

# Measuring Mortgage Credit Availability: A Non-Parametric Frontier Approach

\*Preliminary and Incomplete, Do Not Cite\* \*

Elliot Anenberg<sup>†</sup>    Aurel Hizmo<sup>‡</sup>    Edward Kung<sup>§</sup>  
Raven Molloy<sup>¶</sup>

August 24, 2015

## Abstract

We construct a new measure of mortgage credit availability from the early 2000s to the present. Borrowing from the literature on frontier estimation, we estimate the features of the mortgage contract that extends the maximum amount of credit to a borrower with a fixed set of characteristics. We argue that changes in this mortgage credit frontier reflect changes in credit supply rather than changes in demand for credit, whereas many existing measure of credit availability are affected by demand. An additional advantage of our framework is that it allows us to examine changes in availability for different types of borrowers and in different housing markets, dimensions that have not yet been explored. In a panel of 80 metropolitan areas, we study how changes in our measure for credit availability affect housing construction and house prices.

---

\*The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors.

<sup>†</sup>Board of Governors of the Federal Reserve System

<sup>‡</sup>NYU Stern

<sup>§</sup>UCLA

<sup>¶</sup>Board of Governors of the Federal Reserve System

# 1 Introduction

The Great Recession of 2008 demonstrated in a striking fashion the important linkages between mortgage markets, housing markets, and the macroeconomy. Since then, there has been a renewed push to understand the role of mortgage credit supply in housing markets and other economic outcomes, including among policy makers, as the Federal government has become increasingly involved in regulating the mortgage market.

A common narrative associated with the housing market over the past decade is that mortgage credit supply was loose in the years leading up to the crisis and subsequently tightened, contributing to the boom and bust in housing prices and activity over this same period. The price of mortgage credit, or the interest rate, declined steadily during this time period, which runs counter to the narrative that credit supply loosened and then tightened. However, in addition to the interest rate, mortgage credit supply is a function of *mortgage credit availability*, or the quantity of credit that a borrower with a particular set of characteristics can obtain, which may have moved up and down over the past decade even as interest rates fell.<sup>1</sup> While several existing papers have found convincing ways to measure exogenous changes in interest rates and then estimate relationships of interest – e.g. the elasticity of house prices with respect to mortgage rates<sup>2</sup>, the analogous literature that focuses on changes in mortgage credit availability is scant despite the common perception that it played a prominent role in explaining the recent boom and bust episode.<sup>3</sup> We are not aware of any estimates, for example, of the elasticity of house prices

---

<sup>1</sup>Mortgage credit availability is not necessarily independent of interest rates. For example, a decrease in interest rates might provide incentives for lenders to expand the quantity of credit that they offer.

<sup>2</sup>See Adelino et al. (2012), Kung (2015).

<sup>3</sup>Several studies present evidence that certain elements of mortgage credit availability loosened during the 2000's (e.g. Mian and Sufi (2009) and Keys et al. (2010) provide evidence that lending standards to subprime borrowers loosened), but we are not aware of any studies that consider a comprehensive, well-defined measure of mortgage credit availability and estimate its relationship with housing market outcomes of interest. Glaeser et al. (2010) find that higher approval rates and relaxed LTV ratios cannot explain much of the recent home price boom, but they acknowledge considerable concerns regarding endogeneity and measurement error in their proxies for credit availability.

with respect to mortgage credit availability<sup>4</sup>; in fact, there are only a couple recent studies that take on the challenge of how to properly measure mortgage credit availability in the first place.

As noted by Glaeser et al. (2010) and Li and Goodman (2014), a major challenge facing researchers studying mortgage credit availability is that there are few direct measures that can accurately disentangle mortgage credit supply conditions from mortgage demand. Consider, for example, one commonly used measure of mortgage availability: the mortgage approval rate. The approval rate of loan applications would seem to be a natural measure of underlying credit availability (if credit is harder to obtain we should see a higher rate of denial) but it suffers from a selection bias driven by the endogeneity of the borrower's decision of what loan to apply for (potential borrowers are unlikely to apply for loans for which they are likely to be rejected). This selection bias can result in counter-intuitive patterns in the data. For example, Figure 1(a) shows that the mortgage approval rate in the U.S. actually went down significantly from 2000 to 2006, which is counter to the evidence that mortgage credit expanded during this period (see Mian and Sufi (2009)).

Consider now an alternative measure, the median borrower credit score of originated loans. The motivation for this measure is that if lenders became more discriminating and credit was harder to obtain, then only individuals who are observably less risky would be able to obtain it. There are two problems with this measure. First, it is a single-dimensional measure of credit constraint. For example, median borrower credit score could be higher while underlying credit constraints remained constant if the median loan was riskier along other characteristics, such as LTV. A second problem is that the median credit score of borrowers could conflate demand and supply side factors. For example, median credit scores could be higher without a change to underlying credit conditions if there was an increase in housing demand among high-credit borrowers for reasons unrelated to credit supply. Figure 1(b) shows

---

<sup>4</sup>Recently, Favara and Imbs (2015) estimate the total effect of mortgage credit supply on house prices, but they are not able to isolate the effect due to mortgage credit availability. In a unique approach, Fuster and Zafar (2015) using survey evidence to estimate the elasticity of willingness to pay for a house with respect to changes in the downpayment requirement.

that the median borrower credit scores did not move much between 2000 to 2006. Again, this would indicate that credit did not loosen over this period, which is against our intuition.<sup>5</sup>

In this paper, we propose a new measure of mortgage credit availability that does not suffer from the above concerns, and has a number of additional advantages. We then use our new measure to estimate of the elasticity of house prices and housing construction with respect to mortgage credit availability.

Our method is motivated by the literature on estimating production frontiers. In the mortgage context, we can think of the output of the production process as being the characteristics of the loan that a lender is willing to underwrite, such as loan amount, and the inputs as being borrower characteristics such as credit score and income. We interpret changes in the loan frontier as changes in mortgage availability. The rationale is that for any given set of borrower characteristics, it seems reasonable to assume that at least a small number of potential borrowers would demand a high loan amount. Therefore, changes in the loan frontier reflect lender policy, not borrower demand.

Cazals et al. (2002) develop an approach for estimating a frontier nonparametrically using data on realized production outcomes. In our application, this means using mortgage originations to estimate the maximum amount of credit that a lender would be willing to extend. We apply this estimation approach using originations for home purchase mortgages from two large, commonly-used datasets, which combined cover prime and subprime mortgages, to create our new measure of mortgage availability, the "loan amount frontier". The

---

<sup>5</sup>Recognizing these limitations, two other measures of credit availability are commonly considered. The first is the Federal Reserve's Senior Loan Officer Opinion Survey (SLOOS), a survey of senior loan officers at 60 of the nation's largest commercial banks. In the survey, respondents are asked about whether their institution "tightened" standards on mortgage lending. As such, the SLOOS indicator depends on the respondent's judgment and interpretation of the survey question, and so is too qualitative to be useful for many applications, among other issues. The second commonly used measure is the Mortgage Banker Association's Mortgage Credit Availability Index (MCAI). The MCAI is based on lenders' stated willingness to provide loans of various types to various types of borrowers. While the MCAI has some advantages, its history is limited, which complicates the interpretation of the current index level and limits its usefulness in economic analysis. Furthermore, the data used to compute the index are only available for loans purchased by investors, not loans held in portfolio by banks, and so it may not give a complete picture of aggregate credit conditions.

interpretation of the frontier is the set of loan amounts that a borrower can obtain given her FICO score, downpayment amount, and income. Like the other measures of mortgage availability discussed earlier, we focus on measuring the quantity of mortgage credit available, and not the price of mortgage credit, but the methodology we develop is general and could be easily applied to measure the price of credit.

There are a number of advantages to our new measure. First, our methodology is entirely data driven and relies on minimal assumptions about the data generating process for mortgage originations.<sup>6</sup> Second, because it uses data on mortgage originations, which are typically available back to the early 2000's, we are able to compare current mortgage availability conditions to the pre-crisis period. A third advantage of our measure is that it is disaggregated and so it can be used to examine the heterogeneity in credit conditions that potential buyers face. We are not aware of any other measure that could report, for example, that credit access is loose for one segment of the population, but tight for another.

Using two large, comprehensive datasets on mortgage originations, we compute our measure for 80 large U.S. metropolitan areas. We find substantial differences in the loan frontier across locations and borrower types. For many borrowers with low credit scores and low downpayments, the loan frontier is zero, indicating that these borrowers are unable to access mortgage credit at all. In most cities, for a given income the slope of the frontier is steepest at low to mid fico scores and downpayment levels, and then flattens out somewhat at higher downpayment and fico levels. The loan frontier is also generally increasing in income. This result supports our claim that the loan frontier is determined by credit supply rather than credit demand. If the frontiers were driven by demand, one might expect higher income borrowers to be associated with lower frontiers conditional on FICO and downpayment, as higher income households tend to be wealthier and would not want or need to lever

---

<sup>6</sup>In contrast to the MCAI as well as other measures that use origination data, our measure does not use data on mortgage performance and so it does not rely on any assumptions that future mortgage performance will be similar to past performance.

themselves as much as poorer individuals.

We also characterize aggregate mortgage availability conditions at the metro level by assigning weights to each borrower type and integrating under the loan frontier. For most metropolitan areas, we find that mortgage availability increased steadily from 2001 to 2006 and fell below 2001 levels from 2006 to 2011. Recent experience has varied more across locations, with availability expanding somewhat in some metropolitan areas, but contracting further in others. Because the time-series pattern is broadly similar across metropolitan areas, areas with tighter credit in 2001 tended to still have tighter credit in 2014. However, the contraction in mortgage credit supply was more pronounced in locations that started out with tighter standards, leading to an increase in the dispersion across locations over this period.

To close the paper, we illustrate the economic value of the loan frontier by showing that it offers new insights on the relationship between credit supply, house prices, and construction activity, a topic of active research over the last several years (see, for example, Favara and Imbs (2015), Adelino et al. (2012), Kung (2015)). Exploiting variation in the loan frontier across metropolitan areas and over time, we find that an increase in the loan frontier for above-average FICO borrowers leads to a significant increase in both construction and prices, but that increases in the frontier for below-average FICO borrowers have little effect on these variables. Thus, the housing cycle appears to be much more sensitive to credit availability for less-risky borrowers than to availability for risky borrowers. We show that alternative, naive measures of credit availability give the researcher a very different impression of the relationship between credit availability and housing activity than the one suggested by the loan frontier.

To address the potential endogeneity of the loan frontier, we instrument for the loan frontier using a Bartik-style instrument. The main identification idea is to use the fact that shocks to the national credit markets are exogenous to the local conditions in one particular metropolitan area, and that a shock to national credit markets will have different effects on different metropolitan areas depending on the distribution of potential borrowers in the area. We are

able to create this Bartik-style instrument because, unlike other measures of credit availability, the loan frontier describes credit availability conditions for different types of borrowers. Using our instrument, we find a strong positive effect of the loan frontier on house prices and construction. Taken together, our results illustrate the value of using a measure of mortgage availability that (i) highlights the heterogeneity in conditions across borrowers and (ii) reflects mortgage supply rather than demand conditions.

## 2 Methodology

Consider a mortgage origination process in which a borrower of characteristics  $x \in \mathbb{R}^p$  (i.e. credit score, income) obtains a loan of characteristics  $y \in \mathbb{R}^q$  (i.e. interest rate, loan amount). The set of possible mortgage originations is given by:

$$\Psi = \left\{ (x, y) \in \mathbb{R}^{p+q} \mid \text{Borrower } x \text{ can obtain loan } y \right\}$$

We assume an ordinal ranking for  $x$  and  $y$  such that that if  $(x, y) \in \Psi$  then  $x' \geq x$  and  $y' \leq y$  implies  $(x', y') \in \Psi$ , where the inequality is taken element by element. In words, a borrower with better characteristics can always obtain all the loans available to a borrower with worse characteristics. Similarly, if a borrower could obtain a loan with good characteristics, then the same borrower could also obtain a loan with worse characteristics.<sup>7</sup>

Formulated in this way, the mortgage origination process is equivalent to a production process with free disposal in which the borrower characteristics are inputs and the loan characteristics are outputs. The econometric problem is to estimate  $\Psi$  from a random sample of mortgage originations,  $\{x_i, y_i\}_{i=1}^n$ . Cazals et al. (2002) (henceforth CFS) describe a robust non-parametric approach to this problem, which we adopt in this paper.

To illustrate the CFS method, we begin with the case of a single output  $y \in \mathbb{R}$  and multiple inputs  $x \in \mathbb{R}^p$ . We note that the possibility set  $\Psi$  can

---

<sup>7</sup>For loan characteristics where smaller is better (i.e. the interest rate), we can simply redefine  $y$  as measuring the negative of that characteristic. We can also do this with borrower characteristics where smaller is better, such as other debt holdings.

equivalently be described by the efficient output frontier  $\varphi(x)$ , defined as:

$$\varphi(x) = \sup \{y \mid (x, y) \in \Psi\}$$

Suppose the data,  $\{x_i, y_i\}_{i=1}^n$ , are drawn from the joint distribution  $(X, Y)$ . Let us define the expected maximum output function of order  $m$ ,  $\varphi_m(x)$ , as:

$$\varphi_m(x) = E[\max\{Y_1, \dots, Y_m\} \mid X \leq x]$$

Intuitively,  $\varphi_m(x)$  is the highest expected level of output that would be observed with inputs less than  $x$ , out of  $m$  draws. CFS show that  $\varphi_m(x) \rightarrow \varphi(x)$  as  $m \rightarrow \infty$ . The construction of  $\varphi_m(x)$  is therefore useful because it approaches the efficient output frontier as  $m$  grows large, and also because it has an easy to compute finite sample analog.

Following CFS, let us construct:

$$\hat{S}_{c,n}(y|x) = \frac{\frac{1}{n} \sum_{i=1}^n I[y_i \leq y, x_i \leq x]}{\frac{1}{n} \sum_{i=1}^n 1[x_i \leq x]}$$

which is the empirical analog of  $P(Y \leq y \mid X \leq x)$ . Noting that:

$$P(\max\{Y_1, \dots, Y_m\} \leq y \mid X \leq x) = P(Y \leq y \mid X \leq x)^m$$

we can compute the empirical analog of  $\varphi_m(x)$  by the following procedure. First, let  $n(x)$  be the number of observations with  $x_i \leq x$ . Then, denote by  $y_j^x$  the  $j$ th smallest value of  $y_i$  that is observed with  $x_i \leq x$ ; i.e. for  $x_i \leq x$  we have  $y_1^x < y_2^x < \dots < y_{n(x)}^x$ . Then, we compute:

$$\hat{\varphi}_{m,n}(x) = \hat{S}_{c,n}(y_1^x|x)^m y_1^x + \sum_{j=2}^{n(x)} [\hat{S}_{c,n}(y_j^x|x)^m - \hat{S}_{c,n}(y_{j-1}^x|x)^m] y_j^x$$

CFS establish the asymptotic properties of the estimator, but the key point to note is that  $\hat{\varphi}_{m,n}(x)$  is a  $\sqrt{n}$ -consistent estimator for  $\varphi_m(x)$ . Therefore, as  $m$  and  $n$  grow large,  $\hat{\varphi}_{m,n}(x)$  approaches  $\varphi(x)$ , the efficient output frontier.

The reason to use a finite  $m$  is that choosing a smaller  $m$  makes the estimator robust to outliers that may actually fall outside the possibility set (i.e. due to measurement error) while still maintaining the interpretation as an expected minimum out of  $m$  draws.

$\hat{\varphi}_{m,n}(x)$  is therefore a consistent estimator of the maximum level of output that inputs  $x$  could achieve. To illustrate further how this frontier can be interpreted, consider an application where the output is loan amount and the inputs are the borrower's credit score and income.  $\hat{\varphi}_{m,n}(x)$  is therefore interpreted as the highest loan amount that a borrower with credit score and income  $x$  could obtain.

Of course, there could be many other relevant outputs and inputs in the mortgage origination process, such as the interest rate of the loan and the appraisal value of the collateral. If these inputs and outputs are ignored in the computation of  $\hat{\varphi}_{m,n}(x)$ , then  $\hat{\varphi}_{m,n}(x)$  measures the maximum loan amount that could be obtained by a borrower with credit score and income  $x$ , *irrespective of interest rate and appraisal value*. So if the maximum loan amount obtainable is increasing in the appraisal value, then  $\hat{\varphi}_{m,n}(x)$  is not representative of the average borrower; rather it measures the maximum loan amount obtainable by borrowers with the highest appraisal values. Moreover, if appraisal value is correlated with income, then  $\hat{\varphi}_{m,n}(x)$  is not a structural estimate of the effect of income on maximum borrowing amount; rather it conflates the effects of both income and appraisal value. Whether or not this omission is problematic depends on the application.

Setting aside for now the problem of omitted variables, suppose we know the distribution of characteristics over the population of potential borrowers,  $F(x)$ .<sup>8</sup> We can then compute the expected maximum output over the popu-

---

<sup>8</sup>It is important that we know the distribution of potential borrowers, as opposed to just the distribution of  $x_i$ 's observed in the data. The reason is that the  $x_i$ 's in the data are a selected sample of individuals who were willing and able to borrow. In times of tight credit availability, it is possible that many individuals are unable to borrow at all, and it is important not to exclude them when constructing a measure of credit access.

lation of potential borrowers as:

$$\hat{\psi} = \int \hat{\varphi}_{m,n}(x) dF(x)$$

$\hat{\psi}$  is therefore an aggregate measure of mortgage credit availability. In the above example with loan amount as output and credit score and income as inputs, it has the clear interpretation as the maximum loan amount obtainable by individuals with average credit score and income (but possibly high levels for other omitted inputs).

So far we have limited our discussion to the case of a single output and multiple inputs. To extend the method to multiple outputs (i.e. both loan amount and interest rate are outputs), one simply notes that there is no special distinction between inputs and outputs other than in the assumption:  $(x, y) \in \Psi$  implies  $(x', y') \in \Psi$  if  $x' \geq x$  and  $y' \leq y$ . If one were to take the negative of an output, it would be interpreted as an input according to the above definition. Therefore, one can estimate the efficient frontier for a single output as a function of the all the inputs and other outputs, simply by recasting the other outputs as negative inputs. To illustrate, let the outputs be  $y^* = (y, z)$  where  $y$  is a scalar and  $z$  is a vector, and let the inputs be  $x$ . Then, the estimator  $\hat{\varphi}_{m,n}(-z, x)$  is a consistent estimator for the maximum level of  $y$  obtainable when inputs are less than  $x$  and non- $y$  outputs are greater than  $z$ .

Because outputs are loan characteristics—and are therefore choices to the borrowers—there is no exogenous population distribution of loan characteristics the way there is for borrower characteristics. To construct an aggregate measure of credit availability, one can simply compute the volume of outputs available at input  $x$ , and then integrate over the population distribution of  $x$ . The aggregate measure is thus computed as:

$$\hat{\psi} = \int \left[ \int \hat{\varphi}(-z, x) dz \right] dF(x)$$

## 2.1 Example

To aid the reader’s understanding of the methodology, consider the example in Figure 2, which shows a frontier using loan amount as the output and the borrower’s FICO score as the input. The dots represent individual mortgage originations and the solid line is an estimate of the frontier for  $m = 1000$ . The data sources will be described below.<sup>9</sup> Note that the frontier is not literally the outer envelope of the data. A higher choice of  $m$  would result in fewer observations that lie beyond the frontier.  $m = 1$  would produce a frontier that is a horizontal line at the sample mean loan amount.

Also note that the frontier does not decline at very high FICO scores (e.g. above 830) even though there are few originations with large loan amounts associated with such credit scores.<sup>10</sup> This result occurs because our method assumes that a borrower with a given FICO is as least as credit worthy as an otherwise-similar borrower with a lower FICO, and so the estimate of the frontier at a given FICO uses the data on originations made to lower FICO borrowers. That said, our methodology does not impose monotonicity on the frontier, and the frontier could have declined at FICO equal to 830 if there were a sufficiently large number of originations with lower loan amounts at such a FICO score.

This example shows a frontier in only two dimensions. In the analysis below, we will focus on four dimensions: credit score, borrower income, downpayment and loan amount. Intuitively, this frontier measures the maximum amount that a bank is willing to lend given a borrower’s credit score, income and downpayment. One important issue worth clarifying is how house prices fit in to this analysis. If we did not condition the frontier on downpayment amount, then the frontier would necessarily be higher in higher-priced areas even if lending standards were the same. Consequently, we would incorrectly

---

<sup>9</sup>The example shown here uses data from the Los Angeles metro area for the year 2005.

<sup>10</sup>One might prefer a more parametric frontier estimator that allows the frontier to continue increasing at very high FICO scores where the data are sparse. However, the estimate of the frontier at areas of the FICO distribution where there is not much population mass has only a small effect on the level of and changes in the *aggregate* frontier, which is what we use in most of our subsequent analysis.

attribute increases in house prices to an expansion of mortgage credit availability. However, we do condition the frontier on a borrower’s downpayment, and the required downpayment would also be larger in higher-priced areas if lending standards are the same. Thus, higher house prices do not necessarily imply a higher frontier. In other words, because variation in house prices affects downpayment size as well as loan amount, it implies movement along the frontier, not shifts in the frontier.

### 3 Data

In applying the CFS methodology to mortgages, we combine two sources of loan-level data. The first source is BlackKnight (formerly named Lender Processing Services), which collects data from a large number of mortgage servicers, including 19 of the 20 largest servicers. Since 2005, BlackKnight has covered roughly 65 to 75 percent of agency loans (i.e. loans subsequently purchased by the GSEs or the FHA), and 20 to 40 percent of loans held on banks’ portfolios.<sup>11</sup> BlackKnight covered fewer servicers in the first half of the 2000s. However, the proportions of GSE, FHA, and portfolio loans in the BlackKnight data are fairly similar to the comparable proportions in the aggregate market, so we are reasonably confident that changes in BlackKnight’s coverage of these three segments of the market will not influence our results, at least through 2013. In 2014, BlackKnight’s coverage of the market may be somewhat less representative than in earlier years because the market share of small banks and nonbank servicers expanded, and these types of institutions are not included in the BlackKnight data.<sup>12</sup>

The second dataset that we use is compiled by CoreLogic and covers loans that were subsequently sold into non-agency mortgage-backed securities. This dataset has covered more than 90 percent of these loans since 2000. Conse-

---

<sup>11</sup>We determine market coverage by comparing total loan volumes for each market segment to aggregate loan volumes published by Inside Mortgage Finance.

<sup>12</sup>The mortgage risk index computed by the American Enterprise Institute suggests that loans originated by nonbanks have riskier characteristics than loans originated by large banks, so our measure may understate mortgage availability in 2014 to some extent.

quently, when we combine these two data sources, we obtain a dataset that provides a comprehensive picture of all of the major segments of the residential mortgage market since 2000.<sup>13</sup>

Our combined dataset includes many variables of interest related to the mortgage origination including the loan amount, the loan-to-value (LTV) ratio, the borrower’s credit score, and the zip code of the property associated with the mortgage loan. One limitation of the loan level data is that the borrower’s income is not directly reported. To obtain the borrower’s income, we merge our loan level data with the confidential version of the Home Mortgage Disclosure Act (HMDA) data.<sup>14</sup> For more information on the HMDA data, see Bhutta and Ringo (2014). Another limitation of the loan level data is that the loan amounts account for first liens only, and there is no reliable way to match first liens to second liens in order to obtain a more comprehensive picture of leverage.

In this paper we compute the frontier using the loan amount and downpayment level as outputs, and the borrower’s FICO score and income as the inputs. We measure the loan amounts, downpayment levels, and incomes in real terms by converting the nominal levels into to 2014 dollars using the price index for personal consumption expenditures. We compute the frontier separately for the 100 largest (by average population between 2001 and 2013) metropolitan areas. In this way, one could think of the borrower’s geography as an additional input. With a more comprehensive dataset, one could calculate a measure of credit availability that incorporates additional inputs and outputs,

---

<sup>13</sup>Although the BlackKnight dataset also includes some non-agency securitized loans, we exclude these loans to avoid double-counting.

<sup>14</sup>Mortgages were matched based on the zip code of the property, the date when the mortgage was originated, the loan amount, and the loan purpose (e.g. purchase, refinance). The match rate was approximately 90 percent and all matches are required to have at a minimum the same loan amount, the same four digit zip code, the same loan purpose, and origination dates within 45 days of each other. Priority is given to matches with the same loan type (e.g. FHA, GSE), the same occupancy status, the same 5-digit zipcode, and smaller absolute differences in origination dates. Flexibility in the match on zip code and origination date is permitted because HMDA reports census tracts rather than zip codes, and some error is introduced when translating census tracts to zip codes, and some origination dates are missing and must be imputed using the closing date of the loan.

which would refine our estimates of credit availability. For example, including information on second liens might boost our estimates of credit availability for the mid-2000s relative to more recent years. Incorporating information on non-mortgage debt would also likely be helpful. Another aspect of credit availability that we do not consider in this analysis is the mortgage rate. Incorporating the interest rate is complicated because the interest-related costs to the borrower depends on other terms of the loan, such as the amount of origination fees paid (e.g. “points”), the length to maturity, and how the loan amortizes, and not all of these terms are observable in our data. Examining a sample of loans with the same terms, such as 30-year fixed rates, is also problematic because loan type is highly correlated with borrower characteristics. Thus, focusing on only one type would exclude an important segment of the market.

For the analysis that follows, we focus exclusively on purchase originations because we are interested in the extension of new credit to households, and refinances are primarily a way for households that already have mortgage credit to obtain better terms.<sup>15</sup> After dropping a small number of loans with LTVs > 120 and loans with appraisal amounts below \$10,000 or above \$5 million, we are left with a sample of *17 million loans* originated between 2001 and 2014 that we use to compute our frontier.

## 4 The Loan Amount Frontier

In this section, we apply the methodology developed in Section 2 and the data introduced in Section 3 to construct the “loan amount frontier”, which is the maximum loan amount that borrowers are able to obtain in a particular period given their FICO score, income, and downpayment amount.<sup>16</sup> We set the

---

<sup>15</sup>We could easily expand the sample to include refinances, and then include a dummy variable for loan purpose type as an additional input.

<sup>16</sup>Although we think of downpayment amount as an output, we can still calculate the loan amount frontier as conditional on a given downpayment.

parameter  $m$  – defined in Section 2 – equal to 1000.<sup>17</sup> We assign FICO scores, downpayments, and incomes to equally-sized bins and estimate the frontier for each bin in each year and each metropolitan area.<sup>18</sup> We limit the sample to the largest 100 metropolitan areas (as ranked by average population between 2001 and 2013) because cell sizes become too small to reliably estimate a frontier in metropolitan areas with fewer mortgage originations. In particular, using weights that we describe below, we aggregate the loan amount frontiers by metropolitan area and bootstrap the aggregated frontiers using 100 repetitions to compute standard errors around our estimates. The 95 percent confidence intervals associated with the aggregate frontier for the 10th, 50th, 100th, 150th, and 200th largest MSAs are shown in Figure 3. Our estimates of the frontier are fairly precise up until around the 100th largest MSA, but beyond the 100th largest MSA, it seems that our dataset does not have enough loans to precisely characterize changes in credit availability given our choice of time frequency for the loan frontier (yearly) and our choice of bin size for the input variables.

Returning to the disaggregated loan frontiers, Table 1 presents some basic facts about the variance of the loan frontier. The average loan frontier is \$270k and the standard deviation is \$185k. One half of the variance in the frontier can be explained by fixed effects for each FICO bin, illustrating that credit supply is strongly affected by a borrower’s credit score. Income is also an important determinant of credit supply, accounting for an additional 12 percent of the variation in the frontier. Metropolitan area fixed effects explain 10 percent of the variation, indicating that there are persistent differences in credit supply across locations even conditional on income and credit score. These differences could reflect persistent differences in economic conditions that are not captured by income. They could also reflect differences in the market structure of banks or geographic variation in the types of lenders.

---

<sup>17</sup>When computing the loan frontier at a given fico, income and downpayment, we first drop the  $\min(5, 0.0001 * nobs_j)$  largest loan balances (where  $j$  indexes the MSA) to minimize the influence of any measurement error. This drop does not have any effect on the asymptotic properties of the frontier.

<sup>18</sup>We use a FICO grid of 480 to 840 with bins of length 20; income bins of 10,000 from 40,000 to 180,000 with additional bins for 200,000, 250,000 and 1,000,000; and a downpayment grid of 0 to 300,000 with bins of length 10,000.

Overall, the dimensions of credit that we consider account for 80 percent of the variation in the frontier, with 20 percent reflecting idiosyncratic variation within these categories.

Figure 4 shows a contour plot of the loan amount frontier by FICO and downpayment for eleven large and diverse metro areas for the year 2004, holding income fixed at \$150,000. Not surprisingly, the frontiers indicate that lenders are generally willing to extend larger loans to borrowers with better credit scores and higher downpayments. The dark blue areas of the frontiers indicate that borrowers with very low credit scores were essentially unable to obtain a loan at all in 2004. Los Angeles and San Francisco have larger loan frontiers for borrowers at the upper ends of the credit score and downpayment distributions relative to the other metro areas shown, perhaps because they contain many neighborhoods with very high house price levels. Although this result might partly reflect demand for larger mortgages in high-priced areas, the fact that banks are willing to make these larger loans for a given downpayment suggests that mortgage credit was more available for these types of borrowers in these cities.

Figure 5 shows the contour plots by FICO and income, holding downpayment fixed at \$50,000. The frontier generally rises with income, suggesting that lenders are willing to supply more credit to higher income borrowers, even holding constant credit score and downpayment amount. This result supports our claim that the loan frontier is determined by credit supply rather than credit demand. If the frontiers were driven by demand, one might expect higher income borrowers to be associated with lower frontiers conditional on downpayment, as higher income households tend to be wealthier and would not want to lever themselves as much as poorer borrowers, all else equal. The slope of the frontier with respect to income is not as steep for lower FICO borrowers, suggesting that borrowers with low credit scores were unable to get large loans, even if their incomes were high.<sup>19</sup>

To more completely describe the loan amount frontiers across years and

---

<sup>19</sup>This result should be taken with a degree of caution because there are few potential borrowers with high incomes and low FICO scores.

the dimensions of credit that we consider (credit score, income, downpayment and location), we aggregate the frontiers across all dimensions but one, and then examine how the frontier changes along the remaining dimension of credit. Downpayments are given equal weights, income and credit score are weighted according to the joint distribution of income and credit score across all observations in our sample, and metropolitan areas are weighted by population.<sup>20</sup> For example, to assess the importance of credit score we calculate the average frontier for each FICO bin across all downpayments, incomes, and metropolitan areas.

As shown by Figure 6, consistent with the contour plots the frontier is higher for higher credit scores. Changes over time are striking. From 2001 to 2005 the frontier expanded by 20 to 30 percent for all credit scores above 560. It then contracted for all credit scores, but by much larger amounts for borrowers at the lower end of the distribution. Whereas decreases between 2005 and 2011 were in the range of 12 to 15 percent for borrowers with a credit score above 700, the frontier fell by nearly 40 percent for borrowers with a credit score around 620 and by 72 percent for borrowers with scores around 600. For borrowers with even lower scores, the frontier fell to zero, indicating that borrowers with these scores were no longer able to obtain credit.

Turning to income, Figure 7 shows the relationship between income and the frontier in 2001, 2004 and 2013.<sup>21</sup> The frontier shifted up similarly at all incomes from 2001 to 2004, indicating that standards eased by similar amounts for borrowers at all income levels. The frontier shifted back down during the financial crisis, and this shift was larger for lower incomes. For higher income borrowers, the 2013 frontier was fairly close to its 2001 level. For borrowers

---

<sup>20</sup>Because our weights are constant over time and across locations, the aggregated frontiers are not a function of changes in observed borrower characteristics over time or differences across locations. An alternative weighting scheme would weight income and credit score according to their shares in the aggregate population, rather than their shares only among mortgage borrowers. However, doing so puts a lot of weight on cells with low credit scores and low incomes, and these cells are imprecisely measured because they contain few mortgage originations.

<sup>21</sup>We examine 2004 to reflect the peak of the housing boom rather than 2005 or 2006 because income misreporting was common in 2005 and 2006. Nevertheless, the income-loan amount frontiers for 2005 and 2006 are quite close to that estimated for 2004.

with incomes below \$60,000, standards in 2013 were somewhat tighter than in 2001.

Figure 8 shows the relationship between downpayment and the frontier in 2001, 2004, 2008 and 2013. The loan amount frontier is increasing and concave in downpayment, illustrating that borrowers can lever themselves more by increasing their downpayment at low downpayment levels, whereas at high downpayment levels larger downpayments are not generally associated with larger loan amounts. The frontier shifted up from 2001 to 2004, and this shift was similar across downpayment sizes.<sup>22</sup> Maximum loan sizes decreased substantially in the first few years of the housing market contraction. This decrease was more pronounced for larger downpayment sizes, flattening the downpayment-loan amount frontier. In particular, after the first 100,000 putting more money down did not allow the borrower to obtain much additional credit. Over the next four years the frontier shifted down further and retained this flatter shape. Thus, in 2013, on average, the maximum attainable loan-to-value ratio was even lower than it had been in 2008. For downpayments below \$70,000, the loan frontier in 2013 was somewhat higher than it was in 2001, whereas for downpayments above \$120,000, it was a touch lower. One way to interpret this result is to consider the inverse of this relationship—the minimum downpayment required for a given loan amount. At low loan amounts below \$300,000 the minimum downpayment in 2013 was a little lower than it had been in 2001, whereas for larger loan amounts the minimum downpayment was higher in 2013 than in 2011.

Figure 9 depicts geographic variation in the frontier by aggregating the frontier by metropolitan area and year, and plotting percentiles of the distribution across metropolitan areas in each year. Differences across metropolitan areas tend to be very stable over time, with metropolitan area fixed effects explaining 72 percent of the variation of the frontier at this level of aggregation.

---

<sup>22</sup>The frontiers shown in the figure may seem low because they are lower than the maximum loan size allowed by the GSEs. This result can be explained by the fact that the figure shows the average of maximum loan sizes across a range of borrower characteristics, and so puts some weight on maximum loan sizes available to borrowers that do not qualify for the maximum GSE-backed loan.

Aggregate shocks also appear to be important, in that the frontier shifted up by similar amounts during the housing boom in most locations, and then subsequently fell by similar amounts in most locations. Indeed, year fixed effects account for 21 percent of the variation at this level of aggregation. In many metropolitan areas, credit availability in 2013 was roughly back to its 2001 level; in 2013 the frontier was within 5 percent of its 2001 value in 81 out of these 100 metropolitan areas. In most other metropolitan areas, the 2013 frontier was somewhat lower than in the early 2000s, suggesting that in those locations lending standards have tightened somewhat relative to the early 2000s, on net.

In summary, the loan amount frontiers are consistent with a number of standard predictions about mortgage credit availability: credit score, income and downpayment are important factors influencing the amount of credit that a borrower can obtain, with more credit available to borrowers with higher scores, higher incomes and larger downpayments. Holding these factors constant, credit availability expanded during the first half of the 2000s and contracted during the financial crisis. Our measure also provides some new insights into credit availability. For example, increases in credit availability during the boom were fairly similar across borrower types and metropolitan areas, whereas the contraction in credit was much sharper for low-score and low-income borrowers. On net, for low-score and low-income borrowers credit was more difficult to obtain in 2013 than in 2001, while for high-score and high-income borrowers the reverse is true. Another noteworthy result is that there are differences in credit availability across metropolitan areas, even for borrowers with the same credit scores, incomes, and downpayments. It is this variation that we will use below to study the effects of mortgage credit availability on housing market outcomes.

## 5 Application: The Effect of Credit Availability on House Prices and Construction

To close the paper, we illustrate the economic value of the loan frontier by showing that it offers new insights on the relationship between credit supply, house prices, and construction activity.

First, we investigate the sensitivity of house prices and construction activity to credit availability, focusing on the differential effects of credit availability conditions across borrower types. Existing measures of credit availability cannot satisfactorily address this issue because they either (i) capture credit demand in addition to credit supply (ii) lack long enough histories to compute precise correlations with prices and construction or (iii) are aggregate measures and so do not characterize the heterogeneity in credit conditions across borrowers.

To this end, we estimate regressions of the following form:

$$y_{jt} = \gamma_1 F_{jt}^{620,680} + \gamma_2 F_{jt}^{680,740} + \beta X_{jt} + \alpha_j + \delta_t + \epsilon_{jt} \quad (1)$$

where  $y_{jt}$  is the log quality adjusted house price level<sup>23</sup> or log single family permits in metro  $j$  at year  $t$ ,  $F^{x,y}$  is the weighted average log loan frontier for individuals with  $x < FICO \leq y$  and income of \$105,000 and a downpayment level equal to \$ 50,000,  $X$  are additional controls,  $\alpha_j$  is a set of metro area fixed effects, and  $\delta_t$  is a set of year fixed effects. For the results presented here, in the matrix  $X$ , we include log income, log employment to population rate, and log delinquency rate, which are included to control for time-varying MSA level fundamentals that may affect both housing market activity and credit availability.<sup>24</sup> We estimate (1) using the estimated loan frontiers for the 80 largest metro areas in our data. Standard errors are clustered at the metro

---

<sup>23</sup>The house price index comes from CoreLogic.

<sup>24</sup>The permits data come from the census' building permits data. The employment rate and income measures come from the Bureau of Economic Analysis. The delinquency rate is computed using our loan level data described in Section 3. The BEA data are not yet available for 2014, so data associated with 2014 are dropped from the regression equation (1).

area and account for first stage estimation error in the frontier [the latter is not actually done for this draft].

Tables 2 and 3 show the results with permits and house prices as the dependent variable, respectively. In our preferred specification (Column 6), we also interact each frontier with the measure of housing supply inelasticity developed by Saiz (2010), as the effect of credit availability on prices and construction should depend on the slope of the housing supply curve. The loan frontier for higher FICO borrowers is significantly related to both permits and prices. For a metro area with a mean supply inelasticity, a one percent increase in the loan frontier increases prices and permits by 0.6 and 0.9 percent, respectively. The effect changes to 0.7 and 0.4 percent, respectively, for a supply inelasticity equal to one standard deviation above the mean. As expected, the effect on prices (permits) is stronger (weaker) for more inelastic metros. There appears to be no effect of the frontier on prices or permits for borrowers with lower FICO scores.<sup>25</sup> The coefficients on the control variables all have the expected signs.

We could not have drawn the insight that the housing cycle is more sensitive to credit availability for less-risky borrowers than to availability for risky borrowers using existing measures of credit supply such as median FICO or the loan approval rate, as these measures do not describe the heterogeneity in credit availability conditions across borrowers. One could construct a naive measure that does allow for heterogeneity across individuals in a way that is similar to the loan amount frontier. Consider, for example, the median loan amount for a particular FICO score and downpayment level. Tables 4 and 5 presents the results of the price and permit regressions where the median loan amount is used as the measure of credit availability instead of the loan frontier. Comparing Tables 2 and 4 with 3 and 5, respectively, we see that the estimated effects are quite different. Using the median loan amount, one might

---

<sup>25</sup>This latter result is particularly interesting in light of recent commentary pointing to tight mortgage credit for lower credit score borrowers as an explanation for the lackluster housing market recovery in recent years. Our results suggest that mortgage credit has indeed been tight for borrowers with lower credit scores in recent years, but that the historical correlation between credit to these borrower types and construction activity is very weak.

conclude that credit availability has no effect on permits and only a small effect on prices. Appendix 1 provides a simple model that illustrates why using the median instead of the frontier gives biased estimates of the effect of credit availability on house prices. The attenuation that arises in practice can occur because the median is an imperfect proxy for the supply of mortgage credit, which effectively introduces noise akin to measurement error in equation (1).

One potential issue with interpreting the results presented above is that  $F^{620,680}$  and  $F^{680,740}$  from equation (1) are *endogenous*. For example, suppose household portfolios in a particular metro area are overly exposed to stocks (relative to the national average) and the stock market increases. This would increase both the local loan frontier (to the extent that lenders use household wealth as an input) and local house prices (to the extent that house prices are increasing in household wealth). To address this potential endogeneity issue, we exploit the disaggregated nature of our frontier measure to create an instrument for credit availability in the spirit of Bartik (1991). The main identification idea is to use the fact that shocks to the national credit markets are exogenous to the local conditions in one particular metropolitan area. Additionally, a shock to national credit markets will have different effects on different metropolitan areas. If, for example, willingness to lend to subprime borrowers increases nationally, the impact of such change will be greater in MSA's where there are a large number of people with low credit scores. Controlling for what happens to the credit markets as a whole through year fixed effects, we can exploit the cross-sectional variation in how these shocks affect different locations as exogenous shocks to local credit supply.

Table 7 presents results where we instrument for the aggregate frontier,  $F_{jt}^{480,840}$ , in MSA  $j$  and year  $t$  using

$$Z_{jt} = s_{jt}^{680} * \frac{1}{J-1} \sum_{i \neq j} F_{it}^{680} + (1 - s_{jt}^{680}) * \frac{1}{J-1} \sum_{i \neq j} F_{it}^{840} \quad (2)$$

where  $s_{jt}^{680}$  is the share of the population in MSA  $j$  at time  $t$  with *FICO* < 680. A one percent increase in the aggregate frontier increases prices and permits by 0.5 and 1.1 percent for an MSA with the mean housing supply

elasticity.<sup>26</sup> The point estimates are not very different from the estimates shown in Table 6, which runs the analogous specification using OLS instead of IV. Taken together, our evidence strongly suggests that easier mortgage credit has a significant positive effect on both house prices and construction activity.

## 5.1 Robustness

One issue with interpreting the results above is that a 50,000 downpayment might be considered small in some markets (such as Los Angeles where the house price level is high) but large in other markets (such as Detroit where the house price level is low). Furthermore, even within a market, the meaningfulness of a 50,000 downpayment could be changing over time. In currently unreported results, we run an analogous set of specifications to those shown in Tables 1 and 2, except we (i) set the downpayment level so that it is fixed at 20 percent of the median MSA house price level in 2001, (ii) set the downpayment level for 2001 at 20 percent of the median MSA level house price in 2001, and then grow this downpayment level over time by the MSA level house price index (iii) set the downpayment level for 2001 at 50,000 in 2001, and then grow this downpayment level over time by the MSA level house price index. The results are qualitatively similar. We also experimented with varying the income level in this way and found that the results are also qualitatively similar.

---

<sup>26</sup>The first stage is strong with a coefficient on the instrument of 2.12 and a t-stat of 10.7.

# A Understanding the Implications of Using the Loan Frontier Versus the Mean Loan Amount in House Price Regressions

A housing market is populated by  $i = 1, \dots, N$  types of households. A household's type  $i$  determines the amount of downpayment it has available,  $d_i$ , and other personal characteristics,  $x_i$  (i.e. credit score). The mass of each type in the housing market is given by  $\pi_i$ . Each household has a maximum borrowing amount, which depends on its type,  $\bar{\ell}_i$ .

Within a type, households are heterogeneous on the amount of housing quality-units they desire,  $h$ . We assume that the amount of housing desired by each household is inelastic. Households either purchase  $h$  units of housing or not at all. The distribution of  $h$  within a household type  $i$  is given by  $G_i(h)$ .

Housing is available at unit price  $p$ . Therefore, a type  $i$  household who demands  $h$  units can afford to purchase the house if and only if:

$$d_i + \bar{\ell}_i \geq ph$$

We can therefore write:

$$\bar{h}_i = \frac{d_i + \bar{\ell}_i}{p}$$

as the maximum quantity of housing that type  $i$  can afford. Also note that household  $i$  only needs a loan to purchase housing if  $d_i < ph$ , so let us write:

$$\underline{h}_i = \frac{d_i}{p}$$

as the quantity of housing at which household  $i$  needs to start borrowing.

For type  $i$  households, the amount borrowed will depend on the quantity of housing desired. Define  $\ell_i^*(h)$  as the loan amount that is originated by a

type  $i$  household who desire  $h$  housing units:

$$\ell_i^*(h) = \begin{cases} 0 & \text{if } h \leq \underline{h}_i \\ ph - d_i & \text{if } \underline{h}_i < h \leq \bar{h}_i \\ 0 & \text{if } h > \bar{h}_i \end{cases}$$

as the loan amount taken out by a type  $i$  household who desires  $h$  housing units when prices are  $p$ .

As the econometrician, we observe loans amounts for loans that are actually originated. Therefore, within a type  $i$  we observe the distribution of loan amounts:

$$\ell_i \sim \ell_i^*(h) \mid \underline{h}_i < h \leq \bar{h}_i$$

The mean of  $\ell_i$  within household type  $i$  is therefore:

$$\begin{aligned} E[\ell_i] &= \int_{\underline{h}_i}^{\bar{h}_i} [ph - d_i] dG_i(h) \\ &= pE[h \mid \underline{h}_i < h \leq \bar{h}_i] - d_i \end{aligned} \quad (3)$$

On the other hand, the maximum order statistic of  $\ell_i$  is:

$$\begin{aligned} E[\max \ell_i] &= p\bar{h} - d_i \\ &= \bar{\ell}_i \end{aligned} \quad (4)$$

This is the essentially what we are estimating in the paper by taking the expected maximum out of  $N$  draws within a household type.

Now suppose that prices in a market are a function of the borrowing limits,  $\bar{\ell}_{1:N}$  and the distribution of housing demand within each household type,  $G_{1:N}$ . We write:

$$p = f(\bar{\ell}_{1:N}, G_{1:N}) \quad (5)$$

Equation (5) illustrates why using the mean (or median) loan amounts as a proxy for credit supply in price regressions is inappropriate. The expected value of  $\ell_i$  depends not only on credit supply  $\bar{\ell}_i$ , but also demand-side factors

$G_i$ . If some aspects of  $G_i$  are not observable, then a regression of price on average loan amount would give biased estimates. The bias arises mechanically from the relationship between  $E[l_i]$  and  $G_i$ . In contrast,  $E[\max l_i]$  correctly estimates  $\bar{\ell}_i$  so no mechanical bias results.<sup>27</sup>

---

<sup>27</sup>There may still be correlation between  $\bar{\ell}_i$  and unobserved housing demand factors, through general equilibrium interactions between the housing and mortgage market. Although important, addressing these issues is not the concern of this paper.

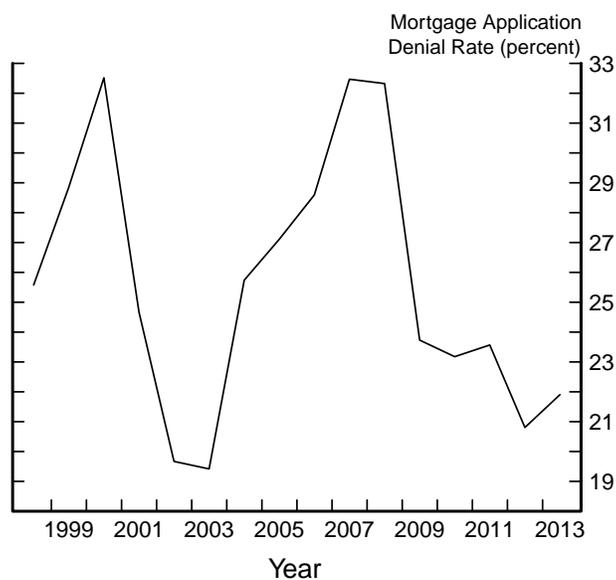
## References

- Adelino, Manuel, Antoinette Schoar, and Felipe Severino**, “Credit supply and house prices: evidence from mortgage market segmentation,” Technical Report, National Bureau of Economic Research 2012.
- Bartik, Timothy J**, “Who benefits from state and local economic development policies?,” *Books from Upjohn Press*, 1991.
- Bhutta, Neil and Daniel R Ringo**, “The 2013 Home Mortgage Disclosure Act Data,” *Federal Reserve Bulletin*, 2014, 100 (4).
- Cazals, Catherine, Jean-Pierre Florens, and Leopold Simar**, “Non-parametric frontier estimation: a robust approach,” *Journal of Econometrics*, 2002, 106, 1–25.
- Favara, Giovanni and Jean Imbs**, “Credit Supply and the Price of Housing,” *American Economic Review*, 2015, 105 (3), 958–92.
- Fuster, Andreas and Basit Zafar**, “The sensitivity of housing demand to financing conditions: evidence from a survey,” 2015.
- Glaeser, Edward L., Joshua D. Gottlieb, and Joseph Gyourko**, “Can Cheap Credit Explain the Housing Boom?,” NBER Working Paper No. 16230 2010.
- Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig**, “Did Securitization Lead to Lax Screening? Evidence from Subprime Loans\*,” *The Quarterly journal of economics*, 2010, 125 (1), 307–362.
- Kung, Edward**, “The Effect of Credit Availability on House Prices: Evidence from the Economic Stimulus Act of 2008,” Research Report, UCLA 2015.
- Li, Wei and Laurie Goodman**, “Measuring Mortgage Credit Availability Using Ex-Ante Probability of Default,” Research Report, The Urban Institute 2014.

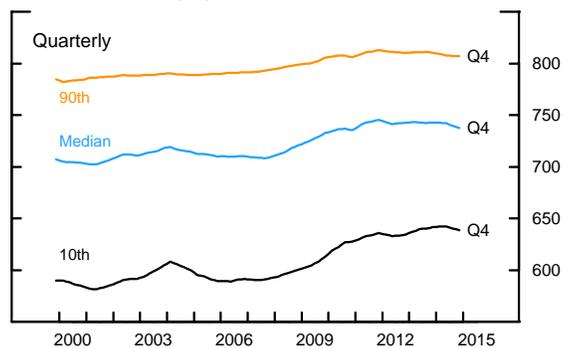
**Mian, Atif and Amir Sufi**, “The Consequences of Mortgage Credit Expansion: Evidence from the US Mortgage Default Crisis\*,” *The Quarterly Journal of Economics*, 2009, *124* (4), 1449–1496.

**Saiz, Albert**, “The geographic determinants of housing supply,” *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Figure 1: Alternative Measures of Credit Availability



(a) Mortgage Application Denial Rate



Source: FRBNY CCP/Equifax  
 Note: Four-quarter moving average. Quarters are quarters of origination.

(b) Credit Score Distribution on New Owner-Occupied Purchase Mortgages

Figure 2: Frontier Example

The dots represent individual mortgage originations and the solid line is an estimate of the frontier using the methodology described in Section 2 for  $m = 1000$ . The loan frontier, shown on the y-axis, is in thousands of dollars.

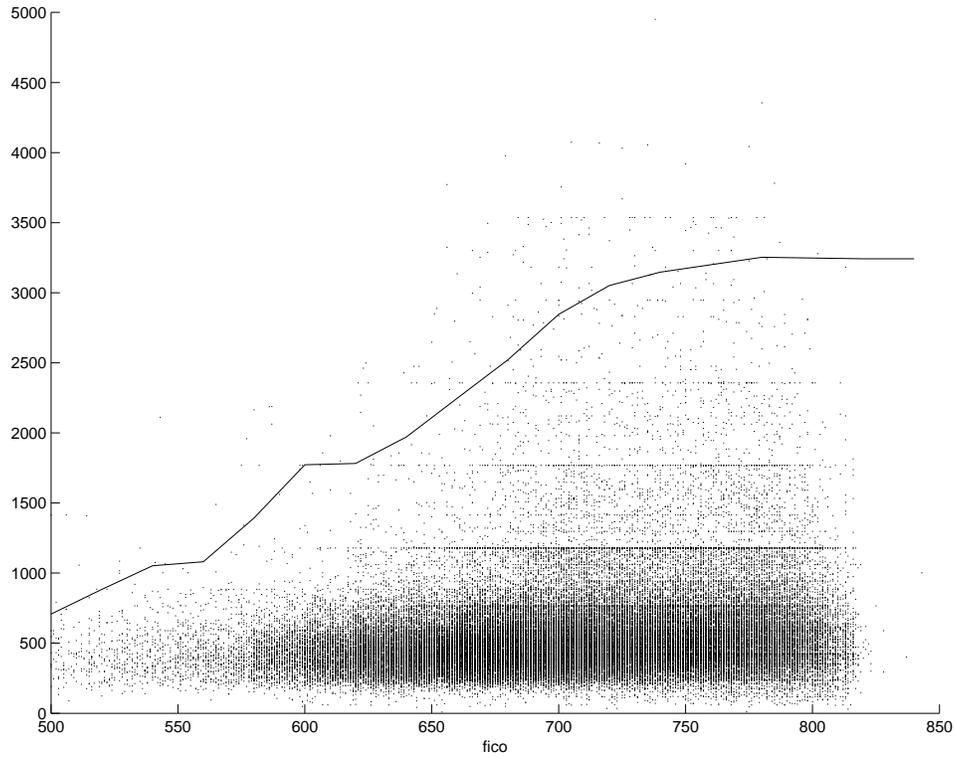


Figure 3: Confidence Intervals for Select Aggregate Loan Frontiers

Dotted lines are 95 percent confidence intervals around the aggregate loan frontier for select metro areas. Standard errors are computed through bootstrap with 100 repetitions. Aggregate loan frontier is the area under the loan amount frontier for each year and city given the choice of weights for each FICO, income, and downpayment described in Section 4. Frontiers for each metro area are normalized to one in 2001.

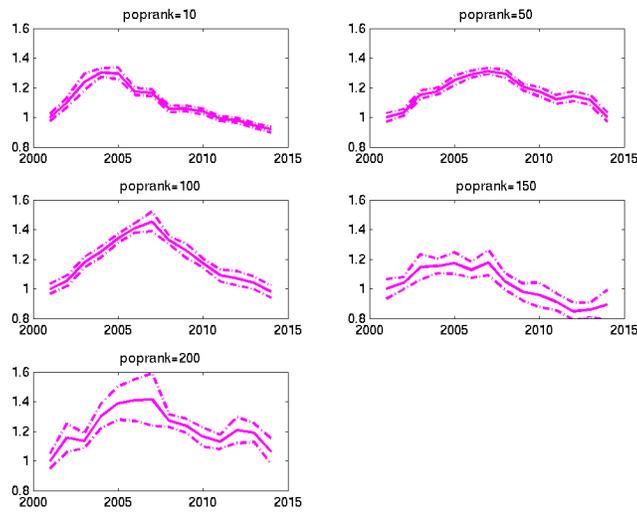


Figure 4: Loan Frontier, Year=2004, Income=150000

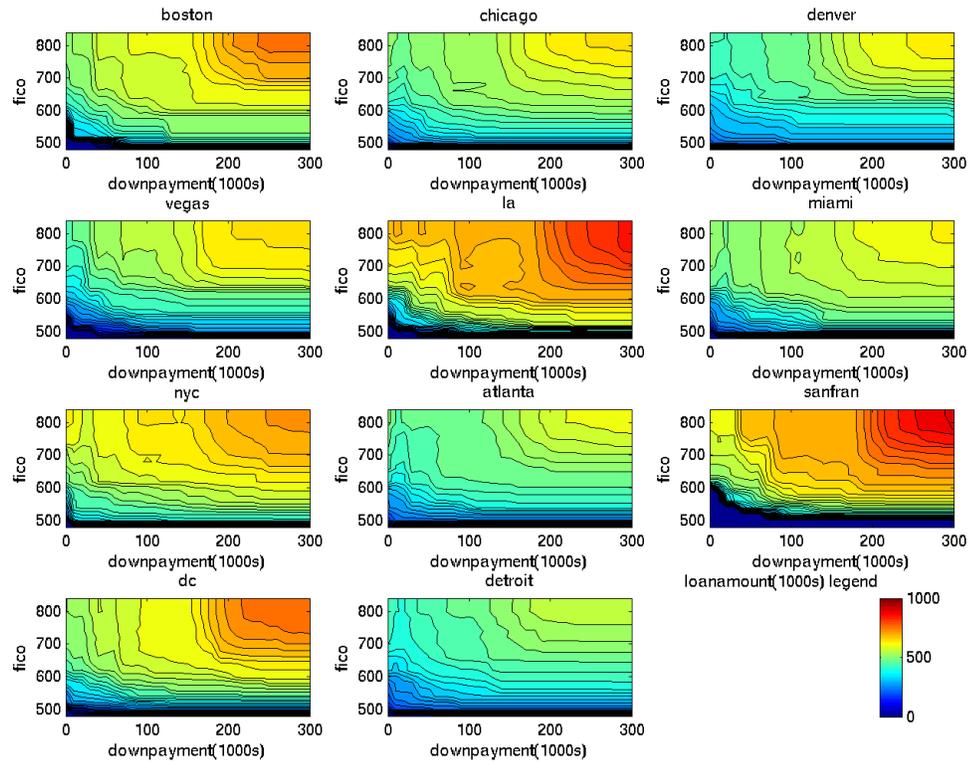


Figure 5: Loan Frontier, Year=2004, Downpayment=50000

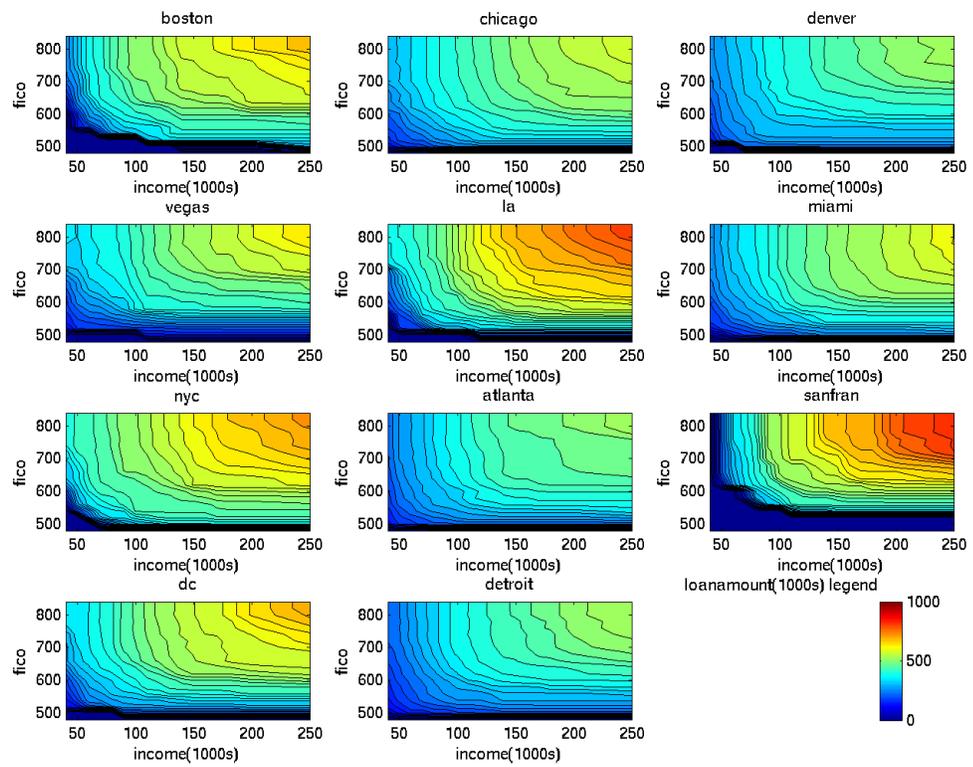


Figure 6: Aggregate Loan Frontiers by FICO

The loan frontier is aggregated over metro areas, incomes, and downpayments using the weights described in Section 4. The loan frontier is in thousands of 2014 dollars.

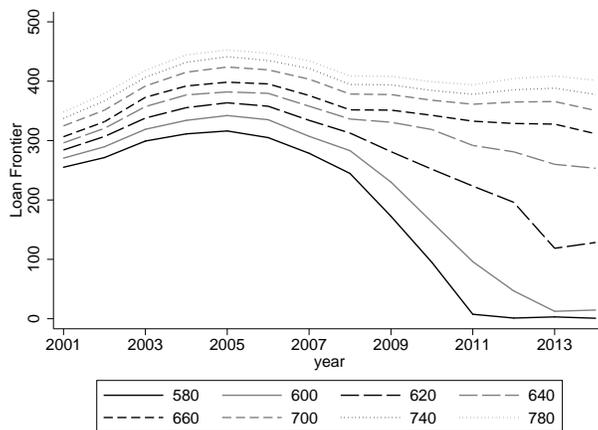


Figure 7: Aggregate Loan Frontiers by Income

The loan frontier is aggregated over metro areas, FICO scores, and downpayments using the weights described in Section 4. The loan frontier is in thousands of 2014 dollars.

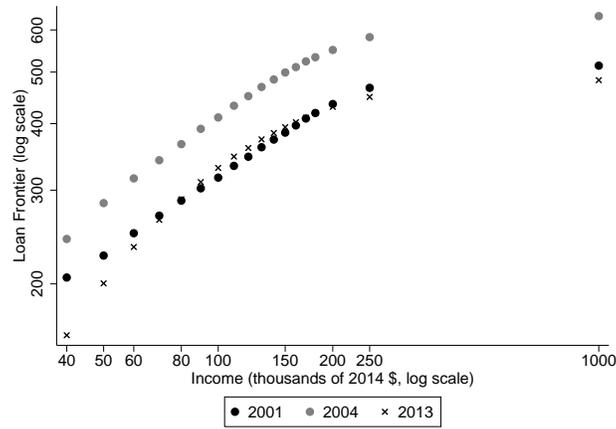


Figure 8: Aggregate Loan Frontiers by Downpayment

The loan frontier is aggregated over metro areas, incomes, and FICO scores using the weights described in Section 4. The loan frontier is in thousands of 2014 dollars.

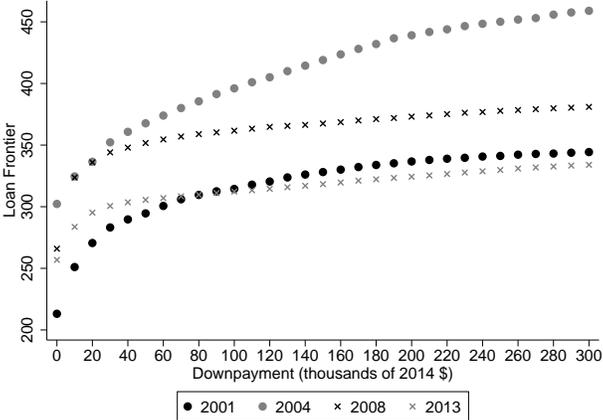


Figure 9: Aggregate Loan Frontiers by Metro Area

The loan frontier is aggregated over downpayments, incomes, and FICO scores using the weights described in Section 4. The loan frontier is in thousands of 2014 dollars. “pX” denotes the Xth percentile of the loan frontier across metro areas within each year.

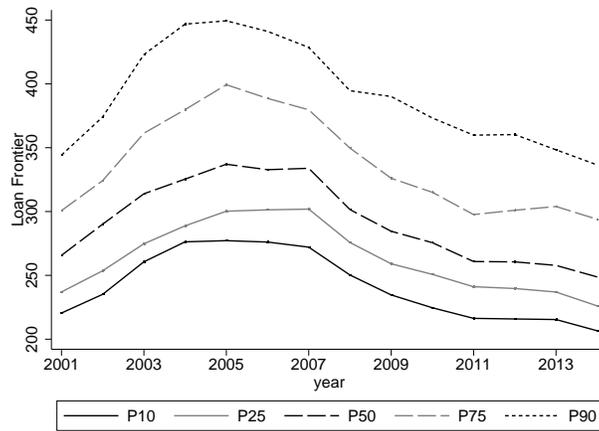


Table 1: Analysis of Variance for Loan Frontier

The average loan frontier is \$270k and the standard deviation is \$185k.

	Dependent Variable: Loan Frontier				
	(1)	(2)	(3)	(4)	(5)
Rsquared	0.52	0.53	0.65	0.7	0.8
FICO F.E.	x	x	x	x	x
Downp F.E.		x	x	x	x
Income F.E.			x	x	x
Year F.E.				x	x
MSA F.E.					x

Table 2: The effects of loan frontier on permits for single family units

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln Frontier_{620,680}$	0.068 (0.071)		-0.031 (0.020)	0.536*** (0.151)		0.033 (0.145)
$\ln Frontier_{680,740}$		0.648** (0.308)	0.694** (0.322)		0.898*** (0.281)	0.874*** (0.304)
$Inelastic \times \ln Frontier_{620,680}$				-0.339*** (0.092)		-0.035 (0.093)
$Inelastic \times \ln Frontier_{680,740}$					-0.510*** (0.138)	-0.480*** (0.165)
Log Delinquency Rate	-0.197*** (0.045)	-0.182*** (0.041)	-0.183*** (0.041)	-0.194*** (0.040)	-0.160*** (0.037)	-0.162*** (0.037)
Log Income	2.592*** (0.537)	2.208*** (0.582)	2.190*** (0.592)	2.199*** (0.547)	2.139*** (0.521)	2.113*** (0.533)
Log Employment	1.693*** (0.497)	1.367** (0.529)	1.359** (0.530)	1.476*** (0.480)	1.269** (0.494)	1.266** (0.498)
Observations	1040	1040	1040	1040	1040	1040
$R^2$ overall	0.872	0.876	0.876	0.880	0.886	0.886

Note - All the variables in this regression are in logs. The dependent variable is the log single family permits in a metropolitan area. The  $Frontier_{x,y}$  is the loan frontier for people with  $x < FICO \leq y$  weighted by the FICO shares of the particular CBSA in a particular year. The sample consists of annual data from 2001 to 2013 for 80 metropolitan areas (CBSAs). All specifications include CBSA and year level fixed effects. The clustered robust standard errors are given in parentheses.

\* statistical significance at the 90% level

\*\* statistical significance at the 95% level

\*\*\* statistical significance at the 99% level

Table 3: The effects of loan frontier on metropolitan area house prices

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln Frontier_{620,680}$	0.047 (0.045)		-0.041*** (0.007)	0.096** (0.043)		-0.040 (0.040)
$\ln Frontier_{680,740}$		0.559*** (0.075)	0.620*** (0.078)		0.497*** (0.077)	0.557*** (0.091)
$Inelastic \times \ln Frontier_{620,680}$				-0.036 (0.025)		-0.002 (0.025)
$Inelastic \times \ln Frontier_{680,740}$					0.121*** (0.029)	0.125*** (0.039)
Log Delinquency Rate	-0.124*** (0.015)	-0.111*** (0.012)	-0.111*** (0.012)	-0.123*** (0.015)	-0.116*** (0.012)	-0.116*** (0.013)
Log Income	1.124*** (0.198)	0.787*** (0.154)	0.762*** (0.149)	1.083*** (0.196)	0.805*** (0.160)	0.777*** (0.155)
Log Employment	0.958*** (0.193)	0.667*** (0.155)	0.656*** (0.154)	0.934*** (0.191)	0.700*** (0.159)	0.688*** (0.157)
Observations	1027	1027	1027	1027	1027	1027
$R^2$ overall	0.823	0.871	0.874	0.825	0.879	0.882

Note - All the variables in this regression are in logs. The dependent variable is the log Zillow house price in a metropolitan area. The  $Frontier_{x,y}$  is the loan frontier for people with  $x < FICO \leq y$  weighted by the FICO shares of the particular CBSA in a particular year. The sample consists of annual data from 2001 to 2013 for 80 metropolitan areas (CBSAs). All specifications include CBSA and year level fixed effects. The clustered robust standard errors are given in parentheses.

\* statistical significance at the 90% level

\*\* statistical significance at the 95% level

\*\*\* statistical significance at the 99% level

Table 4: The effects of median loan amount on permits for single family units

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln Median_{620,680}$	-0.018 (0.128)		-0.086 (0.141)	0.026 (0.133)		-0.075 (0.138)
$\ln Median_{680,740}$		0.029 (0.166)	0.098 (0.208)		0.053 (0.204)	0.136 (0.242)
$Inelastic \times \ln Median_{620,680}$				-0.120 (0.142)		-0.229 (0.152)
$Inelastic \times \ln Median_{680,740}$					-0.036 (0.154)	0.103 (0.151)
Log Delinquency Rate	-0.201*** (0.044)	-0.199*** (0.044)	-0.198*** (0.044)	-0.202*** (0.045)	-0.198*** (0.042)	-0.202*** (0.043)
Log Income	2.628*** (0.555)	2.591*** (0.577)	2.602*** (0.575)	2.627*** (0.547)	2.582*** (0.581)	2.596*** (0.573)
Log Employment	1.743*** (0.499)	1.718*** (0.492)	1.745*** (0.501)	1.683*** (0.498)	1.704*** (0.487)	1.670*** (0.495)
Observations	1040	1040	1040	1040	1040	1040
$R^2$ overall	0.871	0.871	0.872	0.872	0.871	0.872

Note - All the variables in this regression are in logs. The dependent variable is the log single family permits in a metropolitan area. The  $Median_{x,y}$  is the median loan amount for people with  $x < FICO \leq y$ . The sample consists of annual data from 2001 to 2013 for 80 metropolitan areas (CBSAs). All specifications include CBSA and year level fixed effects. The clustered robust standard errors are given in parentheses.

\* statistical significance at the 90% level

\*\* statistical significance at the 95% level

\*\*\* statistical significance at the 99% level

Table 5: The effects of median loan amount on metropolitan area house prices

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln Median_{620,680}$	0.227*** (0.033)		0.080** (0.032)	0.193*** (0.031)		0.077** (0.032)
$\ln Median_{680,740}$		0.276*** (0.037)	0.212*** (0.047)		0.262*** (0.044)	0.195*** (0.054)
$Inelastic \times \ln Median_{620,680}$				0.091*** (0.030)		0.049 (0.031)
$Inelastic \times \ln Median_{680,740}$					0.021 (0.027)	-0.008 (0.028)
Log Delinquency Rate	-0.122*** (0.016)	-0.116*** (0.016)	-0.117*** (0.016)	-0.121*** (0.016)	-0.116*** (0.016)	-0.116*** (0.016)
Log Income	0.967*** (0.184)	0.918*** (0.178)	0.909*** (0.178)	0.966*** (0.185)	0.923*** (0.179)	0.912*** (0.180)
Log Employment	0.838*** (0.188)	0.865*** (0.178)	0.842*** (0.181)	0.882*** (0.188)	0.873*** (0.179)	0.861*** (0.183)
Observations	1027	1027	1027	1027	1027	1027
$R^2$ overall	0.846	0.854	0.855	0.849	0.854	0.856

Note - All the variables in this regression are in logs. The dependent variable is the log Zillow house price in a metropolitan area. The  $Median_{x,y}$  is the median loan amount for people with  $x < FICO \leq y$ . The sample consists of annual data from 2001 to 2013 for 80 metropolitan areas (CBSAs). All specifications include CBSA and year level fixed effects. The clustered robust standard errors are given in parentheses.

\* statistical significance at the 90% level

\*\* statistical significance at the 95% level

\*\*\* statistical significance at the 99% level

Table 6: The OLS effects of the loan frontiers on prices and permits for single family units

Dep. variable:	<i>lnPrice</i>		<i>lnPermits</i>	
	(1)	(2)	(3)	(4)
<i>lnFrontier</i> <sub>480,840</sub>	0.514*** (0.069)	0.471*** (0.072)	0.731** (0.284)	0.891*** (0.266)
<i>Inelastic</i> × <i>lnFrontier</i> <sub>480,840</sub>		0.114*** (0.032)		-0.445*** (0.131)
Log Delinquency Rate	-0.108*** (0.011)	-0.108*** (0.011)	-0.174*** (0.039)	-0.172*** (0.037)
Log Income	0.739*** (0.139)	0.766*** (0.146)	2.049*** (0.568)	1.942*** (0.559)
Log Employment	0.770*** (0.145)	0.765*** (0.146)	1.435*** (0.477)	1.476*** (0.467)
Observations	1027	1027	1040	1040
<i>R</i> <sup>2</sup> overall	0.876	0.883	0.879	0.886

Note - All the variables in this regression are in logs. The *Frontier*<sub>480,840</sub> is the loan frontier for people with  $480 < FICO \leq 840$  weighted by the FICO shares of the particular CBSA in a particular year. The sample consists of annual data from 2001 to 2013 for 80 metropolitan areas (CBSAs). All specifications include CBSA and year level fixed effects. The clustered robust standard errors are given in parentheses.

\* statistical significance at the 90% level

\*\* statistical significance at the 95% level

\*\*\* statistical significance at the 99% level

Table 7: The IV effects of the loan frontiers on prices and permits for single family units

Dep. variable:	<i>lnPrice</i>		<i>lnPermits</i>	
	(1)	(2)	(3)	(4)
<i>lnFrontier</i> <sub>480,840</sub>	0.556*** (0.102)	0.501*** (0.110)	0.940** (0.461)	1.087** (0.477)
<i>Inelastic</i> × <i>lnFrontier</i> <sub>480,840</sub>		0.170*** (0.038)		-0.451*** (0.130)
Log Delinquency Rate	-0.106*** (0.010)	-0.106*** (0.011)	-0.166*** (0.039)	-0.165*** (0.038)
Log Income	0.706*** (0.150)	0.741*** (0.159)	1.888*** (0.704)	1.791*** (0.677)
Log Employment	0.753*** (0.150)	0.742*** (0.152)	1.350*** (0.480)	1.398*** (0.472)
Observations	1027	1027	1040	1040
<i>R</i> <sup>2</sup> overall	0.875	0.880	0.878	0.886

Note - All the variables in this regression are in logs. The *Frontier*<sub>480,840</sub> is the loan frontier for people with  $480 < FICO \leq 840$  weighted by the FICO shares of the particular CBSA in a particular year. The instrument is a Bartik type instrument that translates national shocks to the frontier for subprime borrowers ( $FICO \leq 680$ ) and prime borrowers ( $FICO > 680$ ) to CBSAs where such borrowers are located. The sample consists of annual data from 2001 to 2013 for 80 metropolitan areas (CBSAs). All specifications include CBSA and year level fixed effects. The clustered robust standard errors are given in parentheses.

\* statistical significance at the 90% level

\*\* statistical significance at the 95% level

\*\*\* statistical significance at the 99% level