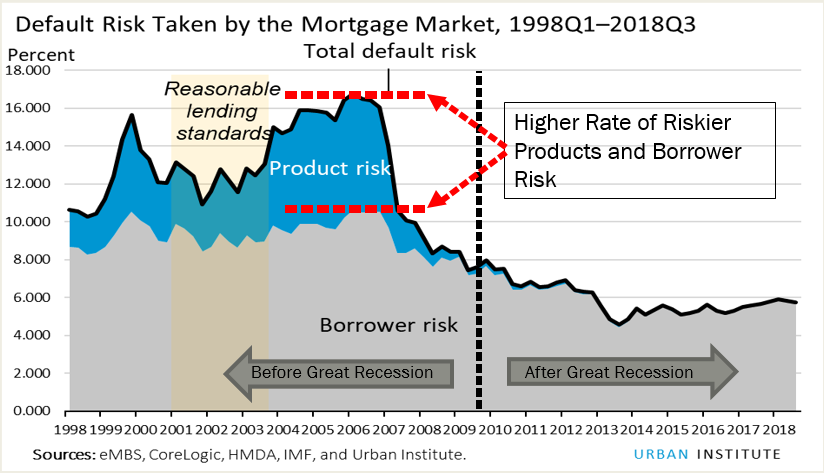
**Introduction:**

It is well known that the great recession was heavily induced by the mortgage market. Relaxed lending standards in the early 2000’s was a primary driver of the mortgage collapse, relaxed lending standards comes in two forms, borrowers and product risk.

Examples of both are as follows, Product Risk, Adjustable Rate Mortgages (ARM), offers an upfront “teaser” rate which is lower than fixed traditional. The idea behind these is that for let’s say the first five or seven years the monthly payments are lower due to the interest rate, however it steps up after the five or seven year mark which makes the monthly payment higher. Higher monthly payments can potentially introduce financial difficulties for the borrower. Second Liens, when a borrower purchases a home generally they will take out a first lien mortgage. A first lien mortgage is secured by the actual home, that is, if the borrower defaults the owner of the first lien mortgage would be able to recoup the amount loan amount by repossessing the home. A second lien is an additional lien taken out on a home, normally these are taken out for home improvements and can be in the form of a Home Equity Line of Credit (HELOC) which acts like a credit card that you borrower against your home. Furthermore, some of the products required “Low” to “No” documented financials from the borrower. Borrower Risk, examples, lending to a borrower with a poor credit history or with no income history.

***Graph 1 Housing Credit Availability Index***

 [*https://www.urban.org/policy-centers/housing-finance-policy-center/projects/housing-credit-availability-index*](https://www.urban.org/policy-centers/housing-finance-policy-center/projects/housing-credit-availability-index)

Graph 1 is a depiction of lenders appetite for risk. A higher rate for the Housing Credit Availability Index (HCAI) means that a bank or lending institution is more willing to lend to borrowers with lower income or poor credit history. Additionally, this index also includes the product risk. Which represents a lending institutions appetite for products that have a higher default rate (HELOC’s and ARM’s). Clearly, after the great recession lending institutions tightened their origination requirements.

Subprime lending is a type of mortgage where the borrowers risk or product risk is higher or both. Case in point, a lending institution lends to a borrower with a lower FICO score, say below 640. Subprime lending can also be called “non-conforming” where “conforming” is defined by the Federal Housing and Finance Agency (FHFA)1. In 2005, a conforming loan had to be less than $359,650. Loans over this amount would be considered non-conforming. Risker borrowers or non-conforming loans have a higher probability to default or not be able to pay their mortgage. Due to this, lending institutions will charge the borrower a higher interest rate for their loan.

The focal point of this paper is to associate subprime lending to higher interest rate loans (higher priced). In general, lending institutions utilize a prime rate2 which is a benchmark for interest rates for mortgages. For loans that are in the first lien position and the interest rate is three percentage points above the prime rate the loan is classified as a subprime higher priced loan. For second liens, if the interest rate charged is five percentage points above the prime rate the loan is classified as a subprime higher priced loan.

Utilizing the subprime higher priced definition, I will use various external factors to attempt to capture the borrower characteristics these subprime loans were being originated too.

**Literature Review:**

During a ten year stretch, from 1994 to 2004 home ownership rose from 63.8% to 69.2%. Mainly the increase was due to relaxed lending guidelines by financial institutions, as well as more diversified products, like Adjustable Rate Mortgages (ARMs) and Piggy Back loans (Dimartino and Duca). The relaxation of lending led many borrowers to deviate from traditional mortgages and be able to obtain loans with little or no financials and purchase outside of conforming loan limits.

Freddie Mac (FHLMC) and Fannie Mae (FNMA) are Government Sponsored Enterprise (GSE) which historically had stricter standards for loan originations than private institutes. GSE’s focused on traditional mortgage originations, which required loans to be within a conforming limit ($359,650 at the national level in 2005) and was created to make housing affordable to underserved areas, mainly in inner cities (Wallison). The Government requires GSE’s to hold a certain percentage of loans be originated to the underserved community, in 2007 38% of originations needed to be granted to the underserved community and 25% to low-income or very low income borrowers (Wallison).

Second, the literature shows conflicting theories concerning GSE’s and the financial crisis, for example, Passmore and Shrelund state GSE loans were less likely to be affiliated with subprime. Whereas Wallison states that due to government requirements during the 2005 through 2007 period forced the GSE’s into subprime territory.

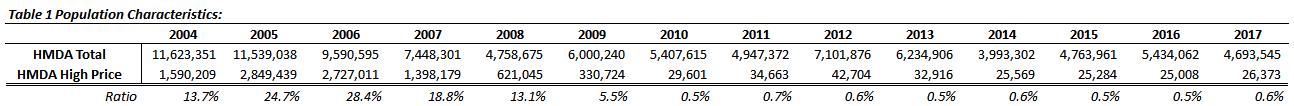
1 <https://www.fanniemae.com/content/fact_sheet/historical-loan-limits.pdf>

2 <https://fred.stlouisfed.org/series/MPRIME>

**Data:**

The primary source of the data comes from the Home Mortgage Disclosure Act (HMDA). The HMDA was enacted by Congress in 1975 and implemented by the Federal Reserve Board’s Regulation C. On July 21, 2011 the HMDA was transferred to the Consumer Financial Protection Bureau (CFPB)3. HMDA’s primary purpose is to ensure fair lending practices are occurring within the mortgage origination realm. It is required for lending institutions to submit various data characteristics when a borrower requests a loan application, all loan applications are from these institutions are submitted to the Loan Application Registrars (LAR). LAR files cover approximately 80% of all mortgage originations (Avery, Brevoort and Canner, 2007).

The LAR files contains characteristics concerning a loan application for a borrower. For this analysis the data is subset to applications that were originated, first lien, one to four family homes conventional loans. Using this population a subprime definition is constructed. Loans that were originated with an interest rate that was 3% over the industry standard are classified as a “Higher Priced” subprime loan. The definition of “industry standard” is set by the Federal Financial Instructions Examination Council (FFIEC) which uses a national average for 30 and 15 year fixed interest rate loans and adds the Treasury yield4. Table 1 represents the frequency of total mortgages and higher priced subprime loans between 2004 and 2017 reported in the HMDA LAR data.



Notes: Subset to First Lien Mortgages originated loans, excludes Puerto Rico, Hawaii, Alaska. Excludes loans not identified as high priced which were originated by the Federal Housing Administration (FHA) or Veterans Administration (VA).

Additional data was sourced from the Federal Reserve Economic Data (FRED). Equifax generates a scoring algorithm which assess a borrower’s credit score. Credit scores generally range from 300 to 850 with 850 representing excellent credit. The FRED stores this data at the county level and presents it as a percentage of the population within a county that has a credit score below 660.

Attributes associated to the higher priced classification were derived from the data. These attributes are as follows.

1. Income Quintile: Income quintiles are created using the aggregated LAR’s data for a given year. If the stated income reported for the LAR record is between the falls within the average quintile at the state level a value of “1” would be assigned, otherwise “0.” These binary variables are then aggregated by the state and county.
2. Equifax Credit Score: is the percentage of the population within the county which fall below 660.
3. Percentage Black: is the number of African Americans applicants approved for a mortgage in the county.
4. Percentage Hispanic: is the number of Hispanic applicants approved for a mortgage in the county.
5. Ownership Rate: is the number of approved mortgages which is the applicants primary residence.

3 <https://www.ffiec.gov/hmda/history2.htm>

4 <https://www.ffiec.gov/ratespread/newcalchelp.aspx#9>



Table 2 statistics represent cohort data generated from account level. The cohort data is aggregated by State, County and year. Income quintiles, race and ownership are summations at the County level, Equifax Credit Score percentage of persons below 660 for the County. Lastly, a weighted approach was taken, which is the total housing units in the County divided by the total housing units in the State.

Review of table 2 shows that the Equifax percentage of low credit score borrowers reflects the changes in FICO version 8 to version 9 scoring methodology. The release of FICO 9 was in August of 20145 which we can see the shift, in 2010 the median Equifax percent subprime population in aggregate was 30.5% and in 2014 it decreased by 300 basis points, to 26.95%. Another interesting trend is the amount of originations for minorities, Black and Hispanics was above 0.20 prior to 2007. 2008 onward this was cut in half and revolved around 0.10, one could conclude that originations to minorities decreased after 2008.

**Regression Analysis:**



Notes: Italics are reported p-values

5 <https://www.moneyunder30.com/fico-9>

Table 3a reports a pooled OLS regression, with the dependent variable being a ratio of total number of High Priced Subprime loans in a given County divided by total number of housing units within the County. Here, income quintile 5 is the reference group and is withheld from the model. Interpreting these results should be as follows, for a 1% point increase in the population within a County which is below the Equifax credit score 660 threshold one would expect to see the a 0.029% (weighted, year 2005) increase in High Priced Subprime mortgages within the County. If there is a 1% increase in Income Quintile 1 for a County this would relate to a 0.253% (weighted, year 2005) increase in High Priced Subprime loans in the County.

Through visual inspection, in 2010 onward the coefficients are considerably smaller, likely this is due to the number of high priced subprime loans is essentially zero and is due to the change in lending standards post the great recession. Generally speaking, the income quintiles has the most impact prior to 2010, however using a pooled regression the coefficients are not stable across years.



Notes: Italics are reported p-values

Using the same predictors from the pooled regression, Table 3b represents coefficients using State level fixed effects. Like the pooled results from Table 3a, the coefficients decrease substantially after 2010.



Notes: only weighted regressions were used.

Table 3a represents a pooled OLS for 2009 and prior and post 2009. Additionally, a State and Year Fixed effects was ran for the same time periods. Splitting the data between 2009 and prior versus 2010 and forward shows the impact of tightened lending standards that came during the great recession. The results from table 3c closely align graph 1 HCAVI which one could conclude that a structural break in lending standards occurred. Focusing on the on the 2004 through 2009 timeframe, overall the coefficients for both Pooled OLS and Fixed effects are similar. Lower stated income borrowers will result in a higher probability of high priced subprime originations relative to the highest income borrowers. Counties which have a higher percentage of minorities, being Black or Hispanic would result in a higher rate of high priced subprime mortgages within a county.

**Conclusion:**

Using High Priced to identify a subprime mortgage has many caveats associated with it. It was shown that there was a 13% increase in the number of high priced loans in 2004 and 2005 solely due to the flattening of the yield curve (Avery, Brevoort and Canner, 2007). This alone raises question into the validity of the high priced definition, furthermore as table 1 represented, 2010 onward shows that the number of high priced subprime loans essentially went to zero.

The main result one could infer from this paper is that, in general the root cause of higher priced subprime loans were primarily originated to minorities and/or lower income borrowers.

**Future Work:**

There was a big scramble to understand why the housing market was collapsing in the late 2000’s. Most work reviewed was using a Housing and Urban Development (HUD) Subprime definition or Higher Priced. The HUD stopped producing the subprime definition in 2006. The high priced subprime definition can be used, however for all intensive purposes this definition is really tied to market characteristics and heavily susceptible to the yield curve flattening, which is the case in late 2018 and early 2019.

I would propose the following:

The LAR data has a field called “respondent\_id” and “agency\_code” in combination one can use these to get the lending institution. A new definition of Subprime would be created using the lending institution and will be based on a list of known failed institutions from the financial crisis. The failed institutions were known for using questionable listing standards in the years prior to the crisis. Using the defined failed institutions, a comparison will be performed to the HUD subprime classification for comparison purposes. A failed institution list was obtained from Wikipedia <https://en.wikipedia.org/wiki/List_of_banks_acquired_or_bankrupted_during_the_Great_Recession>. Review of each institution will provide lineage, example the lending institutions named changed. These institutions are identifiable in the LAR’s data.

Predictors to be included:

1. The allocation of Total Personal Income for the County to the State
   1. County: <https://fred.stlouisfed.org/series/PI01001>
   2. State: <https://fred.stlouisfed.org/series/ALOTOT>
2. The spread between Per Capita Personal Income for the County to the State
   1. County: <https://fred.stlouisfed.org/series/PCPI01001>
   2. State: <https://fred.stlouisfed.org/series/ALPCPI>
3. The spread between All-Transactions House Price Index for the County to the State
   1. County: <https://fred.stlouisfed.org/series/ATNHPIUS01001A>
   2. State: <https://fred.stlouisfed.org/series/ALSTHPI>
4. Equifax Subprime Credit by County
   1. Extracted, review of table 2 simple statistics shows that the definition of the credit score seems to have changed in 2013, or on aggregate for all Americans has increased. Understand why this decrease occurred.
5. The allocation of SNAP Benefits from the County to the State
   1. Small Area Income and Poverty (SNAP), it measures the number of persons receiving government assistance relating to poverty
   2. County: <https://fred.stlouisfed.org/series/CBR01001ALA647NCEN>
   3. State: <https://fred.stlouisfed.org/series/BR01000ALA647NCEN>
6. Rental Vacancy and Home Ownership Rate for the State
   1. <https://fred.stlouisfed.org/series/ALRVAC>
   2. <https://fred.stlouisfed.org/series/ALHOWN>
7. The allocation of All Ages in Poverty (Poverty Universe) from the County to the State
   1. County: <https://fred.stlouisfed.org/series/PEAAAL01001A647NCEN>
   2. State: <https://fred.stlouisfed.org/series/PEAAAL01000A647NCEN>
8. Total People used in the Poverty Universe for State and County
   1. More than likely, will be used as a weighting mechanism. Measures the State and County Total Population. Aligns closely to the Resident population
   2. County: <https://fred.stlouisfed.org/series/PUAAAL01001A647NCEN>
   3. State: <https://fred.stlouisfed.org/series/PUAAAL01000A647NCEN>
9. Spread between the Unemployment Rate from County to the State
   1. County: <https://fred.stlouisfed.org/series/LAUCN010010000000003A>
   2. State: https://fred.stlouisfed.org/series/ALURN
10. Spread between the Median Household Income from County to the State
    1. Needs adjustment using a CPI
    2. County: <https://fred.stlouisfed.org/series/MHIAL01001A052NCEN>
    3. State: <https://fred.stlouisfed.org/series/MHIAL01000A052NCEN>
11. Spread between the Total Building Permits for the County to the State
    1. County: <https://www2.census.gov/econ/bps/County/>
    2. State: <https://www2.census.gov/econ/bps/State/>
12. Spread between the Zillow Single Family Home Value for the County to the State
    1. Only ½ the Counties in the United States are represented in the Zillow data
    2. <https://www.zillow.com/research/data/>
13. Spread between the Zillow Median Price per Square Foot for the County to the State

National levels will potentially be used for normalization of the predictors. As an example, the spread between Poverty levels in Autauga County Alabama compared to the State level would not account for the State having high poverty in general. Post normalization, the index for all predictors would be set to 2001 as the base year, i.e. 2001 would take a value of zero.

A logistic regression at the account level of the LAR’s data will be performed to give a probability that the origination was a subprime loan using the failed institutions list definition. The logistic regression will be trained on the years 2004 to 2007. Scoring the data from 2008 to 2017 (2018 if the LAR’s data is published) will hopefully show that new online origination companies (Quiken Loans as an example) is originating in Counties that that had characteristics similar to the years prior to the Great Recession.

Thus far, all predictors have been gathered, data construction and modeling needs to take place. The new definition of failed subprime institutions during has been created, more testing on this will need to be established prior to modeling.

**References:**

Avery, Robert, Kenneth Brevoort, and Glenn Cannter. 2007. “Opportunities and Issues In Using HMDA Data.” Journal of Real Estate research, 29(4):351-79.

DiMartino, Danielle, and John V. Duca.  “The Rise and Fall of Subprime Mortgages,” Federal Reserve Bank of Dallas *Economic Letter* 2, no. 11 (November 2007): 1-8.

Passmore, Wayne and Shane Sherlund. “FHA, Fannie Mae, Freddie Mac, and the Great Recession,” Finance and Economic Discussion Series, Division of Research & Statistics and Monetary Affairs. Federal Reserve Board, Washington, D.C

Peter J. Wallison (2009) CAUSE AND EFFECT: GOVERNMENT POLICIES AND THE FINANCIAL CRISIS, Critical Review, 21:2-3, 365-376, DOI: [10.1080/08913810902934158](https://doi.org/10.1080/08913810902934158)