

Cunningham, Christopher R.; Gerardi, Kristopher; Shen, Yannan

Working Paper

Fracking and mortgage default

Working Paper, No. 2017-4

Provided in Cooperation with:

Federal Reserve Bank of Atlanta

Suggested Citation: Cunningham, Christopher R.; Gerardi, Kristopher; Shen, Yannan (2017) : Fracking and mortgage default, Working Paper, No. 2017-4, Federal Reserve Bank of Atlanta, Atlanta, GA

This Version is available at:

<https://hdl.handle.net/10419/172917>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Fracking and Mortgage Default

Chris Cunningham, Kristopher Gerardi, and Yannan Shen

Working Paper 2017-4

March 2017

Abstract: This paper finds that increased hydraulic fracturing, or “fracking,” along the Marcellus Formation in Pennsylvania had a significant, negative effect on mortgage credit risk. Controlling for potential endogeneity bias by utilizing the underlying geologic properties of the land as instrumental variables for fracking activity, we find that mortgages originated before the 2007 boom in shale gas, were, post-boom, significantly less likely to default in areas with greater drilling activity. The weight of evidence suggests that the greatest benefit from fracking came from strengthening the labor market, consistent with the double trigger hypothesis of mortgage default. The results also suggest that increased fracking activity raised house prices at the county-level.

JEL classification: Q51, R11, G21

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

For helpful comments and discussions, the authors thank seminar participants at the 2016 Southern Finance Association Conference, seminar attendees at University Catholique de Louvain IRES Macro Lunch, the London School of Economics/SERC, and Maastricht University Finance. The views expressed here are 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, Research Department, 1000 Peachtree Street NE, Atlanta, GA 30309-4470, and University Catholique de Louvain, 3 Place Montesquieu, Box L2.06.01, B-1348 Louvain-la-Neuve, Belgium, chris.cunningham@atl.frb.org; Kris Gerardi, Federal Reserve Bank of Atlanta, Research Department, 1000 Peachtree Street NE, Atlanta, GA 30309-4470, 404-498-8561, kristopher.gerardi@atl.frb.org; or Yannan Shen, Clemson University, College of Business, 170 Sarrine Hall, Clemson, SC 29634, yannan.shen@clemson.edu,

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed’s website at www.frbatlanta.org. Click “Publications” and then “Working Papers.” To receive e-mail notifications about new papers, use frbatlanta.org/forms/subscribe.

1 Introduction

Hydraulic fracturing (fracking), the process by which water, sand and chemicals are injected into certain geological formations to release natural gas and other hydrocarbons, is of considerable concern to environmentalists and homeowners. Since the last quarter of 2007, over 80,000 fracking wells have been permitted and drilled in populated neighborhoods throughout the United States (Figure 1). 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.”¹ Drilling and extracting shale gas presents a number of hazards and dis-amenities that may impose negative externalities on the surrounding population, including ground or surface water contamination, earthquakes and visual blight when industrial equipment replace pastoral scenery. Concerns about the risk of fracking have led a number of countries, states and municipalities to ban the practice.²

Increased fracking activity has also raised important concerns among residual mortgage market participants. Mortgage lenders, borrowers, and policymakers are especially 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 nearby fracking activity, while other homeowners located near fracking wells are unable to obtain house insurance.³

While the negative aspects of hydraulic fracturing have been well-documented, there are

¹Gold, Russell and Tom McGinty, “Energy Boom Puts Wells in America’s Backyards” The Wall Street Journal, October 25, 2013.

²Kaplan, 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; Arenschiold, Lauren, “Fracking industry suing over drilling bans”, The Columbus Dispatch, November 21, 2014.

³See 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/>.

also several potentially positive economic effects that it could bring. For example, fracking activity can generate considerable revenue for landowners and local governments through mineral rents, access fees, and well pad leases. Fracking could also stimulate direct and indirect employment demand, which could increase house prices. Even if fracking does not positively affect home values, a strong labor market may limit default by underwater borrowers (Foote et al. (2008)). Thus, the net effect of fracking on housing markets and mortgage credit risk is uncertain and, ultimately, an empirical question.

In this paper we test the impact of the shale boom on mortgage credit risk in the Pennsylvania housing market. Using detailed, micro data on mortgages originated in the period immediately before the fracking boom (2004–2006) and data on both fracking permits as well as drilling starts, we document a negative relationship between fracking activity and mortgage credit risk. While seemingly straightforward, any analysis of the economic effects of oil and natural gas extraction must deal with potential endogeneity concerns, as the decision to sell mineral rights by property owners, the decision by local governments to permit drilling, and the decision by firms on the precise locations to construct wells may be correlated with the underlying economic characteristics of the region.

To overcome possible endogeneity concerns and establish a causal link between drilling and mortgage default, we use an instrumental variables approach. Our IV approach involves using variation in the underlying geologic properties of the Marcellus Formation and the timing of the Pennsylvania shale boom, which we argue was exogenously determined by innovations in fracking technology. Employing our preferred econometric specification, 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 active fracking reduces the probability of severe delinquency by 0.26 percentage points, on average, which corresponds to approximately 110% of the average monthly delinquency rate in our sample of Pennsylvania mortgages (0.24%), and is roughly equivalent to increasing the borrower’s FICO score at origination from 580 to 700. 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. The mitigating effect of fracking activity on mortgage default appears to be stronger for what were, *ex ante*, more fragile mortgages, and in areas that did not have previous experience with conventional drilling activity. We show that the negative impact on default rates holds for numerous measures of fracking activity, but is especially strong for flow measures of new permits and newly drilled wells. Thus, the weight of evidence suggests that the net effect of the shale boom on Pennsylvania housing markets was—at least initially—positive.

The paper then pivots to an analysis of potential causal mechanisms. We present strong evidence that fracking activity is positively associated with employment gains at the county-level, and that the gains are especially strong in industries that are most likely to be directly impacted by the shale boom, such as drilling/mining and construction. We also present evidence that increased fracking activity leads to moderately higher house price growth, but the correlation with employment is much stronger. This suggests that the fracking boom mitigated mortgage credit risk through its positive effects on labor markets. This is consistent with recent evidence showing that employment loss is a primary determinant of mortgage default (Gerardi et al. (2015)).

There is recent evidence in the literature that shows fracking activity can have negative effects on home values. Using hedonic regression techniques, Muehlenbachs et al. (2015) (hereafter MST) find that proximity to fracking lowered the sale prices of homes with well water, but modestly increased the prices of homes with piped water. While MST is a careful analysis that uses parcel-level data and a credible identification strategy that compares piped versus well water, one important drawback is that it relies on sales to estimate the hedonic parameters, which could introduce potential sample selection bias into the estimation. For example, a prospective home seller might be deterred from completing a sale if she receives

low bids that do not leave enough equity to buy another home or that would require bringing money to the closing (Chan (1996), Ferreira et al. (2010).)⁴. Even homeowners that are not underwater on their mortgages often exhibit loss-aversion –preferring to hold on to a home rather than realize a (nominal) capital loss (Engelhardt (2003).). Furthermore, if a negative externality such as proximity to a drilling pad depresses sales, it could alter the population of sales used to estimate a hedonic model. For example, properties that benefit from the shale boom such as those with revenue generating mineral rights, might sell while their uncompensated neighbors may not. Finally, hedonic methods rely on a careful estimation of implied land rents, however if properties near wells tend to be rural or exurban, the sale price may reflect mostly structure value making unobserved maintenance or innate structure quality a particularly large source of measurement error.

As an alternative to using hedonic or repeat-sales methods that rely on housing sales, in this paper we focus on how the shale gas boom affected mortgage performance in Pennsylvania. We view this question as a complement to MST’s analysis of housing prices, as the decision to default on mortgage debt is related to, but distinct from, choices about the timing and reservation price of a sale. Underwater homeowners subject to a negative price shock may default when the collateral value of the home falls below the mortgage balance as evidenced in Campbell and Dietrich (1983), Deng et al. (2000) and Ambrose et al. (2001). At the same time, determinants of default are of interest in their own right. Did the employment gains associated with the shale boom prevent the “double trigger” (Elul et al. (2010), Foote et al. (2008)) keeping underwater people in their homes? Did uncertainty about the employment market, and the prospect of mineral rents increase the option value of homes that might otherwise have gone into foreclosure?

Studying mortgages in Pennsylvania also offers a few econometric advantages. First, our data encompasses a large fraction of all mortgages in the area, not just properties that sold,

⁴In case the reader is familiar with the debate surrounding negative equity and mobility (Schulhofer-Wohl (2011), Ferreira et al. (2011)) the debate hinged on ability of homeowner to transition to landlords and move, but there was agreement that negative equity delayed sales

which mitigates concerns about sample selection. In addition, we observe the exact timing of default rather than the time that the ultimate sale was recorded. This provides a closer temporal link to drilling activity, which moved quickly during the Pennsylvania fracking boom.

Finally, this study departs from MST by focusing on the effect of drilling at a slightly more aggregated geography. Specifically, we analyze how variation in fracking activity impacts mortgage credit risk, house prices, and employment dynamics at the zip-code level, which is a relatively small area, but allows for wider treatment area from exposure to fracking wells. This yields a better measure of the net benefit or cost of fracking if, for example, the positive spillovers from fracking accrue, if modestly, to the whole region, but the negative spillovers attenuate quickly.

The balance of the paper is as follows: first we briefly review the recent history of hydraulic fracturing and the environmental risks that it entails, before presenting our econometric framework. We then describe the mortgage data and geological instruments before discussing the results. There is a brief conclusion.

2 Overview of Fracking

The extraction of natural gas from tight formation (low-permeability) shale deposits using hydraulic fracturing was first pioneered in the Barnett Shale of north Texas. The process, involves injecting large quantities of water, sand and chemicals under high pressure to create fractures in the shale. Sand or other particulates mixed with the water then become lodged in the newly formed fissures, holding them open, and allowing the hydrocarbons and water trapped inside to migrate up the borehole. Chemical additives can alter the viscosity of injection fluids, preventing corrosion, freezing or the growth of bacteria. While hydraulic fracturing was first pioneered in the late 1940s as a method to revive existing or problematic wells, the higher pressures and volumes (and the pumps necessary to create them) typically

used to release shale gas was not pioneered until the 1990s.

While the basic technology of hydraulic fracturing has been used for decades in places like Texas and California, it did not appear in Pennsylvania until the 2000s. This delay was due to the late emergence of a major innovation that made drilling the Marcellus Formation economically viable: horizontal drilling. In a typical horizontal shale formation, (where there is differential pressure from above), the induced fractures form vertically. Thus, a horizontally drilled well is essential to reach multiple fractures. In addition, the Marcellus Formation is quite deep, often more than a mile beneath the surface. The depth of the shale increases pressure which facilitates gas migration. Furthermore, the simply geometry of horizontal drilling amortizes the fixed cost of the initial bore to reach more shale area. Horizontal drilling was aided, in turn by developments in 3D seismic modeling and the ability to fracture sections of a well in stages. The ability to direct a well horizontally, allows a single well pad to send down multiple wells in different directions, minimizing the cost of ground leases, pad construction and the building of access roads, pipelines, storage tanks and containment ponds.

The Marcellus Formation principally lies under Pennsylvania, but also encompasses part of New York, West Virginia and Ohio. While its location and gas content has been known for some time, it was not thought to be economically viable until the mid-2000s. The first exploratory well in the Marcellus was not drilled until 2002, and economically valuable amounts of gas were not recovered until 2005. Conventional wells drilled through the shale produced relatively little gas. The first commercial wells were drilled in late 2007. Initial signing bonuses (on top of a typical 12.5% royalty rate) were less than \$100 per acre, but would soon exceed \$2,000 for the best properties as wells began producing large quantities of gas in 2008. By 2010 just under 1,400 wells were drilled. Currently ranked as the largest proven wet gas reserves 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 incredibly valuable

in a very short period of time and generated an extraordinary amount of economic activity in the middle of a recession.

2.1 Environmental Risks

Hydraulic fracturing may generate a number of environmental hazards and other dis-amenities that could impact properties near a well. The greatest concern is that hazardous chemicals from the injection fluids, the briny water trapped with the gas, or the gas itself may migrate up the outside of the pile casing and contaminate shallower ground water (Holzman (2011), . Potential contaminants from injection fluids include known or suspected carcinogens including naphthalene, xylene, toluene, ethylbenzene, and formaldehyde. The water released along with gas can include elevated levels of naturally occurring minerals including strontium, barium benzene and radioactive elements including radium 226 (Colborn et al. (2011)). The most likely cause for this contamination is failure to adequately cement the well.

Injection fluids and the co-produced brine that return to the surface as part of the production process must be treated before being released into surface waters, recycled or disposed of. Each method has its own set of risks. Containment ponds can leak or fail, fouling surface or groundwater. Water treatment facilities may not be able to treat or sufficiently dilute hazardous chemicals. If the fracking fluids and co-produced brine, referred to as production water, are disposed of by injecting it into a separate storage well that too risks contaminating ground water. Employing disposal wells has also been associated with increased seismic activity (Ellsworth (2013)).

Methane, the primary molecule in natural gas can leak from wells or pipelines and potentially cause a fire or explosion (Puskar et al. (2015)). More generally the intensity of fracking activity, perhaps because of gas leaks but perhaps from generators and other machinery used in well construction appears to be associated with greater asthma (Rasmussen et al. (2016).) In addition, there is some observed correlation with skin irritaiton (Rabinowitz et al. (2015)).

While these findings suggest that fracking activity may lead to adverse health outcomes, the evidence for a strong causal link between fracking and health is quite limited (Mitka (2012), Werner et al. (2015)). However, even if there is not a substantially elevated direct health risk, the industrial activity involved in creating a well could generate additional trucking volume with its own health risks (Mathews (2015), Graham et al. (2015)).

2.2 Economic Benefits

Despite the potential environmental and health risks described in the previous section, there are a variety of potential economic benefits associated with fracking activities including job creation, income growth, property appreciation, and reduced household energy expenses.

Since 2006, more than 16,000 horizontal fracking wells have been permitted in Pennsylvania. This surge in fracking activity created more than 10,000 jobs in Pennsylvania, which accounts for 20 percent of total statewide employment in the oil and gas industry.⁵ As a result, Pennsylvania went from being the 10th-largest state in terms of oil and natural gas employment in 2007 to being the 6th largest in 2012. The state also had the second-largest employment increase from 2007 to 2012, behind only Texas, which is also a major oil and natural gas producing state. A study conducted by Pennsylvania State University predicts that full development of the Marcellus Shale play in Pennsylvania could support 200,000 jobs.⁶ From 2007 to 2012, the Bureau of Labor Statistics (BLS) reports that employment in the oil and natural gas industry in Pennsylvania increased by 15,114 (259 percent). In addition, wages in Pennsylvania’s oil and natural gas industry rose by \$22,104 (36 percent), to \$82,974 in 2012.⁷

Increases in shale gas extraction activities also creates additional sources of income for local property owners in the form of sign-up bonuses and royalty payments (see Deller and

⁵<http://www.workstats.dli.pa.gov/Documents/Marcellus%20Shale/Marcellus%20Shale%20Update.pdf>

⁶<http://marcelluscoalition.org/wp-content/uploads/2010/05/PA-Marcellus-Updated-Economic-Impacts-5.24.10.3.pdf>

⁷“The Marcellus Shale gas boom in Pennsylvania: employment and wage trends”: <https://www.bls.gov/opub/mlr/2014/article/pdf/the-marcellus-shale-gas-boom-in-pennsylvania.pdf>

Schreiber (2012)). Since fracking shale gas became technically feasible and profitable, property owners in the PA shale gas region have had the option of signing fracking leases with oil and gas companies in exchange for royalty payments of typically 15–25 percent of gas production profits. Pennsylvania law requires a minimum royalty rate of 12.5 percent on all gas leases and some landowners were able to negotiate royalty rates as high as 35 percent.⁸ An average gas lease signing bonus in PA was \$2,400 per acre in 2008, while some signing bonuses have reached \$7,000 per acre in Bradford County, where the shale is thick and fracking is extremely profitable.⁹

Since royalty rights and bonuses are tied to land leases, properties sitting on top of shale gas deposits might also increase in value. For example, Muehlenbachs et al. (2015) found that properties which are exposed to minimal environmental risks from fracking activities appreciate in value when fracking becomes profitable.

Finally, the advancement of fracking techniques also significantly increased oil and natural gas production in the U.S., putting downward pressure on oil and gas prices and making energy more affordable to many households. For example, Hausman and Kellogg (2015) estimate that the U.S. fracking boom is directly responsible for a 47 percent gas price drop. Overall, they estimate that residential consumer gas bills have declined by \$13 billion per year from 2007 to 2013 thanks to the fracking revolution, which amounts to roughly \$200 per year for gas-consuming households.

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 a viable method to extract natural resources as well as the bursting of the housing bubble, which lead to the subprime mortgage foreclosure crisis and subsequent global financial crisis

⁸Act of Jul. 20, 1979, P.L. 183, No. 60: <http://www.legis.state.pa.us/WU01/LI/LI/US/HTM/1979/0/0060..HTM>

⁹“Cash In on the Natural Gas Shale Boom”: <http://www.kiplinger.com/article/business/T019-C000-S002-cash-in-on-the-natural-gas-shale-boom.html>

and recession. 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.¹⁰ For this reason we choose to focus on mortgages originated before 2007 so that our analysis is not contaminated by selection effects related to evolving underwriting standards in response to the emergence of shale drilling activity and the bursting of the housing bubble. Thus, our empirical analysis is focused on quantifying the impact of fracking activity on the default risk of outstanding loans.

We employ a hazard model in most of our empirical analysis, where we relate the monthly hazard of serious delinquency at the individual loan-level to fracking activity in the zip-code. A hazard model allows us to determine if contemporaneous variation in fracking activity affects mortgage default decisions. In addition, a hazard specification naturally accounts for data that is right-censored and also allows for the inclusion of time-varying covariates. Our baseline specification is given by the following linear probability model, where each loan in the sample contributes all monthly observations in which it was active. A loan exits the sample after the first month it becomes 90 days delinquent, is prepaid voluntarily, or is right-censored at the end of 2013. We do not jointly model prepayment and default, so a prepayment is treated identically to a right-censored observation. 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 (Springer and Waller (1993)).

$$Prob(Delinq_{it} = 1) = \alpha + \theta \text{frack}_{zt} + \text{dur}'_{it}\beta_1 + X'_i\beta_2 + \eta_c + \delta_t + \varepsilon_{ict}, \quad (1)$$

where i indexes the individual mortgage, z indexes the zip-code in which each mortgage is

¹⁰The literature has documented how the financial crisis caused a significant tightening of mortgage underwriting standards. For example, the Urban Institute has estimated that approximately tight lending standards resulted in 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>).

originated, t indexes the year-month (in calendar time), and c indicates the county in which the property is located. The term dur_{it} indicates the duration of time since the mortgage was originated and enters as a second-order polynomial. X_i 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. The term $frack_{zt}$ refers to a measure of fracking activity in zip-code z in time period t . Our interest is in 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 use multiple measures of fracking activity in our analysis and also instrument for fracking activity to address potential endogeneity concerns, which we discuss in more detail below.¹¹ Our covariate set, X_{izt} includes detailed mortgage and borrower characteristics at the time of origination, which are typically used by underwriters. These include the origination amount (LOAN AMOUNT), the mortgage interest rate (RATE), loan-to-value (LTV) ratio (at origination), the debt to income ratio (DTI), and the borrower’s FICO score (FICO). We also include dummy variables that indicate whether a loan has the following characteristics: fixed-rate (FRM) or adjustable-rate (ARM), refinance loan (REFI) or purchase loan (PURCHASE), jumbo loan (JUMBO) or conforming loan, 30-year term (TERM30), less than full documentation of income and/or assets (LOW DOC), presence of private mortgage insurance (PMI), an LTV ratio of exactly 80% (LTV80)¹², a prepayment penalty (PENALTY), interest-only payment (IO), and a balloon payment (BALLOON). We also control for whether the mortgage was ultimately retained in the bank portfolio or pooled into a private-label Mortgage Backed Security (MBS) that often had weaker underwriting standards. Mortgages that were packaged into securities with GSE guarantees are

¹¹To address potential spatial and serial correlation, we employ two-way clustering of the standard errors by individual loan ID and county-year. Clustering by loan ID accounts for the fact that we have multiple observations for each loan in the dataset, while clustering by county-year accounts for possible spatial correlation within counties and serial correlation within the calendar year.

¹²Mortgages with an LTV ratio of exactly 80% often had subordinate liens (piggy-back loans) in the pre-crisis period. Since the LPS dataset does not contain any information on subordinate liens, we use this as a proxy.

the omitted category.

3.1 Endogeneity Concerns

The decision to sell mineral rights by landowners or for local governments to permit drill may be endogenous. For example, a struggling community may be more willing to accept drilling in the hopes of attracting employment or tax revenue, whereas a wealthier community may put more weight on health or environmental concerns and withhold drilling rights. Looking across communities we could observe that areas with extensive fracking activity also have higher mortgage delinquency without one causing the other. This could lead to unobserved heterogeneity bias. Blohm et al. (2012) find that 32 percent of the Marcellus Formation is inaccessible because of regulation or current land use.¹³ If local government fails to act, engaged citizens may motivate regional, state or federal agencies to restrict drilling.

In addition, house price declines could themselves encourage fracking. When housing represents a large and un-diversifiable share of a household's total portfolio of wealth, risk-averse homeowners may seek more stringent land use regulation to protect this key asset (Fischel (2001), Saiz (2010)). A decline in house prices, for whatever reason, may make local voters and their representatives more tolerant of the risks posed by fracking, which could introduce simultaneity bias.

The problem with simultaneity in the present context is that, in the absence of a good natural experiment, the only viable econometric strategy is to use geographic proximity to select a suitable control group. If the control and treated populations are similar enough, the assumption of the random assignment of fracking activity becomes less heroic. If, for example, we focus on houses in the same neighborhood then potentially confounding factors like housing demand shocks and zoning regulations can be parsed out, with only variation remaining due to the fact that some houses had wells drilled near them and others did not.

¹³By 2012, 7 local governments in Pennsylvania had banned fracking and 6 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.

Still, there are at least two limitations from this approach. First, even if drilling is effectively exogenous within this carefully selected sample of homes, it is not clear whether the estimated treatment effect has any external validity for more dissimilar housing markets. Sometimes we are fortunate enough to be able to sign the bias. For example, if we think populations with comparatively weak preferences for environmental or health amenities are more likely to tolerate fracking and we still estimate a negative treatment effect from fracking on housing prices, then this is likely a lower bound on the average treatment effect of fracking.

However, there is a second conceptual limitation to using a fine geography method. Certain land uses can generate both positive and negative externalities. If one externality attenuates quickly but the other benefits a much wider area, it can be difficult to estimate the net impact of the treatment. Consider a sewage treatment plant that emits noxious smells. No one would want a plant next to their home, yet the facility may be essential to public health, and clean surface water is a nice amenity for the community as a whole. Using geographic proximity to select the control group is necessarily in tension with the size of the range of the externality. As we expand the bounds of the treatment area, the houses beyond the treatment area become poorer control groups. At the limit, where the externality spans a political jurisdiction, fine geography estimation reverts to naive OLS.

Studying the effect of fracking on housing markets suffers from this obvious tension over how to draw the treatment bounds. While it threatens ground water and clears forest or farm land that would be an obvious concern to immediate neighbors, fracking also generates employment and mineral rents that may spillover to the wider community. Stated more plainly, for a town or county deciding whether or not to ban fracking, observing that house prices fall or foreclosures increase in the immediate vicinity of a well is not a sufficient statistic. Thus, in this paper we utilize a wider geography, the zip code, as our preferred unit of analysis and rely on instrumental variables combined with the technological shock of shale gas fracturing to sever the possible simultaneity between housing markets and fracking wells.

3.2 Identification

Our preferred econometric strategy is to combine the geologic properties of the underlying Marcellus shale formation with the natural experiment coming from innovations in fracking technology to predict fracking activity independent of any housing market or local political economy considerations. Effectively, this is an instrumental difference-in-difference method, a good example of which is [citemoser2014german](#). However, in this paper we adapt it to a hazard model and compare the propensity to default before and after the endogenous timing or intensity of fracking.

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.¹⁴ All else equal, greater shale thickness and greater shale depth leads to higher well production. Figures 2 and 3 display contour maps of the thickness and depth of the Marcellus shale 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, hydraulic fracturing did not become a viable means of extracting oil and gas from the Marcellus Formation until the late 2000s. Therefore, we interact the geologic determinants of fracking productivity in our instrument set with a post-2006 indicator variable. 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 zip-codes on 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-2006 period. Thus, our instrument set is given by:

¹⁴In 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.

$$frack_{zt} = \{thickness_z \times Post2006_t, depth_z \times Post2006_t\} \quad (2)$$

where $Post2006$ is a dummy variable corresponding to the post-2006 period.¹⁵

4 Data

In this section we discuss the two primary sources of data used in this study. We begin with a brief description of fracking measures followed by our mortgage data. As we 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 late-2007. In addition, we focus on mortgage performance through 2012 in order to isolate the impacts of the fracking boom. The impact on mortgage and housing markets from the 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, and will hopefully be the topic of another future analysis.

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 complete permit data for both conventional and fracking wells from 1975 to 2017. The information provided in the dataset includes the unique well identification code (API), the exact longitude and latitude of each well, the type of well (conventional/fracking), and the issuance date of the

¹⁵We 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, $thickness_z \times depth_z \times Post2006_t$ 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.

drilling permit. We use the information on longitude and latitude to compute the zip-code in which each well is located (via the Geographic Information System software ArcGIS). We then match this data on permits and wells to the mortgage data described above at the zip-code-month aggregation level.

Figure 4 displays annual counts of fracking permits and drilled wells in Pennsylvania from 2001–2017. The top panel shows counts for horizontally drilled wells, while the bottom panel displays counts for vertical and directional drilled 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 from the figure 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. Panel A in Figure 4 also supports our choice to focus on the post-2006 sample period when constructing our instrumental variables, as it is clear that horizontal drilling largely did not exist before 2007. In fact, in November 2005, the PA DEP issued the first permit to drill a gas well that utilized both horizontal drilling and hydraulic fracking techniques in Washington County, PA. This exploratory fracking well was then drilled in early 2006 and later plugged by operators due to lack of production. In October 2006, the second horizontal fracking well was permitted and drilled in Fayette County, PA. This proved to be the first producing horizontal fracking well in PA, which signalled the potential of the technology in the state. As a result, the number of fracking permits and wells grew rapidly starting in 2007.

A final notable observation from Figure 4 is the difference in the levels of permits and

spudded 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 a successfully drilled well.

Figure 5 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 3 months of the permit being approved. However, there are also many cases where the lag is significant, as approximately 20 percent of wells in our sample are drilled at least 6 months after the permit is obtained.¹⁶

Table 1 shows the status of all fracking permits issued in the state as of the end of our sample period (January 2017). In total, the PA DEP granted 18,549 fracking permits over the course of our sample period. Approximately 53 percent of the permits were still active at the end of our sample 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 our sample period had been either abandoned, plugged, or temporarily shut down. Finally, almost 39 percent of permits expired prior to the commencement of drilling.¹⁷

Since a significant fraction of fracking permits expire without drilling ever taking place (~39 percent), and the permits that don't expire often take several months to become an active well, measures of fracking activity based on permits and actual drilled wells could diverge significantly. In addition, they are likely to pick up different types of variation in economic activity. For example, measures based on permits are likely to be more correlated contemporaneously with income from leases/signing bonuses as well as with residents' expec-

¹⁶In private discussions with industry experts, we learned that oil and gas companies are motivated to drill within 60 months of permitting because most oil and gas leases expire after 60 months. Thus, oil and gas companies would lose the signing bonuses paid to property owners if the proposed well is not drilled in five years.

¹⁷A 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.

tations about future fracking activity. In contrast, measures based on drilled wells are likely to be more correlated with contemporaneous employment fluctuations and royalty income due to fracking production. For these reasons, we will consider measures of both permits and drilled wells in our empirical analysis below.

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 1 as soon as the first fracking permit is issued in a zip code. “Active Permits” measures the total 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 . “Count of Newly Permitted Wells” measures the number of new permits issued in each month. The final two variables correspond to measures of actual drilling activity. “Cumulative Spudded Wells” is the total number of wells that have ever been drilled in a zip code through the current month t , and “Count of Newly Spudded Wells” is the number of newly drilled wells each month. We discuss the merits of each of these measures below.

4.2 Mortgage Data

The mortgage performance data used in the analysis were obtained from Lender Processing Services (LPS). The LPS dataset covers between 60 and 80 percent of the U.S. mortgage market, and contains detailed information on the characteristics and performance of both purchase-money mortgages and refinance mortgages. It includes mortgages from all segments of the U.S. mortgage market: non-agency securitized loans (PLS); loans purchased and securitized by the GSEs; and loans held in lenders’ portfolios. The LPS 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.

Our primary focus is on a sample of loans originated during the housing boom and before the PA fracking boom in 2007. Specifically, we consider mortgages originated in PA between 2004 and 2006 (inclusive). As mineral rights in Pennsylvania are still closely linked to land ownership, and owners of single-family homes are likely to own the land in the shale gas region, the sample includes only first-lien mortgages on single-family properties. The finest geographic information contained in the LPS dataset is the zip-code corresponding to each mortgaged property.¹⁸ We further restrict our sample to loans that are not located in the Philadelphia or Pittsburgh metropolitan areas. Most fracking wells in PA 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 empirical specifications below we also limit the sample to zip-codes that have not experienced conventional oil or gas drilling, as well as only zip-codes in the Marcellus Formation.

Table 3 presents a portion of the summary statistics used in the analysis. Aside from the fracking measures and the rich set of location and time fixed effects we will employ, most of our covariates are drawn from the mortgage underwriting process. These include the FICO score of the borrower at the time of origination, the LTV ratio, information on whether

¹⁸We 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.

the mortgage was for purchase or refinance of existing debt, and whether it was ultimately insured by a GSE, packaged into a private-label security, or kept within the bank’s portfolio. Column (1) presents mean and standard deviations for some of these variables for our baseline specification that includes all LPS mortgages not in the Philadelphia or Pittsburgh MSAs. Note that 12.6 percent of these mortgages would ultimately become 90+ days delinquent.

In columns (2) and (3) of Table 3 we stratify the sample by zip codes that never had a fracked well and those that would eventually be fracked. First note that less than 10 percent of all mortgages were exposed to fracking within their zip code, 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. The most striking difference between the samples is that given their higher average LTV, fracked homes must have been considerably cheaper than their non-fracked counterparts. While these difference are not particularly large, with the exception of mortgage amount, it is certainly consistent with an endogeneity concern that places or people with more vulnerable mortgages were more willing to embrace the impending fracking boom and that we should have some doubt about the external validity of any specification that treats fracking as exogenous.¹⁹

5 Baseline Results

In Table 4 we present a subset of coefficient estimates for our baseline hazard model of default, in which we gradually add control variables and, in the last two specifications instrument for fracking activity. We start with a simple dummy indicator for whether any fracking has occurred in the zip code, called “any fracking” and regress that on our outcome variable:

¹⁹However, we should note that mortgages in fracking zip codes were somewhat less likely to use low-documentation loans or ARMS, perhaps because these mortgages products were more easily obtained in denser, urban communities.

the first month that a mortgage becomes 90 days delinquent. This variable is zero until the first well is drilled and then one for the balance of the analysis.

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. The coefficient estimate, $\hat{\theta}$, equals -0.0003 and is statistically different from zero at the 5 percent level, implying that the probability of a loan defaulting in a given month-year declines if there is a fracking well in that zip code. In column (2) we include a full set of underwriting variables associated with mortgage i . These include FICO scores, the interest rate at origination, the LTV ratio and the DTI ratio at origination. In addition, indicator variables are included that identify whether the mortgage was an option-ARM, a jumbo loan, had a prepayment penalty, had less than full documentation of income and/or assets, had a balloon payment, and whether the loan ended up in a bank's portfolio, in an agency (GSE) security, or in a private-label security. 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 -0.0008, which suggests that places with observably riskier loans were more likely to allow fracking activity. The likelihood that the true effect of fracking on default risk is zero is now less than one percent.

In column (3) we include calendar year fixed effects and in column (4) we include county fixed effects. Neither appreciably changes 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, which absorbs, so that the effect of fracking activity on default is estimated off of only time-series variation within zip codes. The inclusion of zip

code fixed effects slightly increases the (absolute) magnitude of $\hat{\theta}$ to -0.0009 and is consistent with an interpretation in which communities characterized by riskier mortgages were perhaps also more likely to engage in fracking activities.

In column (7) we instrument for fracking using the variables that we discussed above in section 3.2: shale depth and net thickness of organic matter interacted with a post-2006 indicator variable, *Post2006*. For the instrumental variables specifications, we are forced to revert to county fixed effects, as the 2SLS estimator will not support such a rich set of fixed effects. Instrumenting for θ increases the estimated magnitude of fracking on default by close to a factor of 3 (from -0.0009 to -0.0026). The IV estimate is also highly statistically significant. This result is consistent with endogeneity bias being an important issue in this context, as zip codes characterized by riskier borrowers are more receptive to drilling. Table 5 displays the results from the first stage estimation (column (7)). We include both shale thickness and depth by themselves in both the first and second stages, while we only include the variables interacted with the post-2006 dummy in the first stage. Greater shale thickness and depth in the post-2006 period are strong predictors of increased fracking activity. The Wald F-statistics is greater than 50, which easily exceeds typical thresholds used for weak instruments. In addition, we fail to reject, at standard cut-offs, the over-identification test for whether the first stage residuals are uncorrelated with the errors.

While shale thickness and depth are strong predictors of fracking activity in the post-2006 period, Table 5 shows that they are not statistically significant determinants of fracking before 2007. This comports well with the evidence presented above that showed fracking activity did not begin to pick up in the Marcellus Formation until later in the decade. 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 default.

Finally, in column (8), we display results from an IV specification in which we do not interact the shale variables with the post-2006 dummy variable. The fracking coefficient

estimate declines by almost 50 percent, and is now statistically different from zero at only the 10 percent level. Greater shale thickness and depth still predicts increased fracking activity in the first stage, but the R^2 decreases and the Wald F statistic falls to 20. While the estimated magnitude of the fracking effect falls in this specification, it remains significantly greater than the magnitudes obtained from the OLS specifications.

One explanation for why instrumenting yields larger effects is measurement error, which biases estimates towards zero. The most obvious source of measurement error is that we do not observe the exact location of properties within a zip code. Some properties in a zip code have greater exposure to fracking wells than others, and those borrowers are likely impacted more by fracking activity. However, Timmins et. al. (2006) show that close proximity to a fracking well appears to *lower* house prices, which is a key determinant of default, and so we might expect, a priori, that ameliorating this source of measurement error, would weaken not strengthen the effect of fracking on default. A more consistent explanation is that fracking is not randomly distributed across space and that unobservably more default-prone households or communities are more willing to accept fracking. This conclusion is buttressed by the fact that adding controls for mortgages characteristics (column (2)) increases the magnitude of the fracking coefficient.

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 rigorous OLS specification (column (6) of Table 4), is -0.0009, which implies that mortgages borrowers in zip codes that have at some point experienced fracking activity are 0.09 percentage points less likely to default in a given month compared to borrowers in zip codes without fracking. 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 average default rate, which is a sizeable magnitude.²⁰ The IV estimates, as discussed above, are

²⁰We have also estimated a Cox proportional hazard model and obtained a hazard ratio of 0.79 associated with the “any fracking” variable, which implies that the likelihood of mortgage default is approximately 20

quite a bit larger. The estimate from our most preferred IV specification (-0.0026) is more than 100 percent of the monthly average default rate in our sample.

An alternative way to gauge the economic magnitude of our results is to compare the fracking coefficients with the estimates associated with the FICO score variables. The fracking coefficient from our preferred IV equation, is close in magnitude to the coefficient of the FICO score between 580 and 620 (0.0027). Since we include FICO score as a linear variable as well, the dummy variables corresponding to FICO ranges, which were included to capture any discontinuities arising from underwriting heuristics, can be interpreted as the likelihood of default for a borrower with a FICO score just above the cut-off. Thus, allowing fracking in a zip code reduces the probability of default by about as much as replacing a borrower with a subprime credit score of 580 with a prime credit borrower with a score of 700 (the omitted group is $FICO > 700$).

6 Alternative Fracking Measures and Sub-sample Analysis

In this section we check the robustness of our results to alternative measures of fracking activity and to various sub-samples of interest. These results are presented in Table 6. Each row presents the coefficient estimates from a different measure of fracking activity, while each column denotes a different subset of mortgages restricted either by geography or credit score. The first row contains the estimates for the any fracking variable, which again is zero up to the point that a permit is issued in the zip code and one from that point onward. The second row corresponds to a count of active permits in a zip code, which is our best measure of the count of wells currently or soon to be in production. The third row uses a cumulative measure of permits issued since the beginning of our sample period, which

percent less in fracking zip codes compared to non-fracking zip codes. Those results are unreported due to space considerations, but are available upon request.

captures both active and capped wells. In the fourth row we consider the number of wells in a zip code that have actually been drilled or “spudded,” which is slightly less forward looking than permits, but may better capture economic benefits and environmental risks if some permitted wells are never actually drilled. The final two rows in the table focus on monthly flows of new permits and spudded wells, respectively. These may be the most relevant measures if the economic benefits of fracking come, primarily from signing bonuses or the surge in employment from drilling wells, or alternatively, if the primary disamenity is the noise and blight from constructing the wells. We discuss each of these measures in more detail below in section 7.

Column (1) presents the estimates for the various measures of fracking on the baseline sample. In columns (2)–(4) we restrict the sample to various sub-populations of mortgages in an attempt to construct a better control group for our fracking treatment variable. As detailed in Angrist et al. (1996), all observations in an instrumental variables regression can be decomposed into those that will be treated no matter what, compliers that are impacted by the instrument, potential compliers that are not affected by the instrument, and the never treatable. Without additional assumptions, the external validity of the estimated treatment effect is limited to the second and third groups of potential compliers.

While fracking did not emerge as a viable means of resource extraction in Pennsylvania until the mid-to-late 2000s, there has been conventional oil and gas drilling in the state since the 19th century. In addition, the adoption of fracking was almost certainly spurred by the rise in hydrocarbon prices in the mid-2000s. Thus, it is possible that the ameliorative effect of fracking on mortgage default risk could just be a more general manifestation of the energy boom in the area of the state amenable to drilling. To address this issue, in column (2) we exclude any zip codes that had already experienced conventional drilling. In other words, we attempt to exclude the always treated population of loans from the analysis and focus on areas where the shale boom was a purer economic shock. This reduces our sample of loan-month observations by almost 1.5 million. Doing so reveals that fracking zip codes with no

previous exposure to hydrocarbon drilling actually experienced a greater decline in mortgage delinquency. With the exception of the active and cumulative permits specification (second and third rows), all other fracking measure coefficient estimates are significantly greater in magnitude compared to the estimates from the baseline sample.

There are at least two explanations for this differential impact. The first is that some types of drilling activity, even if not hydraulic fracking, were likely anticipated in areas with existing wells, and thus, the fracking boom may have been less of a positive economic shock in those areas. Homeowners, knowing that there was always some chance of an energy driven recovery in their area, may have been dis-inclined to default even before the first fracked wells came on-line, or having experienced previous energy busts were slightly more likely to walk away when deciding whether to strategically default. Alternatively, households with previous experience with the oil and gas industry may be more aware of the long-term environmental or health risks associated with drilling, which tempered their enthusiasm.

One concern is that space requirements to create and run a drilling pad, and the noise and pollution concerns associated with it, may limit drilling in urban areas. While we do observe suburban fracking activity, it could be in some ways exceptional. If rural areas also had better mortgage performance after the housing collapse, perhaps because they never experienced a run-up in house prices or were less exposed to aggressively underwritten mortgages, we could, despite the instrumental variables setup, be mis-attributing the lower default rate to fracking. Or, in the language of Angrist et al. (1996) we may be pooling the non-compliers with the never treatable. For example, the state of Pennsylvania bans fracking wells within 200 feet of a residence (Blohm et al. (2012)). For this reason, in column (3), we limit the analysis to counties that are not assigned to a metropolitan statistical area. This dramatically cuts our sample (by about 80 percent), but it does not appear to change the estimated effects of fracking, however measured, on the propensity to default. We also note that the first stage becomes more powerful, as evidenced by the Kleibergen-Paap Wald F statistic approximately doubling in magnitude, which is consistent with the idea that

fracking in urban areas may be less responsive to the underlying geology.

We might also worry that the presence of the shale could be spuriously correlated with some other determinants of mortgage default. For example, the shale covers most of the state with the exception of the southeast corner where there is (roughly) an arc 50–100 miles centered between Philadelphia and Baltimore that is shale-free (see Figure 2). Perhaps rural areas on the fringe of the major northeastern cities experienced a speculative run-up and collapse in house prices that made them more prone to default? In column (4) we limit the sample to homes on or near the shale. Specifically, we limit the analysis to counties on or adjacent to (within 25 miles of) shale that had at least 25 feet of net organic matter.²¹ Confining our analysis to areas on or near viable shale again significantly reduces the sample of mortgage-months (to just over 1.8 million), but strengthens the negative correlation between fracking activity and default. For all measures of fracking activity, the coefficients more than double in (absolute) magnitude.

Finally, recalling the significant amount of variation in FICO scores and other mortgage characteristics presented in Table 3, we might worry that shale geology is spuriously correlated with the distribution of vulnerable or poorly underwritten mortgages. In columns (5) and (6) we limit the analysis to mortgages with FICO scores below two conventional underwriting thresholds, 660 and 620 (Keys et al. (2010) and Bubb and Kaufman (2014)). Again we find that despite dramatically shrinking our sample—limiting the analysis to mortgages below 620 leaves just under 600 thousand loan-month observations—the coefficient estimates associated with the fracking variables grow in (absolute) magnitude, suggesting that fracking was more likely to prevent default for the observably riskiest mortgages.

At this point, having found that fracking is negatively associated with default across a wide range of specifications, fracking measures, and sub-samples, we reject the null hypothesis, $H_0 : \theta = 0$ in favor of the alternative $H_a : \theta < 0$; fracking lowered mortgage credit risk in the state of Pennsylvania. We will now dedicate the balance of the paper to trying

²¹From our reading of the trade press this appears to be about the minimum viable depth during the period of analysis.

to distinguish between some of the potential causal mechanisms driving these results.

7 Causal Mechanisms

7.1 Comparing Fracking Measures

We begin by taking a closer look at our various measures of fracking activity to see whether the relative strength of different measures might inform our understanding of the causal mechanisms at work. To do this, we take the coefficient estimates presented in Table 6 and multiply them by a one-standard deviation increase in the fracking measure of interest. We present these results in Table 7. The fracking measures are identical to those in Table 6, with the first row again corresponding to the “any fracking” dummy variable for whether a fracking permit has been issued in the zip code on or before the current month.

Our comparison starts in the second row with the number of currently permitted wells in the zip code. This is, in effect, the total number of permits issued within zip code z up to month t less any expired permits. A permit may expire either because the driller ultimately decided not to construct the well or because an existing well has been capped. A gas company may choose to cap a well if production declines below the variable costs of collecting the gas and treating the effluent or if it expects the price of gas to rise in the future. If the primary deterrent to default is the current or expected flow of royalty payments then this measure would, arguably, be the best proxy (in the absence of measures of well productivity.) With the exception of the sample that excludes areas with conventional wells, the total count of permits is again strongly predictive of fewer defaults. A one standard deviation increase in the number of operating wells is associated with a 0.18 percentage point decline in the likelihood of default in the current month, which is a sizable effect relative to the average, unconditional monthly default rate of 0.24 percent in our sample.

In the third row we consider the cumulative measure of wells ever permitted in the zip code up to month t . As this measure is, effectively, active permits plus a count of wells with

expired permits, we might expect this measure to have a more ambiguous effect. Homes near a capped well do not enjoy any mineral rents but are still exposed to the long run health or environmental risks from the well. However, it too is negatively associated with the probability of default across samples, with the exception of the sample that excludes conventional drilling. A one-standard deviation increase in the cumulative number of permits is associated with a greater reduction in the probability of default than our measure of active permits. This finding is somewhat perplexing. One possibility is that even dormant wells have the potential to be re-opened or even re-fracked, or that the density of ever fracked wells is a better indicator of shale gas potential remaining in the zip code.

In the fourth row we replace the count of cumulative permits with wells that have actually been drilled, or to use an industry term, “spudded” wells. If there is some uncertainty when or if a permitted well will actually be drilled then the physical construction of the well may be a better measure of fracking activity. For example, the flow of royalty payments, which are contingent on production, may provide sufficient income to keep a liquidity-constrained, under-water borrower from defaulting, whereas the issuance of a permit may still require some forward looking behavior by the homeowner. Looking across sub-samples the first stage does not perform quite as well as the permit measures across specifications (fourth row in Table 6, the count of spudded wells is strongly predictive of a lower probability of default and the magnitude of the reduction in default is, on average, similar to the cumulative number of permitted wells.

Finally, in the last two rows, we replace the cumulative measures of permitted and spudded wells with a count of newly permitted and spudded wells, respectively. These measures may better capture variation in employment demand due to fracking activity. While some workers are required to service existing wells, pipelines and water treatment facilities, a much greater number of workers are actually required to construct a well. Both newly permitted wells and newly spudded wells are associated with lower default propensities (Table 6). A one-standard deviation increase in newly permitted wells and newly spudded wells is as-

sociated with a 0.31 and a 0.39 percentage point decline, respectively, in the probability of default for the full sample of mortgages (bottom left corner of Table 7.) These are the largest magnitude declines in the probability of default and is consistent with the hypothesis that labor demand from fracking forestalled the second trigger of default, unemployment. We will investigate this story further in the next two sub-sections. We also note that the magnitude of the reduction in default for a one-standard deviation increase in newly permitted or spudded wells more than doubles when we look at mortgages with low FICO scores. Perhaps the most striking pattern in Table 7 however, is the huge increase in the estimated effect of fracking on mortgage default in the sample of zip codes that have never experienced conventional drilling. For example, a one-standard-deviation increase in the number of spudded wells is associated with a 2 percentage point decrease in the propensity to default.

7.2 Impact of Fracking on Employment and House Prices

In the previous section we observed that the fracking measures that are more closely connected to employment flows appear to generate greater declines in the probability of mortgage default. We now look for this effect directly by seeing whether fracking activity predicts changes in the labor market. We also examine the effect of fracking activity on house prices. These coefficient estimates are presented in Table 8, and each estimate comes from a different econometric specification. Each row denotes a different measure of fracking activity and each column denotes a different dependent variable: either a measure of employment or house prices. The unit of analysis is no longer a loan-month but is instead a county-month, due to the fact that our employment and house price measures are aggregated at the county-level. All specifications include county and year-month fixed effects. The estimated standard errors are clustered by county \times year. We start by presenting results from OLS specifications in Panel A. This re-introduces possible simultaneity bias as areas with weaker labor or housing markets may be more amenable to fracking, but with the likely sign of the bias in mind, the results are fairly compelling. In Panel B we present results from instrumental variables

specifications to address these potential endogeneity concerns. In columns (1) and (7) the dependent variable is the cumulative change in county-level house prices from January 2007 (the approximate peak of the national housing boom) to the current month t . In the rest of the table we focus on employment measures including county-level unemployment rates (columns (2) and (8)) and various employment growth rates broken down by type of industry (columns (3)–(6) and (9)–(12)). Our fracking measures are almost identical to those in Table 8, with the exception being the omission of the measure of cumulative permits issued due to space constraints and the fact that the results for this variable are virtually identical to the measure of cumulative spudded wells.

Our house price measure is a repeat-sales index produced by CoreLogic at the county-level. As this measure requires a certain minimum number of sales to compute it is only available for about 60 percent of our counties (4,340 county-month observations).²² Across our measures of fracking activity, the coefficients in column (1) of Panel A are positive and statistically different from zero. Increased fracking activity is positively associated with house price growth at the county-level. For example, an additional newly drilled (spudded) well is associated with a 0.8 percent increase in cumulative house price growth and an additional permit is associated with a 0.9 percent increase. Moreover, counties characterized by some fracking activity experienced significantly higher house price growth (more than 5 percent on average) than counties with no fracking activity.

In column (2) the dependent variable is the level of the county unemployment rate. The dummy for the presence of fracking activity is associated with an 11 percent decline in the unemployment rate, however the standard error is large and we cannot reject the null hypothesis that the true effect of the any fracking variable is zero. However, our other measures of fracking are all statistically different from zero at the 1 percent level. For example, an additional newly drilled well is associated with a 6 percent reduction in the county-level unemployment rate, on average. In column (3) we explore how our fracking

²²The HPIs do cover more than 90 percent of loan-month observations, highlighting the exclusion of sparsely transacted counties.

measures affect the one-year growth rate in total county employment. Consistent with the previous section, only the flow measures of fracking activity—newly permitted and newly spudded wells—have a statistically significant positive impact on employment growth.

In columns (4)–(6) we consider employment growth of a selection of two digit NAICS industries. In column (4) the dependent variable is employment growth in the “Mining, Quarrying, and Oil and Gas Extraction” industries (NAICS 21). The coefficient estimates associated with the fracking measures are positive and statistically different from zero with the exception of the any fracking dummy variable. In column (5) we consider employment growth in the construction industry (NAICS 23). This variable should reflect employment flows associated with building facilities for the shale industry and may also reflect employment flows related to more indirect fracking effects like surges in housing or office demand. Finally, column (6) shows results for employment growth in the manufacturing sector, which may serve as a rough proxy for the tradeable goods sector, which could be crowded out in a resource boom. None of the coefficients in columns (5) and (6) are statistically significant from zero.

While the IV results presented in Panel B of Table 8 are qualitatively similar to the OLS results, the magnitudes are significantly different. The IV estimates associated with the employment variables are much larger, often by more than an order of magnitude compared to the OLS coefficients. For example, an additional drilled well is estimated to increase county-level employment growth by more than 2 percent according to the IV specification, while in the OLS specification the effect is estimated to be slightly less than 0.2 percent (column (9) vs. column (3)).

The house price results displayed in Panel B are relatively similar in magnitude to the OLS results, but the effects are measured with significantly less precision. This is likely attributable to the geologic instruments not being as strong in the limited sample of counties that have county-level HPIs. We can see this in the table by noting the significantly smaller F-statistics in column (7) compared to columns (8)–(12). The counties that do not have

HPI coverage are the more rural counties that have experienced higher fracking activity on average.

7.3 The Effect of Employment and House Price Growth on Default

We have thus far shown that increased fracking activity lowered mortgage credit risk and also improved labor market conditions and increased housing prices in Pennsylvania. In this section we attempt to directly test whether the causal mechanism behind the reduced default risk was employment gains or home value increases due to fracking. Specifically we isolate variation in employment and house prices that is driven by exogenous variation in fracking activity and test if that variation drives differences in mortgage default. To do this, we regress our indicator for mortgage default on employment and house price variables, which are instrumented for using the geologic variables listed in equation 2 above.

The IV estimates are presented in Table 9. While our instruments performed quite well at predicting fracking activity directly (Tables 4 and 6) and while fracking predicts employment and house prices (Table 8), the instruments are relatively weak in terms of their ability to predict changes in employment or house prices. All but one of the first-stage F statistics in Table 9 is less than 10.²³

Despite the weak instrument issue, the results are still highly suggestive of the underlying causal mechanism. Higher unemployment increases the probability of default (column (1)) while employment growth (column (2)) lowers it. Within industries, employment growth in the oil and gas mining industries lowers default as does growth in construction employment (columns (3) and (4)). Our geologic instruments do not predict changes in manufacturing well, and employment changes in this sector do not appreciably impact the probability of default (column (5)).

²³One possible explanation for this is that in clustering by both county-year and loan-id as part of a loan-month hazard model we have over-estimated the size of the first stage standard errors.

In columns (6)–(8) we test whether exogenous variation in fracking activity mitigates default risk through house price channels. Column (6) considers the cumulative growth of county-level house prices from January 2007, while column (7) instead measures house price growth from the month of mortgage origination. In column (8) we use the county-level HPI, to estimate the probability that a borrower is in a position of negative equity in each month t , and test whether variation in negative equity driven by exogenous variation in fracking potential mitigates default risk. The results in columns (6)–(8) suggest that default risk is not materially affected by house price variation that is driven by exogenous variation in fracking potential. One potential explanation for these negative results is the HPI sample selection issue described, which appears to weaken the geologic instruments.²⁴ Given the inability of the shale measures to predict house price changes or for the endogenous house price measures to predict default, we discount the hypothesis that fracking activity reduced default by increasing the current or expected values of borrowers’ properties.

Finally, in columns (9) and (10) in Table 9 we replicate the unemployment and employment growth specifications using only mortgages in counties with estimated house price indexes from Core-Logic. We do this to determine whether the employment results are being driven by a small number of observations in very rural areas. The estimates are similar both in sign and magnitude to those reported in columns (1) and (2), which suggests that employment gains from fracking activity are lowering mortgage default risk in both rural and urban parts of Pennsylvania.

8 Conclusion

This paper shows that the net effect of the shale gas boom in Pennsylvania was to lower mortgage credit risk. This finding is remarkably robust to the measure of drilling activity

²⁴In addition, we note that the IV specifications that include price as an endogenous variable fail the Hansen-Sargan over-identification test. The probability that the Hansen J statistics are actually χ^2 is less than one percent for the specifications in columns (6) and (7) and is less than 5 percent for the negative equity specification in column (8).

and to the location or fragility of mortgages exposed to the boom. Since the sample of mortgages considered originates from before the first fracking production well was drilled we believe that it is largely free of the types of selection biases that may plague other studies of environmental hazards and energy shocks. Indeed, the effect of drilling on default appears to be greater in zip codes that were not previously exposed to more traditional types of oil or gas drilling. The shale boom also reduced default for the riskiest mortgages based on underwriting criteria like FICO scores. While we are unable to measure the impact on default risk from being in the immediate vicinity of a shale well, using our zip-code aggregate measures we find little evidence in support of the GSE prohibition on insuring mortgages on properties near oil and gas wells.

While the shale boom could have positively impacted housing markets in a number of ways, such as by raising the expected mineral rents from a parcel of land, or by providing much needed liquidity to underwater homeowners in the form of signing bonuses and royalties, the weight of the evidence points to improvements in the labor market associated with the shale boom as the most consequential. There are three pieces of evidence for this conclusion. First, the strong and consistent negative relationship between fracking and default when estimated at the zip code level is somewhat at odds with the fine-geography hedonic estimates of Muehlenbachs et al. (2015), which found, depending on the source of drinking water, large negative house price effects. One way to reconcile these two findings is if the positive effects of fracking have a broader geographic impact than the negative impact of any possible environmental hazards. This broader geography effect is more consistent with a labor demand shock than a mineral wealth shock because the treatment area for gas rents is almost certainly smaller than the footprint for the environmental hazard. To be negatively affected by fracking one needs to be relatively close to a well or at least downstream from one, but to receive a royalty payment the well or drill stem actually has to be on or beneath your property. In addition, the expected flow of payments should have been capitalized into house prices in the Muehlenbachs et al. (2015) paper.

The second piece of evidence is that while most of the fracking measures reduced default across most sub-samples of mortgages, the biggest reduction in default for a standardized increase in activity was associated with the act of drilling the well: the count of newly issued permits or newly spudded wells. If instead, royalty payments and their effect either on the net present value of the home or on the property owner's liquidity was the most important mechanism, then we would expect the total count of active wells to yield the largest impact. Alternatively, if the primary effect of the shale boom was to raise the option value of an underwater mortgage, then we would expect the dummy variable corresponding to any permit activity in the zip code to have the largest impact.

Finally, when we look at the immediate determinants of default from the existing literature—unemployment shocks and house price declines—our geology instruments appear to work better via the employment channel than via house prices. Thus, while this paper is principally about the net spillovers of the shale boom on housing markets it also one of the cleaner tests of the double trigger model of mortgage default. The shale gas boom, driven by the advent of new technology and the slow sedimentation of a former inland sea 270 million years ago created a plausibly exogenous employment shock that limited unemployment and thus, mortgage default.

References

- 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.
- Ambrose, B. W., C. A. Capone Jr, and Y. Deng (2001). Optimal put exercise: an empirical examination of conditions for mortgage foreclosure. *The Journal of Real Estate Finance and Economics* 23(2), 213–234.
- Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association* 91(434), 444–455.
- 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.
- Bubb, R. and A. Kaufman (2014). Securitization and moral hazard: Evidence from credit score cutoff rules. *Journal of Monetary Economics* 63, 1–18.
- Campbell, T. S. and J. K. Dietrich (1983). The determinants of default on insured conventional residential mortgage loans. *The Journal of Finance* 38(5), 1569–1581.
- Chan, S. (1996). Residential mobility and mortgages. *Regional Science and Urban Economics* 26(3), 287–311.
- Colborn, T., C. Kwiatkowski, K. Schultz, and M. Bachran (2011). Natural gas operations from a public health perspective. *Human and ecological risk assessment: An International Journal* 17(5), 1039–1056.
- Deller, S. and A. Schreiber (2012). Mining and community economic growth. *The Review of Regional Studies* 42(2), 121.
- Deng, Y., J. M. Quigley, and R. Order (2000). Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica* 68(2), 275–307.
- Ellsworth, W. L. (2013). Injection-induced earthquakes. *Science* 341(6142), 1225942.
- Elul, R., N. S. Souleles, S. Chomsisengphet, D. Glennon, and R. M. Hunt (2010). What ‘triggers’ mortgage default?
- Engelhardt, G. V. (2003). Nominal loss aversion, housing equity constraints, and household mobility: evidence from the united states. *Journal of Urban Economics* 53(1), 171 – 195.
- Ferreira, F., J. Gyourko, and J. Tracy (2010). Housing busts and household mobility. *Journal of urban Economics* 68(1), 34–45.

- Ferreira, F., J. Gyourko, and J. Tracy (2011). Housing busts and household mobility: An update. Technical report, National Bureau of Economic Research.
- Fischel, W. A. (2001). *The homevoter hypothesis: How home values influence local government taxation, school finance, and land-use policies*. JSTOR.
- Foot, C. L., K. Gerardi, and P. S. Willen (2008). Negative equity and foreclosure: Theory and evidence. *Journal of Urban Economics* 64(2), 234–245.
- Gerardi, K., K. F. Herkenhoff, L. E. Ohanian, and P. S. Willen (2015). Can’t pay or won’t pay? unemployment, negative equity, and strategic default. Technical report, National Bureau of Economic Research.
- Graham, J., J. Irving, X. Tang, S. Sellers, J. Crisp, D. Horwitz, L. Muehlenbachs, A. Krupnick, and D. Carey (2015). Increased traffic accident rates associated with shale gas drilling in pennsylvania. *Accident Analysis & Prevention* 74, 203–209.
- Hausman, C. and R. Kellogg (2015). Welfare and distributional implications of shale gas. Technical report, National Bureau of Economic Research.
- Holzman, D. C. (2011). Methane found in well water near fracking sites. *Environmental health perspectives* 119(7), a289.
- Keys, B. J., T. Mukherjee, A. Seru, and V. Vig (2010). Did securitization lead to lax screening? evidence from subprime loans. *The Quarterly Journal of Economics* 125(1), 307–362.
- Mathews, A. (2015). Modeling atmospheric emissions and calculating mortality rates associated with high volume hydraulic fracturing transportation.
- Mitka, M. (2012). Rigorous evidence slim for determining health risks from natural gas fracking. *JAMA* 307(20), 2135–2136.
- Muehlenbachs, L., E. Spiller, and C. Timmins (2015). The housing market impacts of shale gas development. *The American Economic Review* 105(12), 3633–3659.
- Puskar, J. R. et al. (2015). Fires & explosions in the fracking world—where, why, & how to minimize risks. In *ASSE Professional Development Conference and Exposition*. American Society of Safety Engineers.
- Rabinowitz, P. M., I. B. Slizovskiy, V. Lamers, S. J. Trufan, T. R. Holford, J. D. Dziura, P. N. Peduzzi, M. J. Kane, J. S. Reif, T. R. Weiss, et al. (2015). Proximity to natural gas wells and reported health status: Results of a household survey in washington county, pennsylvania. *Environmental Health Perspectives (Online)* 123(1), 21.
- Rasmussen, S. G., E. L. Ogburn, M. McCormack, J. A. Casey, K. Bandeen-Roche, D. G. Mercer, and B. S. Schwartz (2016). Association between unconventional natural gas development in the marcellus shale and asthma exacerbations. *JAMA Internal Medicine* 176(9), 1334–1343.

- Saiz, A. (2010). The geographic determinants of housing supply. *quarterly Journal of Economics* 125(3).
- Schulhofer-Wohl, S. (2011). Negative equity does not reduce homeowners' mobility. Technical report, National Bureau of Economic Research.
- Springer, T. M. and N. G. Waller (1993). Lender forbearance: evidence from mortgage delinquency patterns. *Real Estate Economics* 21(1), 27–46.
- Werner, A. K., S. Vink, K. Watt, and P. Jagals (2015). Environmental health impacts of unconventional natural gas development: a review of the current strength of evidence. *Science of The Total Environment* 505, 1127–1141.
- Wrightstone, G. (2009). Marcellus shale—geologic controls on production. In *AAPG Annual Convention, Denver, CO, Jun 7-10*.

Figure 1: Fracking Within the US

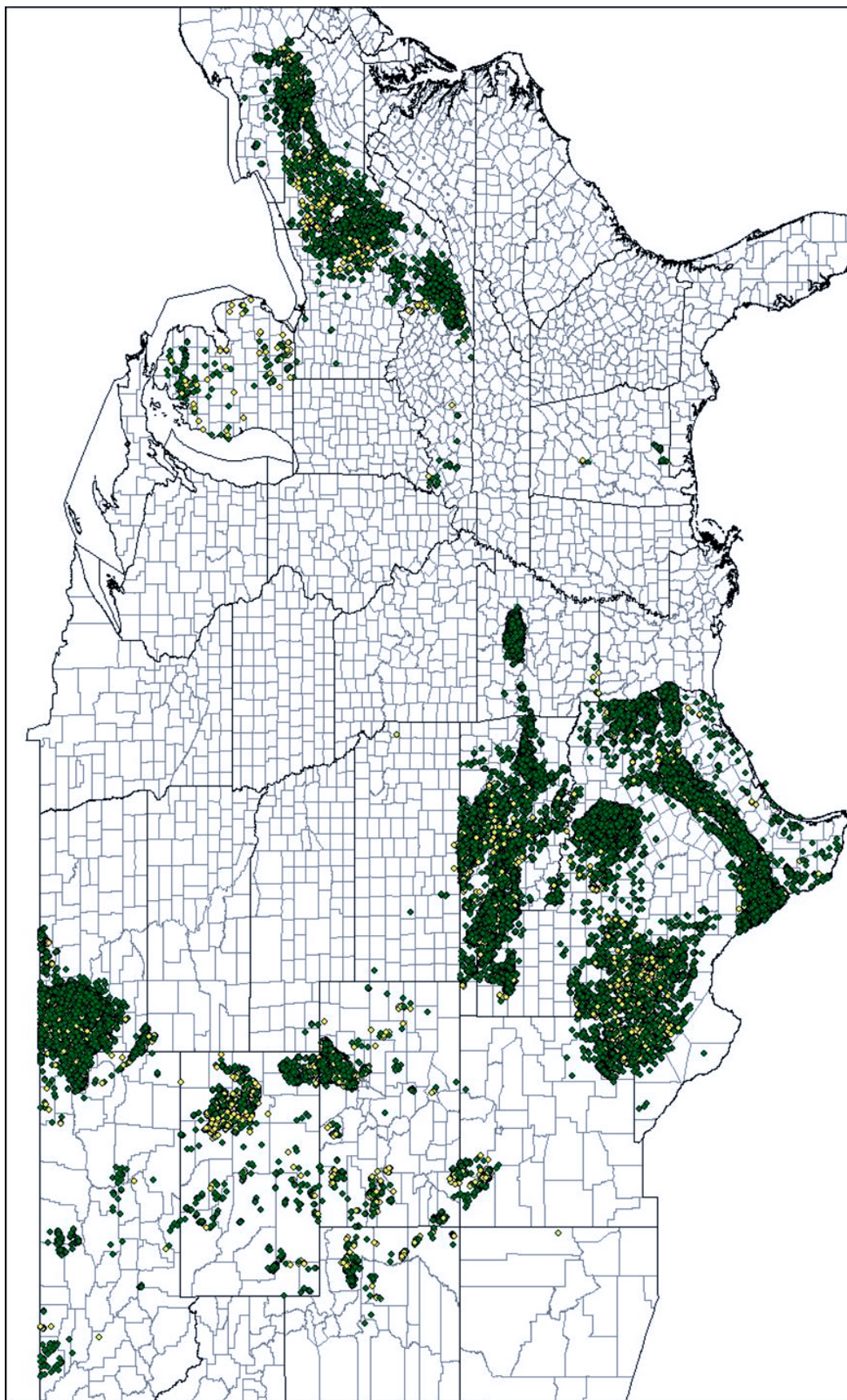
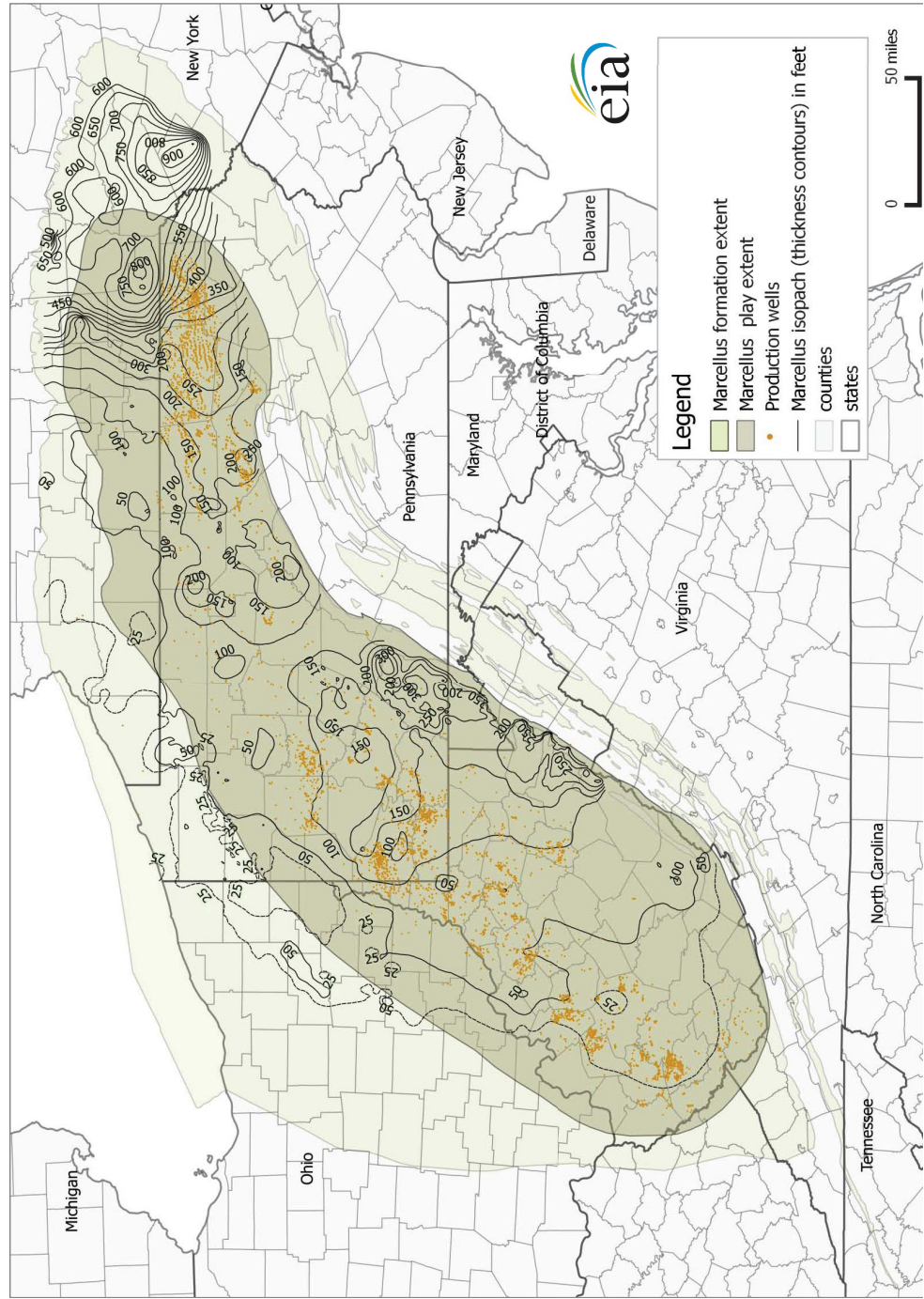
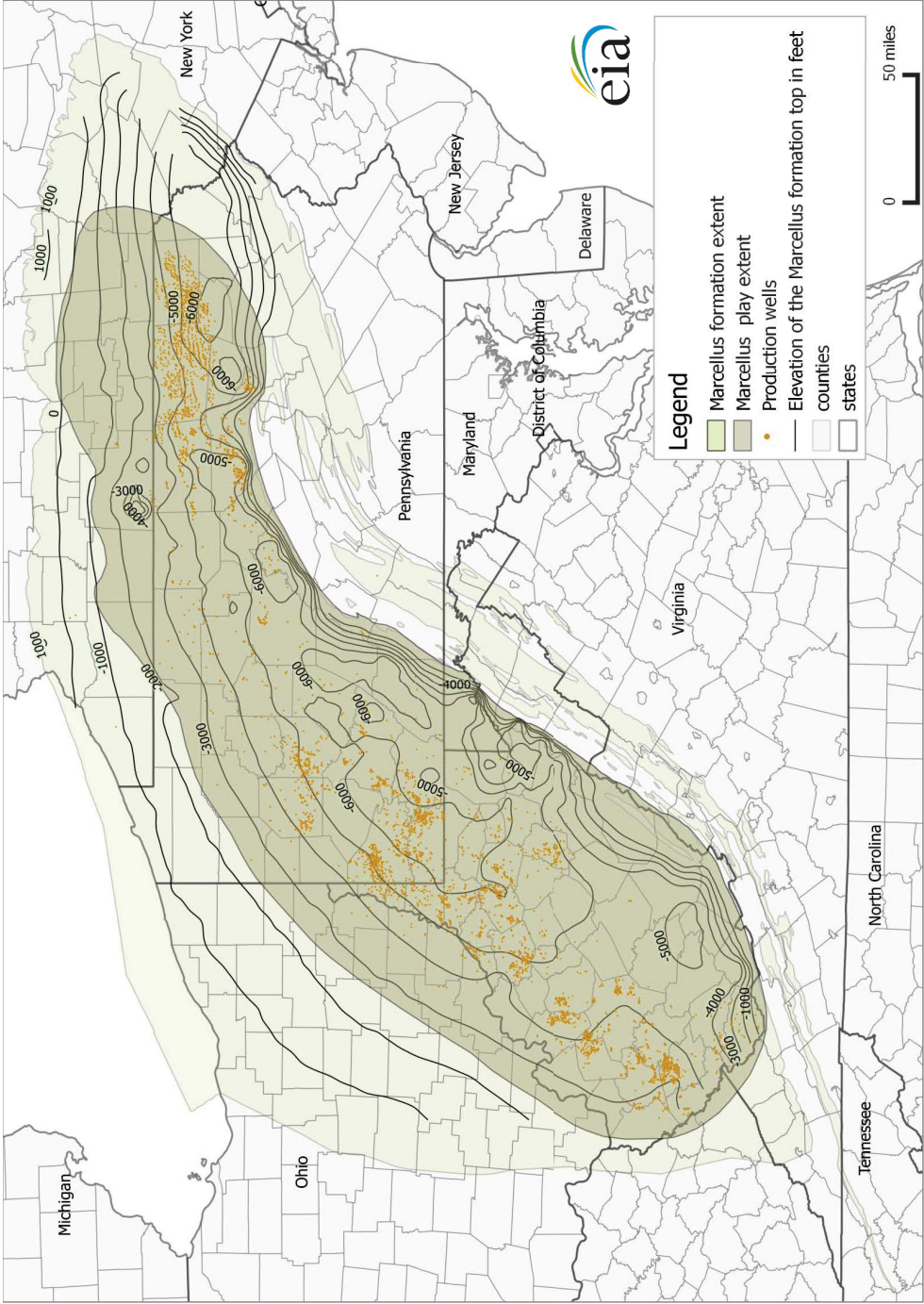


Figure 2: Marcellus Shale Thickness



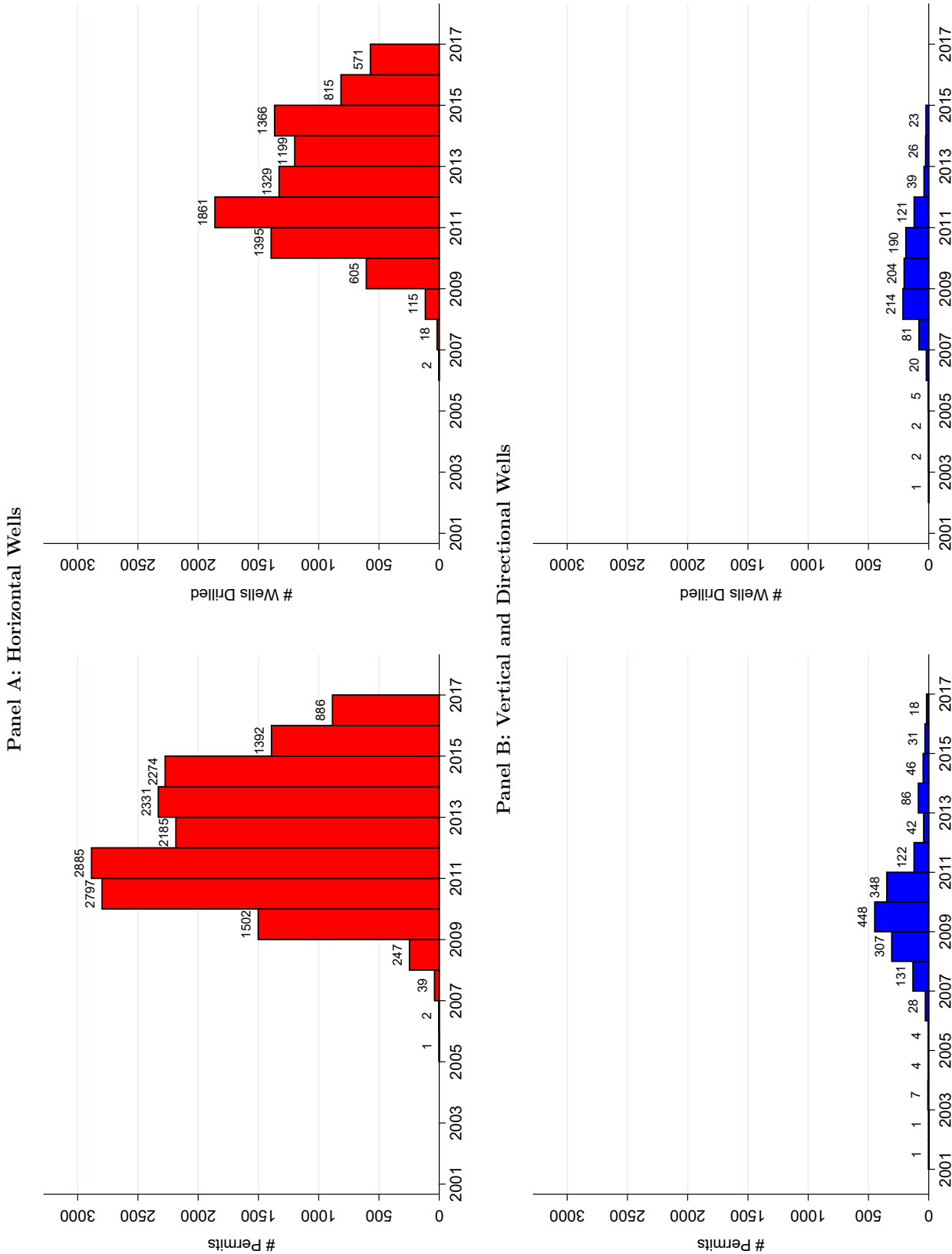
Source: U.S. Energy Information Administration, based on DrillingInfo Inc., New York State Geological Survey, Ohio State Geological Survey, Pennsylvania Bureau of Topographic & Geologic Survey, West Virginia Geological & Economic Survey, and U.S. Geological Survey. The map includes production wells through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

Figure 3: Marcellus Shale Depth



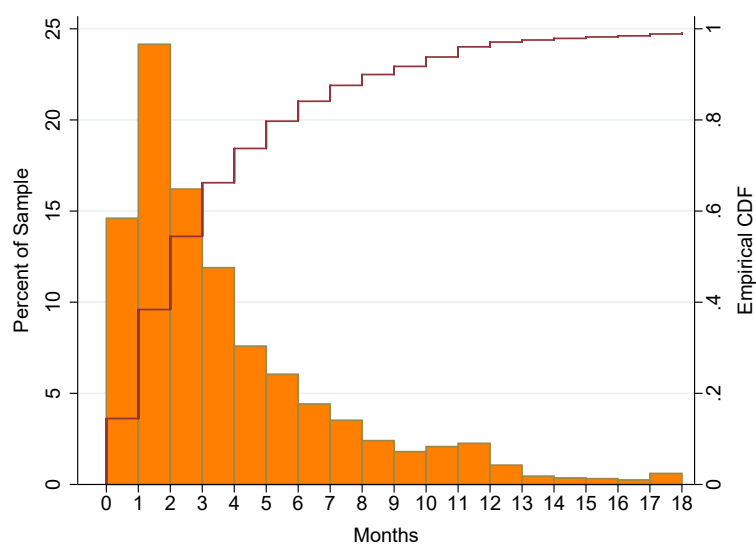
Source: U.S. Energy Information Administration, based on DrillingInfo Inc., New York State Geological Survey, Ohio State Geological Survey, Pennsylvania Bureau of Topographic & Geologic Survey, West Virginia Geological & Economic Survey, and U.S. Geological Survey. The map includes production wells through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

Figure 4: Count of Annual Permits and Newly Drilled Fracking Wells



Notes: This figure displays the annual counts of fracking permits and drilled (or “spudded”) 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 5: Distribution of Time Between Drilling and Permitting



Notes: This figure displays the density and cumulative distribution of the number of months between when a permit is obtained for a fracking well and the well is drilled (or “spudded”). The underlying data were obtained from the Pennsylvania Department of Environmental Protection.

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 | 18,5491 | 100 |

Notes: This table displays the status of all fracking permits issued in PA as of the end of the sample period (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 is drilled and 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; for which equipment necessary for production, extraction or injection has been removed; or considered dry and not equipped for production.

Table 2: Summary Statistics: Fracking Variables

| | Mean | Std. Dev. | Minimum | Median | Maximum |
|---------------------------|------------|-----------|---------|--------|---------|
| Any Fracking (d) | 0.03 | . | 0 | 0 | 1 |
| Active Permits | 0.49 | 6.63 | 0 | 0 | 393 |
| Cumulative Permits Issued | 0.68 | 9.29 | 0 | 0 | 511 |
| Newly Permitted Wells | 0.03 | 0.45 | 0 | 0 | 31 |
| Cumulative Spudded Wells | 0.26 | 4.06 | 0 | 0 | 213 |
| Newly Spudded Wells | 0.01 | 0.24 | 0 | 0 | 23 |
| # Observations | 10,146,844 | | | | |
| # 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: Mortgage Variables

| | Full Sample (1) | Never Fracked (2) | Ever Fracked (3) |
|-----------------------------------|---------------------|----------------------|---------------------|
| Default (90+ days delinquent) (d) | 0.126 (0.33) | 0.126 (0.33) | 0.129 (0.34) |
| FICO \leq 580 (d) | 0.030 (0.17) | 0.030 (0.17) | 0.036 (0.19) |
| 580 < FICO \leq 620 (d) | 0.056 (0.23) | 0.055 (0.23) | 0.067 (0.25) |
| 620 < FICO \leq 660 (d) | 0.137 (0.34) | 0.135 (0.34) | 0.150 (0.36) |
| 660 < FICO \leq 700 (d) | 0.194 (0.4) | 0.194 (0.4) | 0.197 (0.4) |
| FICO | 710 (61.5) | 711 (61.3) | 705 (62.5) |
| DTI | 35.210 (13.82) | 35.228 (13.73) | 35.060 (14.54) |
| LTV | 77.638 (15.29) | 77.448 (15.35) | 79.404 (14.61) |
| Interest rate | 6.166 (0.78) | 6.160 (0.78) | 6.228 (0.78) |
| Mortgage amount (\$) | 139,000 (75,803) | 143,000 (76,437) | 103,000 (58,479) |
| LTV > 80% (d) | 0.129 (0.34) | 0.129 (0.34) | 0.131 (0.34) |
| Low documentation (d) | 0.326 (0.47) | 0.331 (0.47) | 0.283 (0.45) |
| Fixed-rate mortgage (FRM) (d) | 0.877 (0.33) | 0.874 (0.33) | 0.904 (0.29) |
| Balloon (d) | 0.006 (0.08) | 0.006 (0.08) | 0.004 (0.07) |
| Interest-only (d) | 0.037 (0.19) | 0.039 (0.19) | 0.023 (0.15) |
| Private-label (d) | 0.214 (0.41) | 0.214 (0.41) | 0.216 (0.41) |
| Refinanced mortgage (d) | 0.407 (0.49) | 0.405 (0.49) | 0.423 (0.49) |
| Term \neq 30 years (d) | 0.193 (0.4) | 0.190 (0.39) | 0.230 (0.42) |
| # Loans | 194,026 | 175,194 | 18,832 |

Note: This table displays sample summary statistics for most of 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% , option ARMs, 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 | | | | | | | |
|---|-----------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Model: | OLS | OLS | OLS | OLS | OLS | IV shale × post2006 | IV shale |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | | | | | | | (8) |
| Any Fracking (d) | -0.0003** (0.0001) | -0.0008*** (0.0001) | -0.0009*** (0.0001) | -0.0007*** (0.0001) | -0.0007*** (0.0001) | -0.0009*** (0.0001) | -0.0026*** (0.0004) |
| FICO \leq 580 (d) | | 0.0069*** (0.0004) | 0.0069*** (0.0004) | 0.0069*** (0.0004) | 0.0069*** (0.0004) | 0.0070*** (0.0004) | 0.0069*** (0.0004) |
| 580 < FICO \leq 620 (d) | | 0.0027*** (0.0002) | 0.0027*** (0.0002) | 0.0027*** (0.0002) | 0.0027*** (0.0002) | 0.0027*** (0.0002) | 0.0027*** (0.0002) |
| 620 < FICO \leq 660 (d) | | 0.0014*** (0.0001) | 0.0014*** (0.0001) | 0.0014*** (0.0001) | 0.0014*** (0.0001) | 0.0014*** (0.0001) | 0.0014*** (0.0001) |
| 660 < FICO \leq 700 (d) | | 0.0004*** (0.0001) | 0.0004*** (0.0001) | 0.0004*** (0.0001) | 0.0004*** (0.0001) | 0.0004*** (0.0001) | 0.0004*** (0.0001) |
| Low Documentation (d) | | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) |
| Private-Label (d) | | 0.0003*** (0.0001) | 0.0003*** (0.0001) | 0.0003*** (0.0001) | 0.0003*** (0.0001) | 0.0003*** (0.0001) | 0.0003*** (0.0001) |
| Pre-payment Penalty (d) | | 0.0013*** (0.0001) | 0.0013*** (0.0001) | 0.0013*** (0.0001) | 0.0013*** (0.0001) | 0.0013*** (0.0001) | 0.0013*** (0.0001) |
| | | 0.0015*** (0.0002) | 0.0015*** (0.0002) | 0.0015*** (0.0002) | 0.0015*** (0.0002) | 0.0015*** (0.0002) | 0.0015*** (0.0002) |
| Underwriting Covariates | no | yes | yes | yes | yes | yes | yes |
| Calendar Year FE | no | no | yes | yes | no | no | no |
| Year×Month FE | no | no | no | no | yes | yes | yes |
| County FE | no | no | no | yes | yes | no | yes |
| Zip Code FE | no | no | no | yes | yes | yes | no |
| K-P F-stat | | | | | | 51.009 | 21.345 |
| Overidentification $p(\chi^2)$ | | | | | | 0.121 | 0.331 |
| # Observations | 10,146,844 | 10,146,844 | 10,146,844 | 10,146,844 | 10,146,844 | 10,146,844 | 10,146,844 |
| R ² | 0.0003 | 0.0035 | 0.0036 | 0.0036 | 0.0037 | 0.0039 | 0.0037 |

Notes: This table displays results corresponding to the baseline specification (equation (1) in the text). 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 for whether a zip code has experienced any fracking activity during the sample period. 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, option-ARM, prepayment penalty, private-label, portfolio, balloon payment or jumbo. First stage results for columns (7) and (8) are located in Table 5.

Table 5: First Stage Results for 2SLS IV Specifications

| Dependent Variable: Any Fracking | | |
|----------------------------------|-------------------------------------|---------------------------------|
| Model: | IV shale × Post2006 (7) | IV shale (8) |
| Thickness | 0.0003 (0.0002) | 0.001*** (0.0003) |
| Depth | $-6.70e^{-06}$ ($4.73e^{-06}$) | 0.00001*** ($3.66e^{-06}$) |
| Thickness*Post2006 | 0.001*** (0.0002) | |
| Depth*Post2006 | 0.00003*** ($5.50e^{-06}$) | |
| FICO \leq 580 (d) | -0.0001 (0.0001) | -0.0118*** (0.0029) |
| 580 < FICO \leq 620 (d) | 0 (0) | -0.0033 (0.0022) |
| 620 < FICO \leq 660 (d) | -0.0109 (0.0028) | -0.0029* (0.0015) |
| 660 < FICO \leq 700 (d) | -0.0035*** (0.0022) | -0.0014 (0.001) |
| Low Documentation (d) | 0.0012 (0.0008) | 0.0008 (0.0009) |
| Private-Label (d) | -0.0054 (0.0035) | -0.0009 (0.0009) |
| Pre-payment Penalty (d) | -0.0038** (0.0022) | 0.0015 (0.0017) |
| K-P F-stat | 51.01 | 21.35 |
| # Observations | 10,146,844 | 10,146,844 |
| R^2 | 0.397 | 0.327 |

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-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, option-ARM, prepayment penalty, private-label, portfolio, balloon payment or jumbo. The Kleibergen-Paap Wald rank F statistic is displayed for each IV regression.

Table 6: Alternative Fracking Measures and Sub-Sample Analysis

| Dependent Variable: 90+ Days Delinquent | | | | | |
|---|------------------------|------------------------|------------------------|------------------------|-----------------------------|
| Fracking Variables | Full Sample (1) | Never Conventional (2) | Non-Metro (3) | Within 25m Shale (4) | FICO <660 (5) FICO <620 (6) |
| Any Fracking (d) | -0.0026*** (0.0004) | -0.0040*** (0.0011) | -0.0031*** (0.0005) | -0.0059*** (0.0012) | -0.0053*** (0.0013) |
| K-P F-stat | 62.5 | 20.5 | 117.1 | 22.9 | 57.0 |
| | | | | | 47.4 |
| Active Permits | -0.0001** (0.0000) | -0.0002 (0.0004) | -0.0001*** (0.0000) | -0.0002*** (0.0001) | -0.0004** (0.0002) |
| K-P F-stat | 19.7 | 8.9 | 21.3 | 10.5 | 19.9 |
| | | | | | 12.7 |
| Cummulative Permits Issued | -0.0001** (0.0000) | -0.0002 (0.0004) | -0.0001*** (0.0000) | -0.0002*** (0.0001) | -0.0004** (0.0002) |
| K-P F-stat | 19.7 | 8.9 | 21.3 | 10.5 | 19.9 |
| | | | | | 12.7 |
| Cummulative Spudded Wells | -0.0002*** (0.0001) | -0.0014*** (0.0005) | -0.0002*** (0.0001) | -0.0004*** (0.0001) | -0.0005*** (0.0001) |
| K-P F-stat | 17.0 | 7.0 | 5.5 | 9.4 | 14.7 |
| | | | | | 10.8 |
| Newly Permitted Wells | -0.0023*** (0.0005) | -0.0086*** (0.0030) | -0.0023*** (0.0005) | -0.0048*** (0.0012) | -0.0045*** (0.0013) |
| K-P F-stat | 21.5 | 6.8 | 23.1 | 12.2 | 19.0 |
| | | | | | 14.6 |
| Newly Spudded Wells | -0.0054*** (0.0012) | -0.0285*** (0.0094) | -0.0053*** (0.0012) | -0.0099*** (0.0026) | -0.0106*** (0.0031) |
| K-P F-stat | 18.1 | 7.34 | 20.0 | 10.9 | 16.2 |
| | | | | | 11.8 |
| # Observations | 10,146,844 | 8,646,215 | 2,090,264 | 1,837,112 | 1,652,045 |
| | | | | | 571,316 |

Notes: This table displays estimation results for different subsamples and different measures of fracking activity. The dependent variable in all columns is an indicator variable for whether the mortgage is 90 or more days delinquent. 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 county and year-month fixed effects, mortgage age (quadratic), and underwriting characteristics including: 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, option-ARM, prepayment penalty, private-label, portfolio, balloon payment or jumbo. The excluded instruments are shale tickness×post2006 and depth×post2006. The Kleibergen-Paap Wald rank F statistic is displayed for each IV regression.

Table 7: Decline in Mortgage Default from a One-Standard-Deviation Increase in Fracking Activity

| $\hat{\theta} \times \sigma_{fracking}^2$ | Full Sample (1) | Never Conventional (2) | Non-Metro (3) | Within 25m Shale (4) | FICO <660 (5) | FICO <620 (6) |
|---|--------------------|---------------------------|------------------|-------------------------|------------------|------------------|
| Any Fracking (d) | -0.0012 | -0.0019 | -0.0015 | -0.0028 | -0.0025 | -0.0029 |
| Active Permits | -0.0018 | -0.0037 | -0.0018 | -0.0037 | -0.0073 | -0.0147 |
| Cumulative Permits Issued | -0.0026 | -0.0051 | -0.0026 | -0.0051 | -0.0103 | -0.0206 |
| Cumulative Spudded Wells | -0.0025 | -0.0176 | -0.0025 | -0.0050 | -0.0063 | -0.0063 |
| Newly Permitted Wells | -0.0031 | -0.0115 | -0.0031 | -0.0064 | -0.0060 | -0.0064 |
| Newly Spudded Wells | -0.0039 | -0.0204 | -0.0038 | -0.0071 | -0.0076 | -0.0083 |

Notes: This table displays the decrease in mortgage default associated with a one-standard-deviation increase in fracking activity that is predicted from the point estimates displayed in Table 6. Mortgage default is defined to be a loan that is 90 or more days delinquent. The measures of fracking are identical to those in Table 6.

Table 8: Direct Effects of Fracking Activity on Employment and House Price Growth

| Panel A: OLS | | | | | | |
|--------------------------|----------------------------|------------------------|----------------------------|------------------------|---------------------|----------------------|
| Dependent Variables: | HPA Since Jan. 2007 (1) | U.E. Rate (2) | Industry Employment Growth | | | |
| | | | Tot. Emp. (3) | Drilling/Mining (4) | Construction (5) | Manufacturing (6) |
| Any Fracking (d) | 0.0507*** (0.0087) | -0.1164 (0.0710) | 0.0024 (0.0024) | 0.0228 (0.0166) | 0.0128 (0.0091) | -0.0078 (0.0066) |
| R^2 | 0.55 | 0.89 | 0.46 | 0.17 | 0.27 | 0.30 |
| Active Permits | 0.0009*** (0.0003) | -0.0025*** (0.0007) | 0.0000 (0.0000) | 0.0006** (0.0003) | -0.0001 (0.0002) | -0.0000 (0.0001) |
| R^2 | 0.55 | 0.89 | 0.46 | 0.17 | 0.26 | 0.30 |
| Cumulative Spudded Wells | 0.0013*** (0.0003) | -0.0030*** (0.0011) | 0.0000 (0.0000) | 0.0006* (0.0004) | -0.0002 (0.0003) | -0.0000 (0.0001) |
| R^2 | 0.55 | 0.89 | 0.46 | 0.17 | 0.26 | 0.30 |
| Newly Spudded Wells | 0.0078** (0.0034) | -0.0610*** (0.0189) | 0.0018*** (0.0006) | 0.0178*** (0.0050) | 0.0036 (0.0061) | 0.0006 (0.0013) |
| R^2 | 0.55 | 0.89 | 0.46 | 0.17 | 0.26 | 0.30 |
| Newly Permitted Wells | 0.0091*** (0.0024) | -0.0551*** (0.0096) | 0.0013*** (0.0004) | 0.0162*** (0.0044) | 0.0031 (0.0025) | 0.0016 (0.0011) |
| R^2 | 0.55 | 0.89 | 0.46 | 0.17 | 0.27 | 0.30 |
| # Observations | 4,340 | 7,560 | 7,560 | 7,336 | 7,354 | 7,432 |

| Panel B: Instrumental Variables | | | | | | |
|---------------------------------|----------------------------|------------------------|----------------------------|-------------------------|----------------------|-----------------------|
| Dependent Variables: | HPA Since Jan. 2007 (7) | U.E. Rate (8) | Industry Employment Growth | | | |
| | | | Tot. Emp. (9) | Drilling/Mining (10) | Construction (11) | Manufacturing (12) |
| Any Fracking (d) | 0.0298 (0.0305) | -0.5578*** (0.1510) | 0.0123** (0.0047) | 0.0664 (0.0433) | 0.0559** (0.0247) | 0.0128 (0.0188) |
| K-P F-stat | 84.8 | 232.7 | 232.7 | 229.7 | 214.5 | 227.5 |
| Active Permits | 0.0032 (0.0020) | -0.0216*** (0.0062) | 0.0005** (0.0002) | 0.0025 (0.0017) | 0.0022** (0.0010) | 0.0006 (0.00086) |
| K-P F-stat | 10.21 | 33.31 | 33.31 | 31.01 | 30.69 | 30.31 |
| Cumulative Spudded Wells | 0.0051* (0.0031) | -0.0376*** (0.0101) | 0.0008** (0.0003) | 0.0044 (0.0030) | 0.0037** (0.0018) | 0.0009 (0.0013) |
| K-P F-stat | 8.09 | 27.51 | 27.51 | 25.8 | 25.49 | 25.32 |
| Newly Spudded Wells | 0.0247* (0.01392) | -1.0652*** (0.3126) | 0.0238** (0.0094) | 0.1244 (0.0829) | 0.1062** (0.0504) | 0.0290 (0.0390) |
| K-P F-stat | 7.76 | 21.1 | 21.1 | 19.94 | 19.62 | 18.53 |
| Newly Permitted Wells | 0.0977 (0.0652) | -0.5055*** (0.1486) | 0.0117** (0.0046) | 0.0610 (0.0429) | 0.0517** (0.0246) | 0.0159 (0.0195) |
| K-P F-stat | 6.77 | 26.7 | 26.7 | 24.2 | 24.9 | 23.1 |
| # Observations | 4,340 | 7,560 | 7,560 | 7,336 | 7,354 | 7,432 |

Notes: This table displays estimates of the effects of fracking activity on house price growth and employment. The unit of analysis in all regressions is the county-month. Panel A displays OLS estimates while Panel B displays IV estimates in which the excluded instruments are shale tickness \times post2006 and depth \times post2006. House price appreciation (HPA) is measured from January 2007 through the end of the sample period (December 2012). Employment measures are obtained from the BLS. Standard errors are clustered by county \times 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 in Panel B. First stage results are available upon request.

Table 9: Estimated Impact of Employment and House Price Growth on Mortgage Default

| Dependent Variable: 90+ Days Delinquent | | Annual Employment $\Delta\%$ | | | |
|--|-----------------------|------------------------------|----------------------|------------------------|--------------------|
| Endogenous Independent Variable of Interest: | U.E. Rate | Total Employment | Natural Resources | Construction | Manufacturing |
| | (1) | (2) | (3) | (4) | (5) |
| | 0.0033*** (0.0008) | -0.1041** (0.0405) | -0.0123* (0.0065) | -0.0177*** (0.0042) | 0.0091 (0.0064) |
| K-P F-stat | 8.0 | 7.9 | 2.4 | 11.3 | 2.1 |
| # Observations | 10,146,844 | 10,146,844 | 10,081,034 | 10,084,610 | 10,141,235 |
| R^2 | 0.0031 | 0.0030 | 0.0028 | 0.0036 | 0.0035 |

| $\Delta\%$ HPI Since: | | Limit to Counties with HPI | | |
|--|--------------------|----------------------------|-----------------------|----------------------|
| Endogenous Independent Variable of Interest: | Jan-07 | Since Origination | Negative Equity Dummy | Total Employment |
| | (6) | (7) | (8) | (10) |
| | 0.0038 (0.0034) | 0.0042 (0.0053) | -0.0063* (0.0033) | -0.0546* (0.0298) |
| K-P F-stat | 4.8 | 3.1 | 5.3 | 2.6 |
| # Observations | 9,354,140 | 9,354,140 | 9,354,140 | 9,354,140 |
| R^2 | 0.0037 | 0.0037 | 0.0027 | 0.0034 |

Notes: This table displays estimates of the effects of house price growth and employment on mortgage default risk. The unit of analysis in all regressions is the loan-month. In all specifications the endogenous, independent variable of interest is instrumented for using shale tickness \times post2006 and depth \times post2006. Standard errors are calculated using two-way clustering by county \times year, and loan ID (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). All specifications include county and year-month fixed effects and underwriting characteristics: year of origination, Winorized FICO score and dummies for common thresholds, mortgage rate, DTI, LTV, and dummies for whether LTV > 80, fixed rate, 30 year term, refinanced, option-ARM, prepayment penalty, private label, portfolio, balloon payment or jumbo. The Kleibergen-Paap Wald rank F statistic is displayed for each IV regression.