codes. Those results are reported in the Internet Appendix.

sample to mortgages that had low credit scores at origination (below 660 and 620) in case drilling activity was spuriously correlated with lax underwriting. The results are located in Table A.2 in the Internet Appendix. In all subsamples, fracking reduces mortgage default risk.

6.2 Instrument Validity and a Placebo Test

In Section 5 we argued that the increased magnitude of the fracking coefficient in the instrumental variables regressions is likely the result of endogenous fracking activity—wells are more likely to be drilled near homes that were otherwise more at-risk of default. An alternative explanation is that our instrument, the depth and thickness of the shale, though exogenous, is spuriously correlated with some other determinant(s) of mortgage default and it is this correlated, but omitted variable, that is driving the results. We perform two exercises to address this issue.

First, we regress our shale instruments on a number of socio-demographic variables at the ZIP Code level, which we obtain from the 2000 U.S. Census long-form that might plausibly be associated with default. Specifically, we regress the instruments (shale thickness and depth) on self-reported median earnings, the unemployment rate, labor force participation, the share of the population with a bachelor's degree or higher (at least sixteen years of education), the share of the population with less than twelve years of education, median self-assessed home value, and the home ownership rate. We estimate regressions both in levels and the change between 1990 and 2000 in case there is some long-run pre-trend. The estimates are displayed in Table A.3 in the Internet Appendix. All pre-tend estimates are insignificant and all level values are insignificant except for the coefficient estimate corresponding to the regression of shale thickness on the level of median, self-assessed home values in 2000 which is negative and significantly different from zero at the 10 level.

While these findings give us some confidence that the IV results are not driven by omitted, socio-demographic characteristics, there may be other important omitted factors creating spurious correlation between the fracking instruments and the propensity to default. Of particular concern is that fracking activity might be correlated with some unobservable factors that reduce mortgage credit risk during a recession. To address this concern, we implement a placebo specification in which we repeat our baseline LPM and IV specifications using a sample of Pennsylvania loans that were active during the 2001 recession, a time period that predates fracking. According to the National Bureau of Economic Research (NBER) the 2001 recession was an eight-month economic downturn that began in March 2001. Thus, we define 2001–2003 (inclusive) as the post-treatment period and use the same empirical setup as in our baseline analysis.[31] We focus on loans originated between January 1997 and December 2000, before the 2001 recession, and we follow the performance history of those loans from origination through December 2003. Importantly, the sample used in the placebo tests does not overlap with the sample used to estimate our main results. The fracking measures as well as the instrumental variables (shale depth and thickness) in the placebo specification are defined in the same way as they are in our baseline analysis, except that we assume the timing of fracking activity in each ZIP Code occurred seven years earlier than it did in reality. For example, if fracking first appears in location i in January 2008, the placebo "Any Fracking" measure will be set to 1 in January 2001. The results of this placebo exercise are displayed in Table A.4 in the Internet Appendix. In all specifications the coefficient associated with the placebo fracking measures is positive rather than negative indicating that at least in the previous recession, mortgages in fracking areas did not perform better.

6.3 Accounting for Prepayment

Prepayment is another important source of risk in the mortgage market. One concern is that by ignoring prepayment risk, we may be introducing bias in our estimates of the effect

[31]Specifically, our two placebo instruments are the interaction between the shale thickness and depth measures and a $Post2000$ indicator variable (rather than the $Post2008$ indicator used in our baseline specifications).

of fracking activity on mortgage default behavior. For example, if households are more likely to sell their homes and move after fracking activity begins in their neighborhoods, then that could change the composition of the remaining pool of active mortgages in our sample towards homeowners with an unobservably strong desire to stay in their homes. This composition effect could be driving our main results on the negative relationship between fracking and mortgage default risk.

To address this issue we look at the relationship between fracking and prepayment behavior.[32] We first replicate our baseline specifications but change the dependent variable from a default indicator to an indicator for prepayment. These results are displayed in Table A.5 in the Internet Appendix. We do not find that increased fracking activity is associated with increased prepayment.

In a second exercise, we jointly model the borrower's decision to prepay or default using a competing risks, hazard model. Jointly modeling the two risks is a more direct way of controlling for the possibility that homeowners who might otherwise default are more likely to sell their homes and move in response to increased fracking activity. Using a multinomial logit model, which is one of the more commonly used competing risks models in the mortgage literature, we continue to find a statistically significant, negative relationship between fracking and mortgage default. These results are reported in Table A.6 in the Internet Appendix.

7 Causal Mechanisms

Thus far, we have documented a strong, robust, causal link between fracking activity and lower mortgage default in our sample of Pennsylvania mortgages. In this section we investigate the underlying mechanisms driving this result.

[32] It is important to note that we cannot distinguish in our data between a prepayment due to a mortgage refinance and a prepayment due to a home sale. We can only distinguish between voluntary prepayments and involuntary prepayments (i.e. foreclosures, short-sales, etc.)

7.1 The Role of Negative Equity

We begin by analyzing the impact of negative equity on the relationship between fracking and mortgage default. The double-trigger theory predicts that defaults are driven by a combination of negative equity, which makes it difficult for borrowers to sell or refinance to pay off their loans, and a negative liquidity shock such as an unemployment spell, which makes it difficult to make monthly mortgage payments. If the double trigger theory holds, and fracking reduces default by alleviating the liquidity trigger, then we should expect larger effects among homeowners in positions of negative equity. In this section, we test for this effect.

In order to implement such a test, we need to identify borrowers in our sample that are underwater on their mortgages. To do this, we create a new indicator variable for whether a borrower is underwater at time t by comparing the current estimated house value to the outstanding mortgage balance. We predict current home value based on the cumulative change in the monthly (CoreLogic) county house price index since the month of loan origination. The McDash database provides detailed monthly information on the outstanding balance of every mortgage, which we use to construct the following negative equity indicator:

$$d_t^{underwater} = \begin{matrix} 1, \ if \ \frac{balance_t}{value_t} > 1 \\ 0, \ if \ \frac{balance_t}{value_t} \leq 1 \end{matrix}$$

where:

$$value_t = \frac{balance_0}{LTV_0} \times \frac{HPI_t}{HPI_0}.$$

The terms $balance_0$, LTV_0, and HPI_0 correspond to the mortgage balance, LTV ratio, and monthly house price index at the month of loan origination, respectively. County subscripts for the HPI have been suppressed for simplicity.

It is important to note that our estimate of current value may be imprecise and is likely to underestimate the true number of underwater homeowners for at least two reasons. First, there may be a systematic, upward bias in all repeat-sales estimates if homes that are sold more frequently are subject to greater maintenance and renovation than homes that do not turn over as often. This could be especially true in central Pennsylvania where the structure rather than the land may comprise most of a property's value (absent mineral rights). Second, the decline in home sales during the crisis may introduce an upward bias in the HPI (and into our valuation estimates) if underwater homeowners simply choose not to sell, again skewing the population of repeat-sale homes.

In addition, while the outstanding balance of the first mortgage is tracked with precision over time in the McDash database, there is no information on the outstanding balances of subordinate liens. Given the widespread use of "piggy-back" loans during the housing boom, this means that we will overestimate equity and underestimate the number of loans that are underwater.

With these caveats in mind, we modify our baseline LPM specification from column (5) of Table 4, by adding the negative equity indicator to the regression and interacting it with our "Any Fracking" indicator.[33] We present these results in Table 6. In column (1) we first replicate our baseline specification on the subset of mortgages located in a county with a CoreLogic HPI. The "Any Fracking" coefficient estimate of -0.067 is similar in magnitude to the estimate in column (5) of Table 4 (-0.074). In column (2), we add the negative equity dummy to the specification. Loans that are currently underwater are 0.33 percentage points more likely to default in a given month, but controlling for underwater status does not appreciably change the effect of fracking on default. In column (3) we add the interaction term between the "Any Fracking" and underwater dummy variables and find that the ameliorative effect of fracking on default is significantly larger for underwater loans. The coefficient estimate associated with the interaction term is -0.212 percentage points, which is more than four times the effect (in absolute magnitude) of fracking on mortgages

[33]We revert to the LPM as we were unable to obtain a powerful enough first stage in our IV specification once we include the interaction term, $Underwater \times Any\ Fracking$, in our set of endogenous variables.

that we estimate as not being underwater (-0.051 percentage points).

This pattern is consistent with the double-trigger theory and with fracking ameliorating the liquidity trigger of default. It is also notable that the effect of fracking on borrowers with positive equity remains negative and statistically significant. There are at least two explanations for this result. First, it may be that some borrowers are defaulting with positive equity due to transitory liquidity shocks. In the absence of such shocks and other housing market frictions, we would expect all loan defaults to be limited to mortgages that are underwater. A homeowner with sufficient positive equity to cover transactions costs could simply sell the property instead of defaulting on the loan. In reality, however, home sales significantly declined during the Great Recession as buyers became more cautious and banks tightened their underwriting considerably. Thus, it is likely that even homeowners with positive equity who suffered an income and/or employment shock may have become severely delinquent while struggling to sell their homes. If this is the case, then the negative effect of fracking on the default propensity of positive equity homeowners is also consistent with the interpretation that fracking is ameliorating default by alleviating the liquidity trigger.

A second possibility is that we are significantly underestimating the magnitude of negative equity in our sample due to the measurement issues discussed above. If this is the case, then the positive equity result could be driven by some borrowers who are truly underwater but that we are misclassifying as having positive equity. In columns (4) and (5) of Table 6 we attempt to address two sources of measurement error in our underwater indicator variable. First, we try to reduce the number of mortgages in our sample that may have an unobserved subordinate lien by excluding mortgages with a LTV ratio at origination that is exactly equal to 80 percent. The intuition for this is that the primary reason to take out two separate loans at origination is to borrow the bulk of the loan using a conforming, GSE mortgage that typically carries a lower interest rate than a jumbo loan, and then borrow the remaining balance using a more costly second (piggy-back) mortgage. Since a GSE loan cannot have a LTV ratio above 80 percent without private mortgage insurance (PMI), borrowers that used piggy-back mortgages (rather than PMI) to finance their home purchase typically took out a first mortgage with a LTV ratio of exactly 80 percent. In column (4), we re-estimate the specification in column (3), but drop all loans that have a LTV ratio at origination of exactly 80 percent. The results are largely unaffected by this change.

Finally, in column (5) we exclude all refinance mortgages. The intuition for this restriction is that the estimate of the value of the home at the time of origination for a refinance loan is typically based either on an appraisal or on an automated valuation model (AVM). There is significant evidence of inflated home appraisals during the housing boom period, which may cause us to underestimate the true number of underwater refinance mortgages.[34] In contrast, the value of the home at the time of origination for a purchase loan is typically based on the transaction price. The results in column (5) show that excluding refinance mortgages does not appreciably change the estimates.

7.2 Fracking and House Prices

Next, we estimate the causal effect of fracking activity on local house prices. House prices are a key variable to study because fracking could be reducing default through two different channels related to home values. First, fracking could be increasing borrower expectations of future house price gains and hence, reducing strategic mortgage default. If fracking causes increased economic activity at the local level through income effects related to royalties or through increasing employment opportunities in fracking-related industries, it should increase housing demand and hence, increase home values. Potential strategic defaulters in fracking areas would then have a lower incentive to default, as they would expect future house price gains. In addition, if fracking activity increases home prices, it would reduce the extent of negative equity among mortgage borrowers and thus ameliorate the first trigger in the double-trigger theory of default.

[34] A large literature exists on the issue of appraisal fraud during the housing boom in the early-to-mid 2000s. Papers on the topic include Ben-David (2011), Agarwal et al. (2014), Shi and Zhang (2015), Ding and Nakamura (2016), Calem et al. (2017), Conklin et al. (2019), and Kruger and Maturana (2017).

We look at changes in house prices using a county-month, repeat-sales, HPI produced by CoreLogic. This measure requires a certain minimum number of sales to compute and thus, is only available for about 60 percent of the counties in our sample (3,213 county-month observations).[35] In column (1) of Panel A of Table 7, we regress house price appreciation (HPA) since January 2007, which roughly corresponds to the market peak in many areas, on our instrumented "Any Fracking" dummy. In column (2) we consider HPA since January 2009, just after the widespread adoption of fracking (i.e. $\ln(\frac{HPI_t}{HPI_{2007/2009}})$). In neither case is fracking associated with increased house price growth. In fact, both IV estimates are negative but neither is statistically significant.

In column (3) of Panel A we consider HPA since the month of mortgage origination. Unlike columns (1) and (2), where the regressions are estimated at the county-month level, the IV specification in column (3) is estimated at the loan-month level.[36] The point estimate is negative, but is not statistically different from zero. Finally, in column (4) we estimate the effect of fracking activity on negative equity, which we measure in the same manner as we did in the analysis above (Section 7.1). While the point estimate is positive, it is not statistically significant.

Overall, the results reported in Panel A of Table 7 do not support the hypothesis that fracking lowers mortgage default through a house price channel. The point estimates suggest that fracking may actually lower house prices at the county-level, which would serve to increase mortgage default risk, *ceteris paribus*. However, given that the estimates are quite noisy and not statistically different from zero, we do not emphasize such an interpretation.

[35]The HPIs cover more than 90 percent of our loan-month observations, as the missing HPI observations correspond to sparsely populated counties.

[36]For all county-month regressions in Table 7, we aggregate the "Any Fracking" variable to the county level by taking its maximum value across all ZIP Codes in a given county. Thus, if there is fracking activity in one ZIP Code in a county-month, the "Any Fracking" variable takes a value of 1. We also estimated alternative IV regressions where we considered the average of the "Any Fracking" variable across all ZIP Codes in a county (i.e. the share of fracking ZIP Codes in a county) and found similar results.

7.3 Fracking, Income, and Employment

We now turn to an analysis of the effect of fracking on household income and employment. In Panels B and C of Table 7 we display results from IV regressions that estimate the causal effect of fracking activity on local income and employment measures. If fracking raises household income through increased royalty payments or lowers the risk of unemployment, then we would expect it to mitigate ability-to-pay issues and thus, ameliorate the second trigger of default.

Our income variables in Panel B come from the IRS's SOI database, which provides median adjusted gross income (AGI), royalty income, wage and salary income, and income derived from unemployment insurance (UI) benefits from tax returns by ZIP Code and year.[37] We express all income variables in per-household terms, where we divide total income by the number of tax returns in the ZIP Code. Column (1) shows that fracking predicts large and statistically significant increases in per-household AGI. The coefficient estimate of 0.147 suggests that fracking increases average household AGI by approximately 15 percent. Columns (2) and (3) reveal that the AGI effect is a result of fracking significantly increasing per-household royalty income and per-household wage and salary income. Specifically, fracking activity in a ZIP Code raises wage and salary income by almost 7 percent and raises royalty income by more than 29%. At the same time, column (4) shows that fracking substantially reduces income derived from UI benefits. These results suggest that workers in fracking ZIP Codes are being paid more, and are earning more income through royalties, while avoiding more unemployment spells compared to workers in non-fracking ZIP Codes.

In Panel C of Table 7 we estimate the effect of fracking on county-level unemployment rates and county-level employment growth, which we obtain from the Bureau of Labor Statistics (BLS). In column (1) we regress county-level, monthly unemployment rates on our measure of fracking activity while in columns (2)–(4) we focus on employment growth (year-over-year) for a selection of two digit NAICs industries. While our estimate of the

[37]We thank an anonymous referee for directing us to this data source.

relationship between unemployment rates and fracking is negative in column (1), it is not statistically significant. In column (2) the dependent variable is employment growth in the "Mining, Quarrying, and Oil and Gas Extraction" industries (NAICS 21). The coefficient estimate on the "Any Fracking" dummy is positive, large (32 percent) and statistically different from zero at the 10 percent level. In column (3) we consider employment growth in the construction industry (NAICS 23). This variable reflects employment flows associated with building facilities for the shale industry and may also reflect employment flows related to more indirect fracking effects such as growing demand for housing or commercial space. Fracking increases construction employment growth by approximately 13 percent. Finally, column (4) shows results for employment growth in the manufacturing sector, which serves as a rough proxy for the tradeable goods sector that may experience crowding out in a resource boom. The coefficient estimate in this case is negative but is not statistically different from zero.

In summary, the results in Table 7 show that fracking does not increase house prices at the county-level or alleviate negative equity. This suggests that the causal mechanism through which fracking lowers mortgage default is not a strategic/ruthless default channel. At the same time, the table provides evidence that fracking activity significantly strengthened the labor market in Pennsylvania, especially in fracking-related industries, and raised household incomes. Fracking ZIP Codes enjoyed higher royalty income as well as higher wage and salary income, while experiencing significantly lower income from UI benefit claims. These results are consistent with fracking alleviating the second trigger of default.

7.4 Fracking and Default on Other Types of Debt

In this section, we test whether fracking activity reduces the default risk associated with other types of credit instruments, including second mortgages, credit cards, auto loans, and student loans. If our interpretation of the empirical evidence presented above is correct, and fracking ameliorates mortgage default risk by mitigating negative income/employment shocks, then we should expect it to also reduce default rates associated with other types of consumer debt. In contrast, if fracking reduces mortgage default through a strategic channel by lowering the likelihood of negative equity and raising the probability of future capital gains, then we should not expect to find any fracking effects on the propensity of a borrower to default on non-mortgage debt.

To implement this test, we introduce an additional database into the analysis called CRISM (Equifax Credit Risk Insight Servicing McDash Database). CRISM is a dataset that consists of McDash mortgages matched to credit bureau data from Equifax at the borrower level.[38] The Equifax credit bureau data is updated at a monthly frequency and includes information on outstanding consumer loans and credit lines for the primary borrower as well as all co-borrowers associated with the McDash mortgage.[39]

We focus our analysis on the most popular types of consumer debt including first and second mortgages, credit cards, auto loans, and student loans. Identifying defaults in the CRISM database is not as straightforward as it is in McDash, where we have direct, detailed information from the servicer on the payment status of each mortgage. CRISM does not provide detailed information on each credit line and/or loan, but rather it provides aggregate information at the tradeline level on total outstanding balances, total number of loans/lines, total outstanding balances that are "current" and the total number of loans/lines that are "current." We use these variables to identify borrowers who have missed payments and are behind on their loans and/or credit lines. To maintain consistency with our 90–days delinquent mortgage default definition we require the borrower to be behind on payments for at least three consecutive months before considering the borrower to be in default.

We estimate instrumental variable regressions that mirror our baseline IV regressions in Table 4 (column (7)), but now focus on testing whether fracking activity has causal effects on the default risk of a broader array of credit instruments. Similar to our baseline regression,

[38]The matching process was conducted by Equifax using confidential and proprietary data. Coverage begins in June 2005, and according to Equifax, approximately 90 percent of McDash mortgages were matched to a credit bureau account with high confidence.

[39]We keep only observations that pertain to the primary mortgage borrower to avoid double-counting.

they are estimated at the borrower-month level and include year-month and county fixed effects. Since CRISM does not provide details of the underlying loans/credit lines, our set of covariates is much more limited compared to the set that we include in the McDash mortgage default regressions. The set includes a quadratic polynomial for the primary borrower's age and the borrower's credit (FICO) score in the first month (i.e. the month in which the first mortgage was originated according to McDash). Finally, borrower-month observations are only included in the regressions if the number of reported loans/credit lines is greater than zero for a given tradeline.

Table 8 displays the estimation results. Consistent with our results based on the McDash sample, in column (1) we see that fracking activity significantly reduces the propensity to default on first mortgage debt. Specifically, borrowers in ZIP codes with any fracking activity are 0.037 percentage points less likely to default on their first mortgage, which is a sizeable effect given that the average, monthly default rate in the CRISM sample is 0.068 percentage points (displayed in the first row of the table).[40]

In column (2) we see that fracking also significantly mitigates defaults associated with second mortgage debt, while in column (3) we find evidence that fracking activity mitigates credit card defaults. Borrowers in ZIP Codes with fracking activity are approximately 0.15 percentage points less likely to default on their credit cards compared to borrowers in non-fracking ZIP Codes. This is a sizeable effect as it is more than 50% of the average monthly credit card default rate in the CRISM sample (0.28 p.p.). Finally, in columns (4) and (5) we do not find statistically significant effects of fracking activity on auto loan defaults or student loan defaults (although the point estimates are negative in both cases).

[40]The average default rates on first mortgage debt in McDash and CRISM are quite different (0.24 p.p. versus 0.068 p.p., respectively) and the estimated fracking treatment effects are also different (-0.297 p.p versus -0.044 p.p., respectively). This is not surprising, as information on first mortgage balances in CRISM is not always the same as the information on the first mortgage in McDash. If a borrower has multiple properties with a first mortgage on each property, CRISM will capture the aggregate $ amount outstanding across all of the properties. In addition, even if the balances are identical, the information on delinquency in CRISM comes from Equifax, while the information on delinquency in McDash comes from mortgage servicers directly. These differences likely explain the different default rates and estimated fracking treatment effects in Tables 4 (column (7)) and 8 (column (1)).

In summary, we find evidence that fracking activity lowers both mortgage default risk as well as credit card default risk. This is consistent with the double-trigger theory of mortgage default, and specifically with a causal mechanism whereby increased fracking activity reduces the risk of negative income and/or employment shocks.

8 Conclusion

This paper uses a novel instrumental variables approach to estimate the causal effect of hydraulic fracturing on mortgage credit risk using detailed micro data on mortgages originated in Pennsylvania in the period before the fracking boom. The IV approach uses variation in the underlying geologic properties of the Marcellus Formation and the timing of the Pennsylvania shale boom, which was determined by exogenous innovations in horizontal drilling technology that emerged in the late 2000s.

We document a robust, economically significant, *negative* relationship between fracking activity and mortgage credit risk. Pennsylvania ZIP Codes that experienced higher fracking activity were characterized by significantly lower mortgage default rates, *ceteris paribus.* While there are numerous channels by which the shale boom could have positively impacted the Pennsylvania housing market, the weight of the evidence points to fracking strengthening local labor markets and improving household finances through higher royalty payments wages and salaries. In addition, the negative effect of fracking on default is approximately four times greater for borrowers that are in negative equity positions, compared to those that have positive or zero home equity.

Overall, the evidence is consistent with the importance of the double-trigger theory, which holds that mortgage defaults are generated by adverse shocks to borrowers' ability to repay their mortgages, in addition to negative house price shocks.

References

Agarwal, S., B. W. Ambrose, and V. W. Yao (2014). The limits of regulation: Appraisal bias in the mortgage market. *Penn State University Working Paper*.

Agency, E. I. (2010). Summary: U.s. crude oil, natural gas, and natural gas liquids proved reserves 2009. Technical report, Energy Information Agency, Office of Oil, Gas, and Coal Supply Statistics, Department of Energy.

Agency, E. I. (2015). Top 100 u.s. oil and gas fields. Technical report, Energy Information Agency, Office of Oil, Gas, and Coal Supply Statistics, Department of Energy.

Anderson, N. B. and J. K. Dokko (2016). Liquidity problems and early payment default among subprime mortgages. *Review of Economics and Statistics 98*(5), 897–912.

Ben-David, I. (2011). Financial constraints and inflated home prices during the real estate boom. *American Economic Journal: Applied Economics 3*(3), 55–87.

Blohm, A., J. Peichel, C. Smith, and A. Kougentakis (2012). The significance of regulation and land use patterns on natural gas resource estimates in the marcellus shale. *Energy Policy 50*, 358–369.

Boslett, A., T. Guilfoos, and C. Lang (2016). Valuation of expectations: A hedonic study of shale gas development and new york's moratorium. *Journal of Environmental Economis and Management 77*, 14–30.

Bradley, M. G., A. C. Cutts, and W. Liu (2015). Strategic mortgage default: The effect of neighborhood factors. *Real Estate Economics 43*(2), 271–299.

Calem, P. S., L. Lambie-Hanson, and L. I. Nakamura (2017). Appraising home purchase appraisals. *FRB of Philadelphia Working Paper No. 17-23*.

Conklin, J., E. Coulson, M. Diop, and T. Le (2019). Competition and appraisal inflation. *The Journal of Real Estate Finance and Economics Forthcoming.*

Ding, L. and L. Nakamura (2016). The impact of the home valuation code of conduct on appraisal and mortgage outcomes. *Real Estate Economics 44*(3), 658–690.

Gerardi, K., K. F. Herkenhoff, L. E. Ohanian, and P. S. Willen (2017). Cant pay or wont pay? unemployment, negative equity, and strategic default. *The Review of Financial Studies 31*(3), 1098–1131.

Gilje, E. P. (2019). Does local access to finance matter? evidence from us oil and natural gas shale booms. *Management Science 65*(1), 1–18.

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

Guiso, L., P. Sapienza, and L. Zingales (2013). The determinants of attitudes toward strategic default on mortgages. *The Journal of Finance 68*(4), 1473–1515.

James, A. and J. James (2014). A canary near a gas well: Gas booms and housing market busts in colorado. Technical report, University of Alaska Anchorage.

Jiang, W., A. A. Nelson, and E. Vytlacil (2014). Liar's loan? effects of origination channel and information falsification on mortgage delinquency. *Review of Economics and Statistics 96*(1), 1–18.

Kruger, S. A. and G. Maturana (2017). Collateral misreporting in the rmbs market. *Working Paper*.

Muehlenbachs, L., E. Spiller, and C. Timmins (2015). The housing market impacts of shale gas development. *The American Economic Review 105*(12), 3633–3659.

Pence, K. M. (2006). Foreclosing on opportunity: State laws and mortgage credit. *Review of Economics and Statistics 88*(1), 177–182.

Piskorski, T., A. Seru, and V. Vig (2010, September). Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis. *Journal of Financial Economics 97*(3), 369–397.

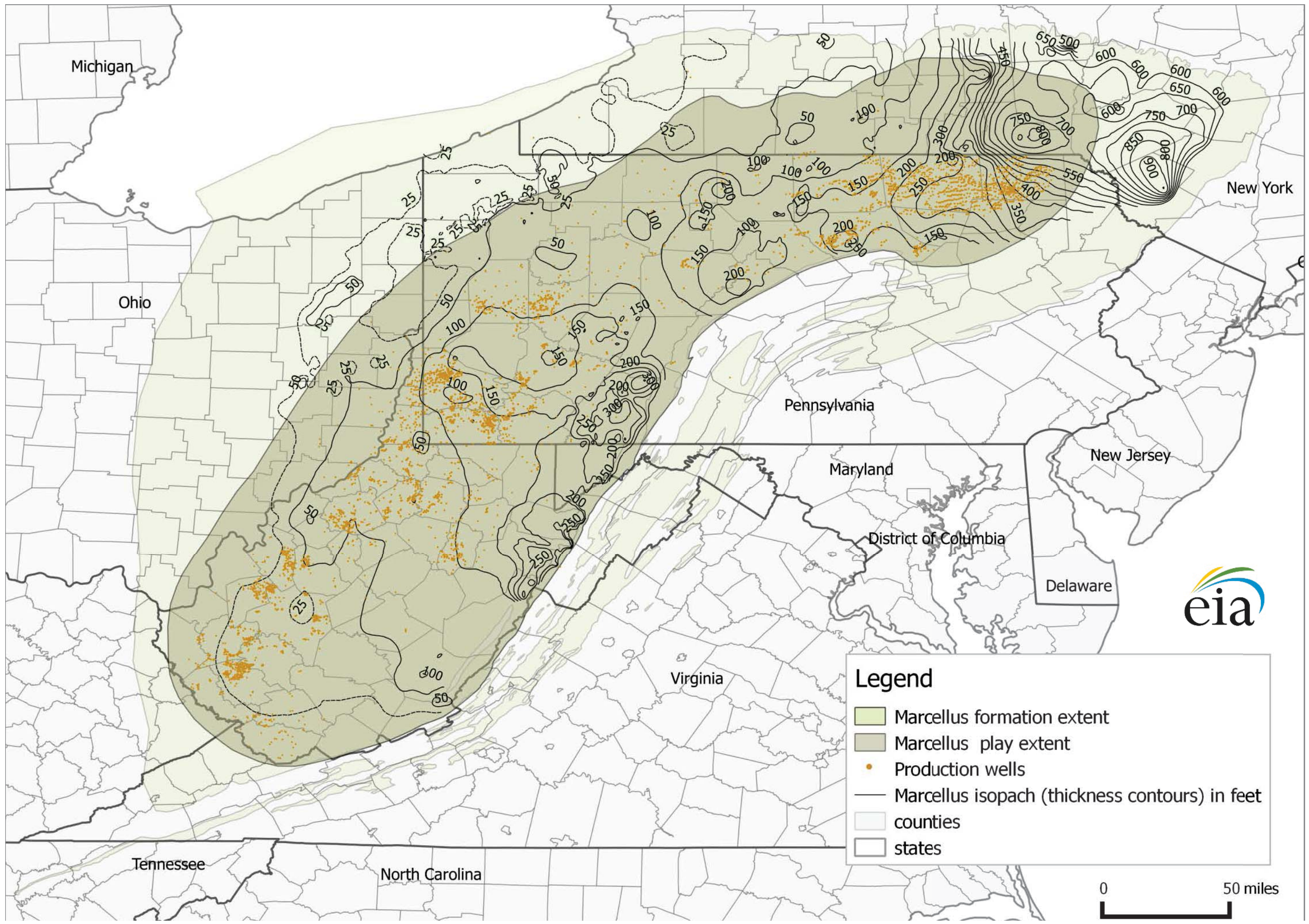
Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy 82*(1), 34–55.

Shi, L. and Y. Zhang (2015). Appraisal inflation: Evidence from the 2009 gse hvcc intervention. *Journal of Housing Economics 27*, 71–90.

Weber, J., J. Burnett, and I. Xiarchos (2014). Shale gas development and housing values over a decade: Evidence from the barnett shale. Technical report, United States Association for Energy Economics.

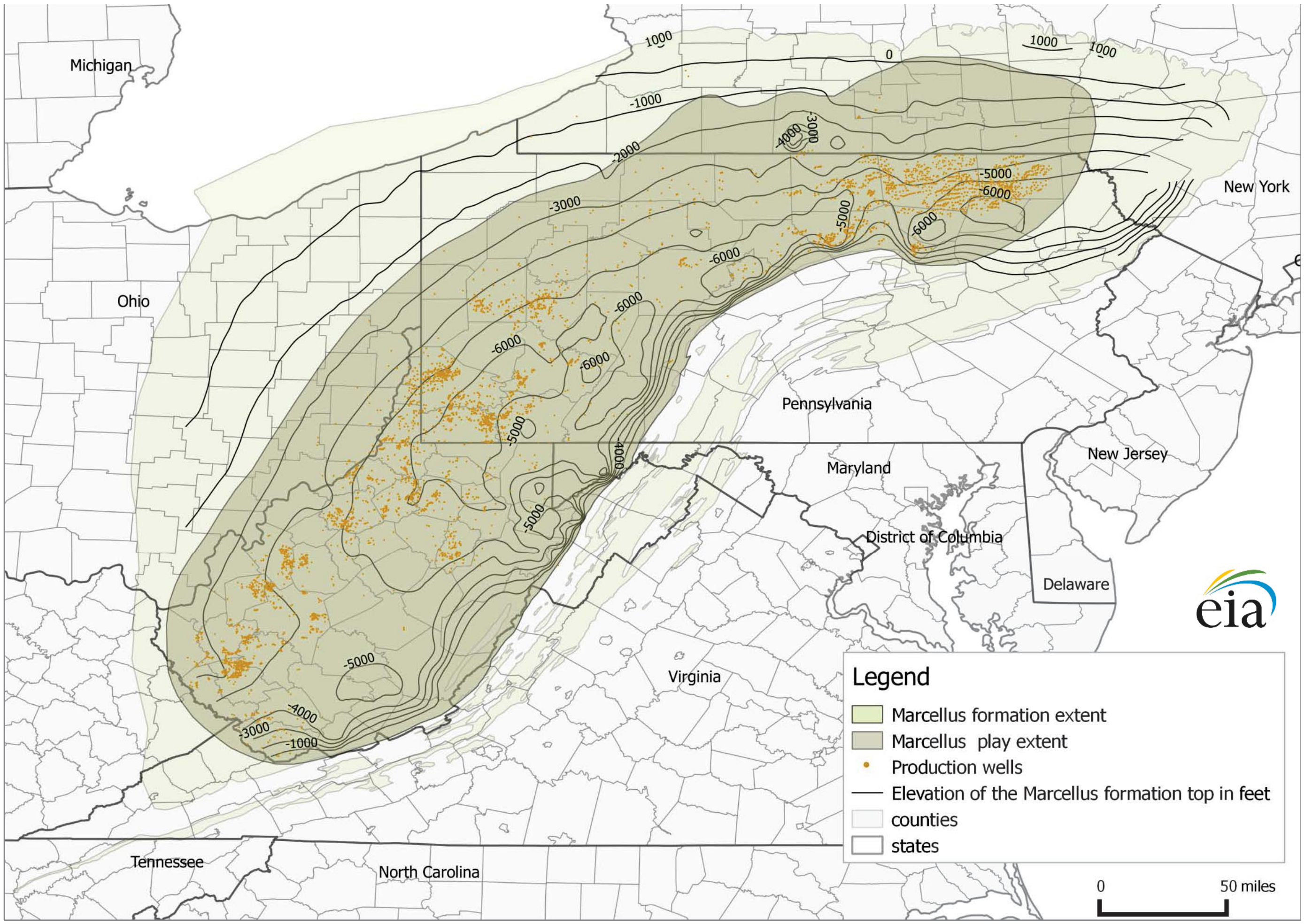
Wrightstone, G. (2009). Marcellus shale–geologic controls on production. In *AAPG Annual Convention, Denver, CO, Jun 7-10*.

Figure 1: Marcellus Shale Thickness



Source: U.S. Energy Information Administration, based on DrillingInfo Inc., New York State Geological Survey, Ohio State Geological Survey, Pennsylvania Bureau of Topographic & Geologic Survey, West Virginia Geological & Economic Survey, and U.S. Geological Survey. The map includes production wells through December 2014.

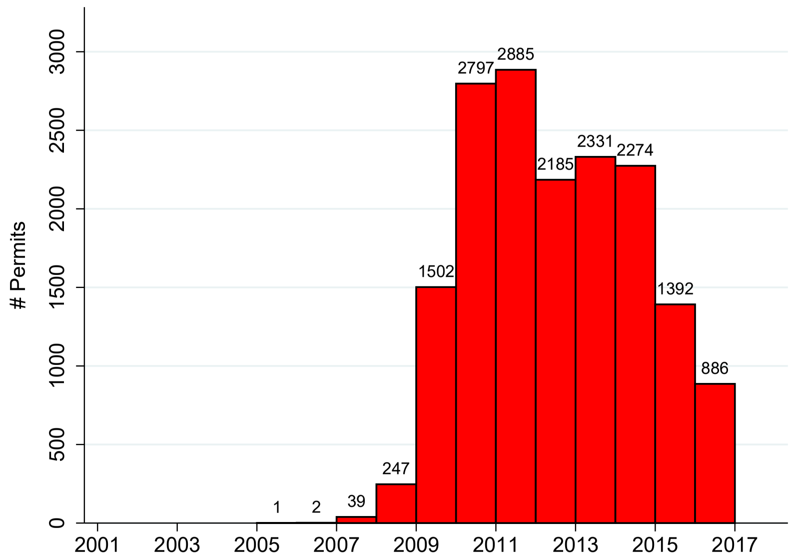
Figure 2: Marcellus Shale Depth

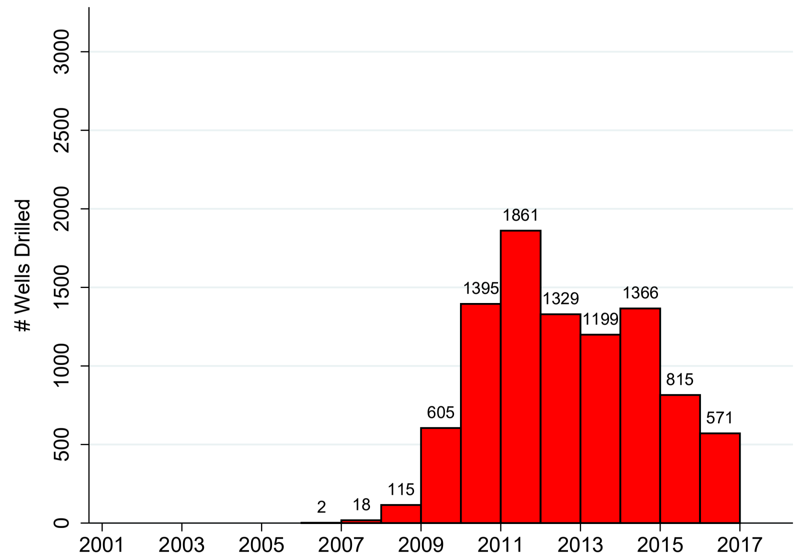


Source: U.S. Energy Information Administration, based on DrillingInfo Inc., New York State Geological Survey, Ohio State Geological Survey, Pennsylvania Bureau of Topographic & Geologic Survey, West Virginia Geological & Economic Survey, and U.S. Geological Survey. The map includes production wells through December 2014.

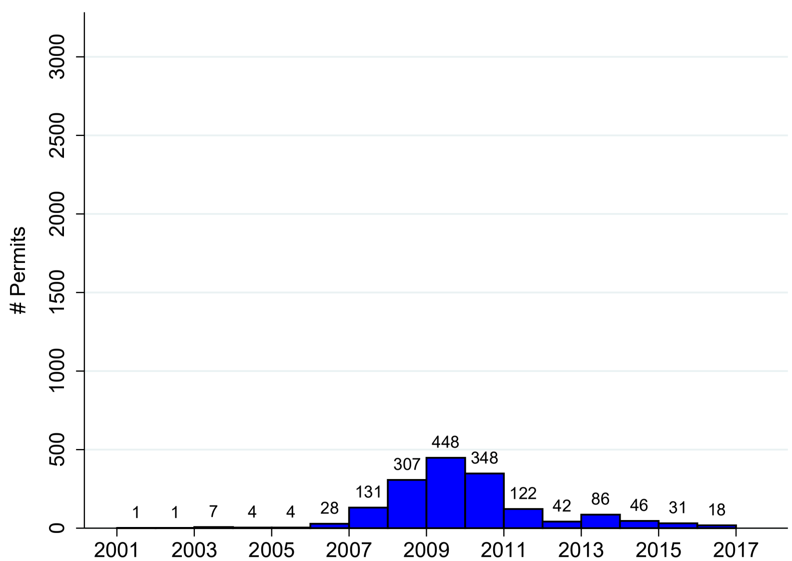
Figure 3: Count of Annual Permits and Newly Drilled Fracking Wells

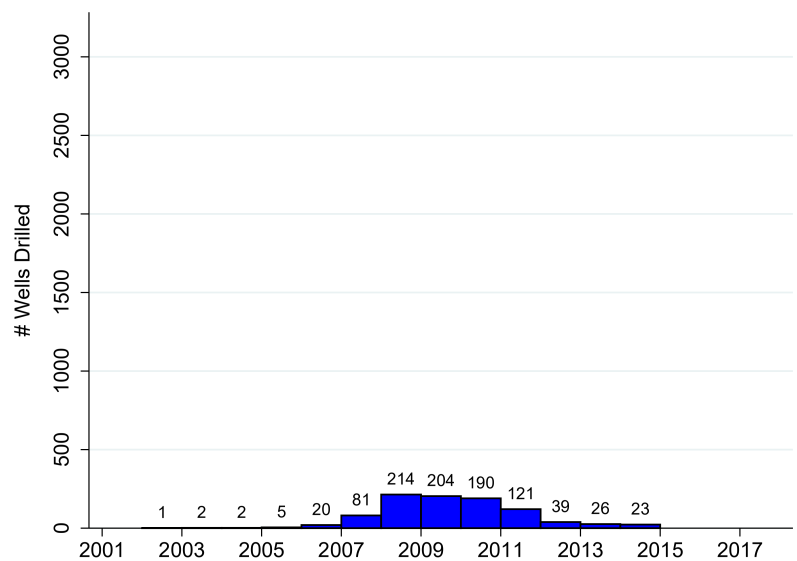
Panel A: Horizontal Wells





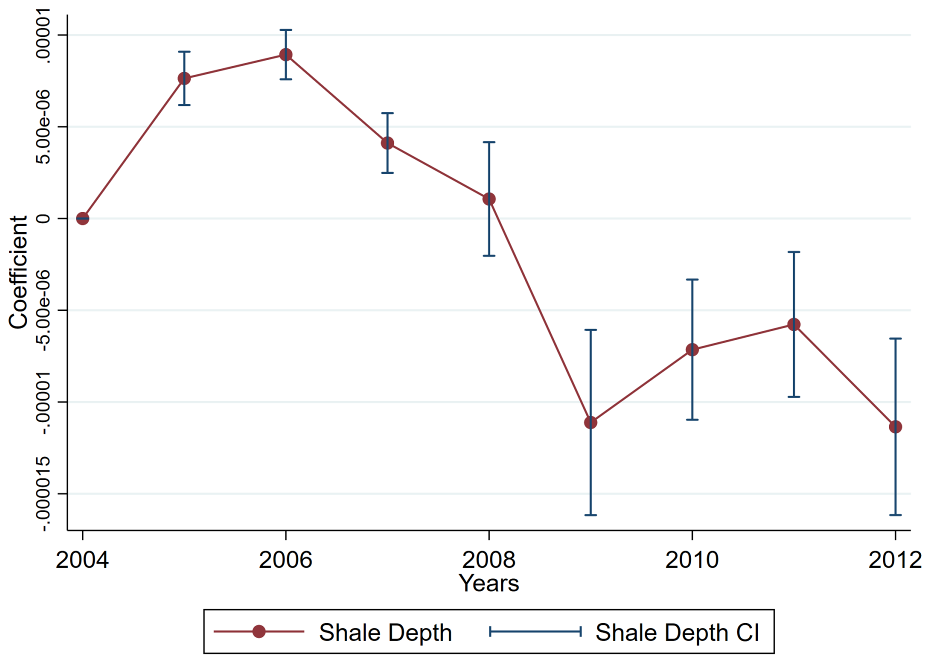
Panel B: Vertical and Directional Wells

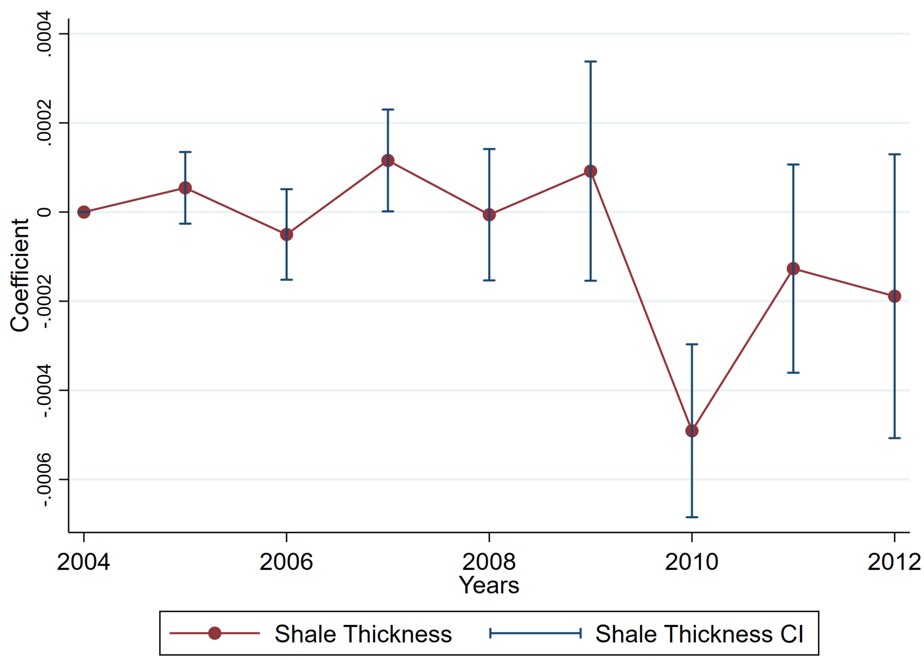




Notes: This figure displays the annual counts of fracking permits and drilled (or "drilled") wells. Panel A displays counts of horizontal wells and Panel B displays counts of vertical and directional wells. The underlying data were obtained from the Pennsylvania Department of Environmental Protection.

Figure 4: Mortgage Default and Shale Depth and Thickness Over Time





Source: Coefficient estimates from a reduced form regression of mortgage default on the instrumental variables shale depth and thickness interacted with calendar year dummies. Standard errors are calculated using two-way clustering by county–year and loan ID. All specifications include mortgage age (quadratic) and mortgage underwriting characteristics: year of origination, winsorized FICO score and dummies for common thresholds, mortgage rate, DTI, LTV, and dummies for whether LTV>80, fixed-rate, 30-year term, refinance, prepayment penalty, private-label, portfolio, balloon payment or jumbo. The regression also includes calendar year and county fixed effects.

Table 1: Fracking Permit Outcomes: Pennsylvania 2001–2017

Well Status	Count	Percent
Abandoned	3	0.02
Active	9,816	52.92
Not Drilled	4,412	23.79
Plugged	653	3.52
Proposed but Never Materialized	2,872	15.48
Regulatory Inactive Status	793	4.28
Total Permits	18,549	100

Notes: This table displays the status of all fracking permits issued in PA through January 2017. "Active" means that a well has been drilled and is either producing or expected to produce in the near future. "Proposed but Never Materialized" means that a permit was issued, but expired prior to the commencement of drilling. "Plugged OG Well" means that a well was drilled but was subsequently plugged (or capped) by the operator. "Operator Reported Not Drilled" means that a well was never drilled. "Regulatory Inactive Status" means that the well was drilled and is capable of producing but is temporarily capped (usually for up to 5 years) at the request of the well operator and approval by the PA DEP. "Abandoned" means that the well was drilled but has not been used to produce, extract or inject any gas, petroleum or other liquid within the preceding 12 months.

Table 2: Summary Statistics for Fracking Variables

	Mean	Std. Dev.	Minimum	Median	Maximum
Any Fracking (d)	0.03	.	0	0	1
Active Permits	0.41	5.73	0	0	295
Cumulative Permits Issued	0.55	7.70	0	0	374
Newly Permitted Wells	0.03	0.44	0	0	31
Cumulative Drilled Wells	0.24	3.73	0	0	189
Newly Drilled Wells	0.01	0.23	0	0	23
# Observations			9,766,708		
# Loans			194,026		

Notes: This table displays sample summary statistics for the various measures of fracking activity included in the analysis. Dummy variables are signified by (d). The sample corresponds to loan-month observations in the 2004–2012 period associated with mortgages originated in PA in the 2004–2006 period.

Table 3: Summary Statistics for Mortgage Variables

	Full Sample (1)	Never Fracked (2)	Ever Fracked (3)
Default (90+ Days Delinquent) (d)	0.126	0.126	0.129
	(0.33)	(0.33)	(0.34)
FICO $\leq$ 580 (d)	0.030	0.030	0.036
	(0.17)	(0.17)	(0.19)
580 $<$ FICO $\leq$ 620 (d)	0.056	0.055	0.067
	(0.23)	(0.23)	(0.25)
620 $<$ FICO $\leq$ 660 (d)	0.137	0.135	0.150
	(0.34)	(0.34)	(0.36)
660 $<$ FICO $\leq$ 700 (d)	0.194	0.194	0.197
	(0.40)	(0.40)	(0.40)
FICO	710	711	705
	(61.5)	(61.3)	(62.5)
DTI	35.2	35.2	35.1
	(13.8)	(13.7)	(14.5)
LTV	77.6	77.4	79.4
	(15.3)	(15.4)	(14.6)
Interest Rate	6.17	6.16	6.23
	(0.78)	(0.78)	(0.78)
Mortgage Amount ($)	139,000	143,000	103,000
	(75,803)	(76,437)	(58,479)
LTV $>$ 80% (d)	0.129	0.129	0.131
	(0.34)	(0.34)	(0.34)
Low Documentation (d)	0.326	0.331	0.283
	(0.47)	(0.47)	(0.45)
Fixed-Rate Mortgage (FRM) (d)	0.877	0.874	0.904
	(0.33)	(0.33)	(0.29)
Balloon (d)	0.006	0.006	0.004
	(0.08)	(0.08)	(0.07)
Interest-Only (d)	0.037	0.039	0.023
	(0.19)	(0.19)	(0.15)
Private-Label (d)	0.214	0.214	0.216
	(0.41)	(0.41)	(0.41)
Refinanced Mortgage (d)	0.407	0.405	0.423
	(0.49)	(0.49)	(0.49)
Term $\neq$ 30 years (d)	0.193	0.190	0.230
	(0.40)	(0.39)	(0.42)
# Loans	194,026	175,194	18,832

Note: This table displays sample summary statistics for the mortgage underwriting variables included in the analysis. Sample means are displayed along with standard deviations in parantheses. Dummy variables are signified by (d). Other variables not shown in the table include dummies for loans with LTV ratios that are exactly equal to 80% , prepayment penalties, primary mortgage insurance, and jumbo loans. All continuous variable are winsorized at the 5% level. Separate dummy variables for any missing values are also included.

Table 4: Baseline Specification: "Any Fracking"

Dependent Variable: 90+ Days Delinquent (d)								
Model:	LPM	LPM	LPM	LPM	LPM	LPM	IV Shale × # Permits	IV Shale × Post2008
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any Fracking (d)	-0.026	-0.079**	-0.088***	-0.070**	-0.074**	-0.093***	-0.289***	-0.297***
	(0.033)	(0.025)	(0.025)	(0.025)	(0.024)	(0.024)	(0.070)	(0.072)
FICO ≤ 580 (d)		0.705***	0.705***	0.709***	0.710***	0.716***	0.708***	0.708***
		(0.109)	(0.109)	(0.108)	(0.109)	(0.107)	(0.108)	(0.108)
580 < FICO ≤ 620 (d)		0.273***	0.273***	0.276***	0.276***	0.279***	0.276***	0.276***
		(0.050)	(0.051)	(0.050)	(0.050)	(0.049)	(0.050)	(0.050)
620 < FICO ≤ 660 (d)		0.142***	0.143***	0.143***	0.143***	0.145***	0.142***	0.142***
		(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.020)	(0.020)
660 < FICO ≤ 700 (d)		0.041**	0.041**	0.040**	0.040**	0.041**	0.039**	0.039**
		(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)
FICO Score		-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Low Documentation (d)		0.030**	0.029*	0.025*	0.026*	0.024*	0.026**	0.026**
		(0.013)	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
Private-Label (d)		0.135***	0.131***	0.133***	0.133***	0.131***	0.133***	0.133***
		(0.028)	(0.026)	(0.026)	(0.027)	(0.026)	(0.026)	(0.026)
Pre-Payment Penalty (d)		0.153***	0.155***	0.153***	0.152***	0.152***	0.152***	0.153***
		(0.052)	(0.052)	(0.051)	(0.051)	(0.051)	(0.050)	(0.050)
Underwriting Covariates	no	yes	yes	yes	yes	yes	yes	yes
Calendar Year FE	no	no	yes	yes	no	no	no	no
Year×Month FE	no	no	no	no	yes	yes	yes	yes
County FE	no	no	no	yes	yes	no	yes	yes
ZIP Code FE	no	no	no	no	no	yes	no	no
K-P F-stat							18.0	12.2
# Observations	9,766,708	9,766,708	9,766,708	9,766,708	9,766,708	9,766,708	9,766,708	9,766,708
R^2	0.000	0.004	0.004	0.004	0.004	0.004		

Notes: This table displays results corresponding to the baseline specification (equation (1) in the text). Columns (1)–(6) display linear probability model (LPM) results, while columns (7)–(8) display instrumental variable results. The dependent variable in all columns is an indicator variable for whether the mortgage is 90 or more days delinquent. "Any Fracking" is a dummy variable equal to 0 until the month the first fracking permit in the ZIP Code was issued. Standard errors are calculated using two-way clustering by county-year and loan ID (*** p<0.01, ** p<0.05, * p<0.1). All specifications include mortgage age (quadratic) and columns (2)–(8) include mortgage underwriting characteristics: year of origination, winsorized FICO score and dummies for common thresholds, mortgage rate, DTI, LTV, and dummies for whether LTV>80, fixed-rate, 30-year term, refinance, portfolio, balloon payment or jumbo. First stage estimates for columns (7) and (8) are located in Table 5.

Table 5: First Stage Results for 2SLS IV Specifications

Dependent Variable: Any Fracking (d)		
	IV Shale × # Permits (7)	IV Shale × Post2008 (8)
Thickness × #Permits	5.40e-07***	
	(1.91e-07)	
Depth × #Permits	1.63e-08***	
	(5.23e-09)	
Thickness × Post2008		0.0012***
		(0.0005)
Depth × Post2008		3.75e-05***
		(1.32e-06)
Thickness	0.0004	0.0005
	(0.0003)	(0.0004)
Depth	-3.75e-06	-6.61e-07
	(5.65e-06)	(5.21e-06)
FICO Score	-0.00002*	-0.00003*
	(0.00001)	(0.00001)
FICO < 580 (d)	-0.0093**	-0.0096**
	(0.0043)	(0.0043)
580 ≤ FICO < 620 (d)	-0.0035	-0.0036
	(0.0029)	(0.0030)
620 ≤ FICO < 660 (d)	-0.0025	-0.0026
	(0.0017)	(0.0017)
660 ≤ FICO < 700 (d)	-0.0015	-0.0015
	(0.0011)	(0.0011)
Low Documentation (d)	0.0004	0.0004
	(0.0009)	(0.0009)
Private-Label (d)	0.0009	0.0007
	(0.0011)	(0.0011)
K-P F-stat	18.0	12.2
# Observations	9,766,708	9,766,708
R^2	0.003	0.003

Notes: This table displays the first stage estimation results corresponding to columns (7) and (8) in Table 4. Standard errors are calculated using two-way clustering by county and year (*** p<0.01, ** p<0.05, * p<0.1). All specifications include mortgage age (quadratic) and columns (2)–(8) include mortgage underwriting characteristics: year of origination, winsorized FICO score and dummies for common thresholds, mortgage rate, DTI, LTV, and dummies for whether LTV>80, fixed-rate, 30-year term, refinance, prepayment penalty, private-label, portfolio, balloon payment or jumbo. The Kleibergen-Paap Wald rank F statistic is displayed for each IV regression. The mean value of the dependent variable,"Any Fracking" is 0.03.

Table 6: Negative Equity and Default

Dependent Variable: 90+ Days Delinquent (d)					
Sample Restrictions:				Exclude LTV = 80%	Exclude Refinances
	(1)	(2)	(3)	(4)	(5)
Any Fracking (d)	-0.067**	-0.065**	-0.051**	-0.051**	-0.058**
	(0.022)	(0.022)	(0.020)	(0.019)	(0.025)
Underwater (d)		0.333***	0.339***	0.344***	0.333***
		(0.056)	(0.055)	(0.061)	(0.057)
Underwater × Any Fracking (d)			-0.212**	-0.240***	-0.201**
			(0.073)	(0.063)	(0.078)
FICO score	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FICO below 580 (d)	0.702***	0.694***	0.694***	0.687***	0.783***
	(0.106)	(0.108)	(0.108)	(0.106)	(0.129)
FICO 580 to 620 (d)	0.282***	0.274***	0.274***	0.278***	0.339***
	(0.050)	(0.051)	(0.050)	(0.050)	(0.051)
FICO 620 to 660 (d)	0.149***	0.145***	0.145***	0.143***	0.164***
	(0.022)	(0.021)	(0.021)	(0.021)	(0.023)
FICO 660 to 700 (d)	0.041**	0.039**	0.039**	0.036**	0.035**
	(0.014)	(0.013)	(0.013)	(0.014)	(0.014)
# Observations	9,009,016	9,009,016	9,009,016	7,854,817	4,914,017
R^2	0.004	0.004	0.004	0.004	0.004

Notes: This table displays results on the estimated relationship between mortgage default and fracking. All columns correspond to linear probability models. Every LPM specification includes county fixed effects, year-month fixed effects, mortgage age (quadratic), year of origination fixed effects, winzorized FICO score plus dummies for common FICO thresholds at 580, 620, 660 and 700, the interest rate of the mortgage, the borrower's debt-to-income-ratio (DTI) at origination, the loan-to-value (LTV) ratio at origination, and whether the mortgage was fixed rate, low documentation, featured a balloon payment, was jumbo, had a prepayment penalty, was a refinance of an existing mortgage, and whether mortgage term was 30 years. Standard errors are calculated using two-way clustering by county-year and loan ID. (*** p<0.01, ** p<0.05, * p<0.1.)

Table 7: Direct Effects of Fracking Activity on House Price Growth, Income, and Employment

Panel A: House Price Variables

Aggregation Level:	County-month		Loan-month	
Dependent Variables:	HPA Since Jan. 2007	HPA Since Jan. 2009	HPA Since Origination	Negative Equity (d)
	(1)	(2)	(3)	(4)
Any Fracking (d)	-0.096	-0.145	-0.047	0.221
	(0.123)	(0.149)	(0.080)	(0.176)
K-P F-stat	10.5	10.5	13.8	13.8
# Observations	3,213	3,213	9,009,016	9,009,016

Panel B: Income Variables

Aggregation Level:	ZIP Code-year			
Dependent Variables:	Log(AGI)	Log(Royalty Income)	Log(Wage + Salary Income)	Log(UI Benefits)
	(1)	(2)	(3)	(4)
Any Fracking (d)	0.147***	0.294***	0.070**	-0.378**
	(0.034)	(0.082)	(0.023)	(0.132)
K-P F-stat	28.3	28.3	28.3	29.2
# Observations	85,912	85,877	85,879	83,713

Panel C: Employment Variables

Aggregation Level:	County-month			
		Industry Employment Growth (year-over-year)		
Dependent Variables:	U.E. Rate	Drilling/Mining	Construction	Manufacturing
	(1)	(2)	(3)	(4)
Any Fracking (d)	-0.735	0.320*	0.131***	-0.023
	(0.423)	(0.154)	(0.033)	(0.029)
K-P F-stat	27.6	25.6	27.2	25.6
# Observations	5,986	5,796	5,818	5,898

Notes: This table displays estimates of the effects of fracking activity on house price growth, income, and employment. All columns display IV estimates in which the excluded instruments are shale thickness × # annual permits and depth × # annual permits. House price appreciation (HPA) is measured through the end of the sample period (December 2012). The income measures in Panel B are obtained from the IRS Statistics of Income (SOI), and the employment measures in Panel C from the Bureau of Labor Statistics (BLS). Standard errors are clustered by county×year (*** p< 0.01, ** p< 0.05, * p< 0.1). All specifications include county and year-month fixed effects. The Kleibergen-Paap Wald rank F statistic is displayed for each IV regression. First stage results are available upon request.

Table 8: The Effect of Fracking on Default Rates of Other Debt

Dependent Variables:	90+ Days Delinquent on:				
	1st Mortgage Debt	2nd Mortgage Debt	Credit Card Debt	Auto Debt	Student Debt
[Mean (p.p.)]	[0.068]	[0.039]	[0.281]	[0.057]	[0.080]
	(1)	(2)	(3)	(4)	(5)
Any Fracking (d)	-0.037**	-0.054**	-0.151***	-0.001	-0.019
	(0.013)	(0.022)	(0.050)	(0.016)	(0.037)
K-P F-stat	17.9	14.5	17.6	16.7	14.2
# Observations	8,694,040	1,554,064	7,743,782	4,789,787	1,562,557

Notes: This table displays results for the estimated relationship between various types of consumer debt and fracking activity. The table considers the effect of fracking activity on first mortgage debt (column (1)), second mortgage debt (column (2)), credit card debt (column (3)), auto debt (column (4)), and student debt (column (5)). The underlying data come from the Equifax Credit Risk Insight Servicing McDash Database (CRISM). Standard errors are calculated using two-way clustering by county and year (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). All specifications are instrumental variables regressions that include county and year-month fixed effects. Additional covariates include a quadratic polynomial for the age of the borrower and the borrower's FICO score in the month of mortgage origination. The excluded instruments are shale Thickness $\times$ # annual permits and Depth $\times$ # annual permits. The Kleibergen-Paap Wald rank F statistic is displayed for each IV regression.