**Abstract:**

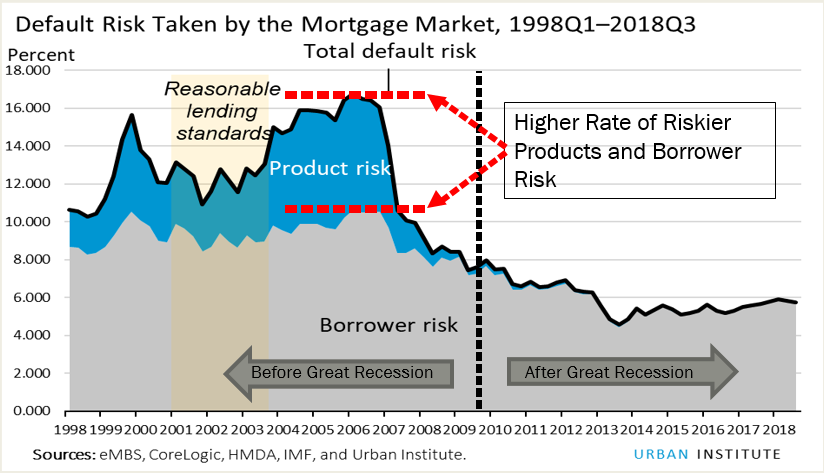
Goes here…

**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 source of the mortgage collapse. Relaxed lending standards comes in two forms, borrowers risk 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 the first few the monthly payments are lower due to the interest rate, however it increases 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 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 lender’s 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 institution’s 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.

Prior to the financial crisis subprime lending had both, higher borrowers risk and product risk versus afterwards, which had lower product risk and borrower risk. 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. Riskier 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 work is to compare two of the traditional subprime loan definitions, Housing and Urban Development (HUD) Subprime and Home Mortgage Disclosure Act (HMDA) Higher Priced Subprime to a derived subprime definition. HUD Subprime loans were associated to specific lending institutions. Mainly, HUD would call a lending institution and ask what percentage of loans are you originating which were deemed subprime. If the institution responded with 50% or more of the originations as subprime, they would be considered a HUD subprime institution. HMDA Higher Priced was classified at the individual loan. ). 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. Lastly, the derived subprime definition utilizes a list of lending institutions which went bankrupt or failed during the financial crisis. Loans originated by these failed investors are marked as subprime.

Utilizing these three defintions, I will compare the 2004 through 2007 originations at the loan level to understand how the definitions overlap. After, I will compare loans originated from 2008 through 2016 to see the amount of subprime loans being originated.

The ability to understand the housing market from a lending perspective is a powerful tool. Given the financial crisis, it is imperative to ensure that lending standards for home loans is healthy and the borrowers can afford the homes they are purchasing. If subprime mortgages are on the rise, as they were in the early 2000’s stricter government oversight would likely be beneficial to ensure the health of the home mortgage market.

**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:**

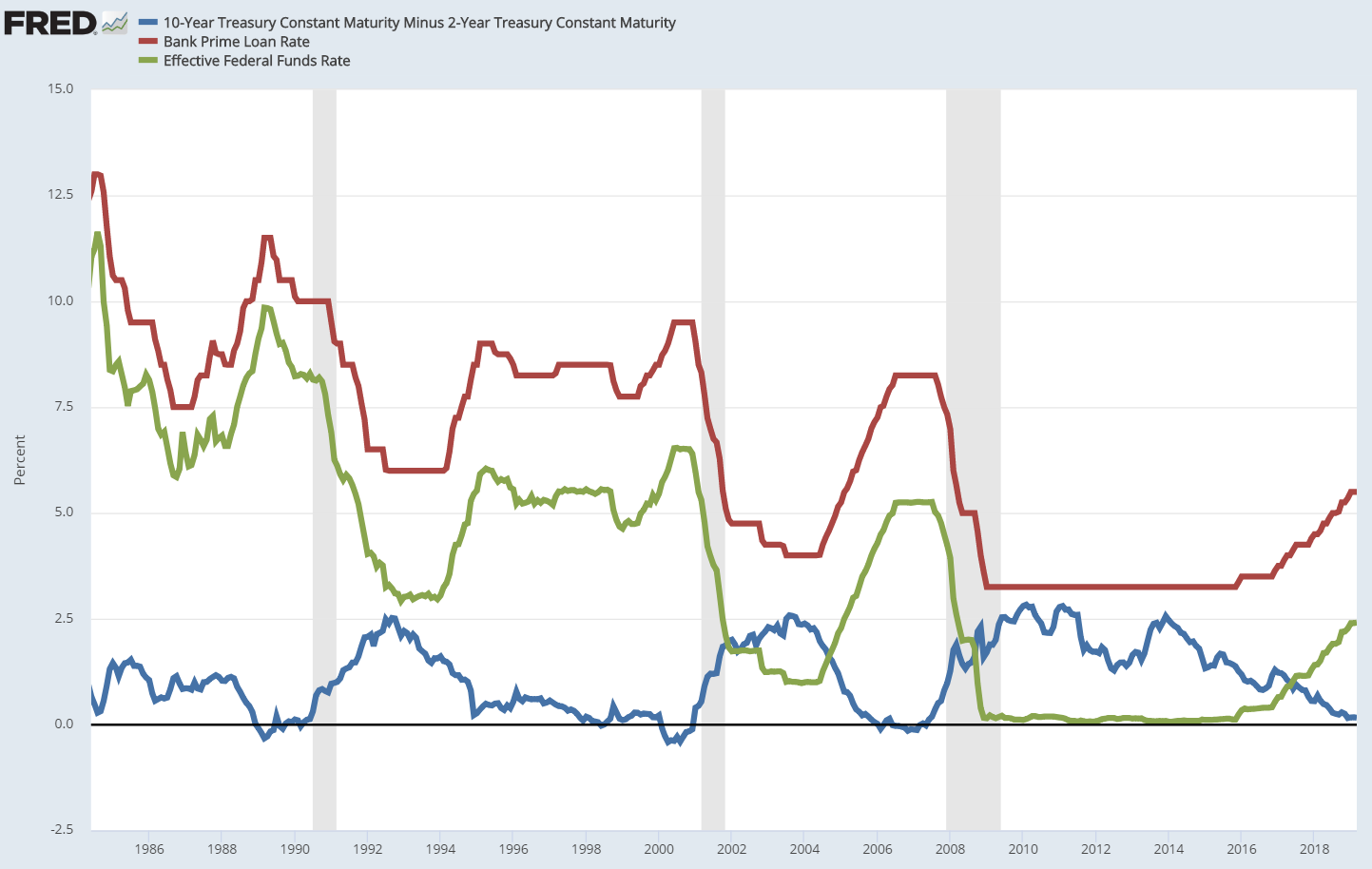
The primary source of the data comes from the Home Mortgage Disclosure Act (HMDA). The HMDA was enacted by Congress in 1975 and implemented by the Federal Reserve Board’s Regulation C. On July 21, 2011 the HMDA was transferred to the Consumer Financial Protection Bureau (CFPB)3. HMDA’s primary purpose 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 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 contain characteristics concerning a loan application for a borrower. For this analysis the data is subset to applications that were originated between 2004 and 2016. Using this population a subprime institution definition was constructed using a list from Wikipedia##. The institution name was associated to the LAR data using the key identifiers and data from the FFIEC##. The literature surrounding subprime definitions is extensive, two other methods are compared, “High Priced” and “HUD Subprime.” High Priced loans were originated with an interest rate that was 3% points over the industry standard (5% points for lines of credit). 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. HUD Subprime## definition used the High Priced methodology as well as called lenders which that specialized in subprime lending. If an institution classified itself as originating 50% or more subprime loans it was marked as a HUD Subprime institution. The HUD stopped producing their definition of subprime investors in 2006.

Table 1 represents the frequency of total mortgages and the three subprime definitions of loans between 2004 and 2007 reported in the HMDA LAR data. After 2005, the HUD stopped producing their version of the subprime definition. The failed institution (Subprime Institution) is classifies a considerably lower number of units in 2004 and 2005 than both the HUD Subprime and the HMDA higher priced loans. In 2007, the HMDA higher priced definition decreased by 50% from the preceding year. This is likely due to the classification of these loans and their susceptibility to the rate environment. As Mayer and Pence# *~~(for footnote,~~* [*~~https://www.federalreserve.gov/pubs/feds/2008/200829/200829pap.pdf~~*](https://www.federalreserve.gov/pubs/feds/2008/200829/200829pap.pdf)*~~)~~* found that the flattening of the yield curve will cause shifts in the rate spreads. In general if the rate spread between the long term and short term maturities converges loans will not be able to be classified as HMDA higher priced. The rational for this being it makes the 3 percentage points (5 percentage points for Lines of Credit) for first lien loans almost non-existent. This can be seen by comparing the bank prime loan rate to the federal funds rate. The bank prime loan rate is the rate in which a bank prices loans for customers. The federal funds rate is the rate in which banks price loans to other banks or lending institutions and the yield curve (here represented as the 10 year treasury less the 2 year) shows the flattening environment.



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).



As the yield curve flattens in 2007 it causes the spread between the 10 year and 2 year to be close to zero or negative which dramatically reduces the number of loan originations that can be priced 3 percentage points above this rate.

Table 1a depicts a cross tab of the Subprime definitions. The discrepancies between the definitions are vast, in general approximately 84% of the Subprime Institution definition is not covered by either High Priced or HUD Subprime. There is a 68% match rate between HUD Subprime and High Priced, however High Priced defines 70% more Subprime loans than the HUD.



The majority of the predictor variables for this analysis were sourced from the LAR data files.

1. Sold To: Many loan originating institutions originate, then sell to larger investors. If a loan was sold in a given year to Fannie Mae, Ginnie Mae, Freddie Mac or Farmer Mac, these were combined and marked as GSE (Government Sponsored Entities). Commercial Bank, such as Bank of America, Wells Fargo. Held For Investment, the originating institution holds the loan on their book. Mortgage Bank or Insurance company, such as AIG. All other types of sales are classified as Other.
2. Purpose of loan, Refinance versus Purchase.
3. Race of the Purchaser, Not reported was marked as Other.
4. Sex of the purchaser, not reported was marked as Other.
5. Preapproval, if preapproval of financing was requested, not reported was marked as Other.
6. Lien position for the loan
7. Income Spread, which is the stated income of the borrower differenced from the median income of all reported originations in the LARS files for the state. This difference is then put into quintiles. Dollars used were adjusted to 2001 dollars.

The predictor variables attributes are depicted in table 2. Reported by time dimension 2004-2007 and 2008 through 2016 between the categorical identifiers which will be used in the model. Post 2008 loans sold to Private Securities consumed 1.1% of the originations, before 2008 was 4.4%. Another considerable shift was to GSE, from a percentage standpoint, loans sold to the GSE investors increased over 100%. As the FRED graph shows, the Prime rate decreased substantially 2008 onward causing a considerable shift in the originations versus refinance between the two periods as home owners would be able to benefit by refinancing for lower monthly payments. Second lien loans pre and post shows another considerable shift, 18% of originations were second lien between 2004 through 2007 versus 5% 2008 onward. Likely, this is due to tightening of lending standards post the great recession.



**Regression Analysis:**

Given the categorical attributes of the data, a simple logit model was used to generate inferences on the data and to give the probability of a Subprime Instution. The data used to fit the model was between 2004 and 2007. As mentioned, this was a heavy period for subprime mortgage origination. The model was built using consisting of approximately a 50/50 split between the defined Subprime Institution estimates and Non Subprime Institution.





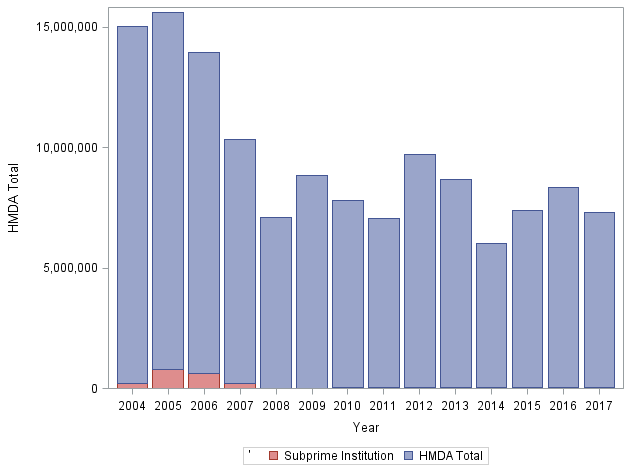
The above regression show interesting characteristics. First off, its rather odd that a first lien loan would have roughly be 2.5 times more likely than to be classified as a subprime than a second lien loan. This is likely due to the overwhelming majority of loans are first lien in the data, further the subprime institutions defined in this work must have been more prodominent in first liens. Table 2 simple statistics show only 366 thousand were classifed as second lien subprime institution. Loans sold to a Private Security was heavily related to a subprime institution. From table 2, we note that this population roughly disapears in after 2008 which would make the ability to use the fit model questionable at future predictions.

In hind sight, and interaction term would likely assist in the model, being loan purpose by lien position. The rational being its rather odd to see the loan purpose of Purchase being less likely to be subprime than refinance and a first lien loan being 2.5 times more likely to be a subprime institution. Here, the rational is that purchased loans are more likely to be a first lien and the modeled results show a contradiction from this when the two variables are modeled individually.

Lastly, preapproval seems to be standard when purchasing a home. As the data shows from the simple statistics one would expect preapproval to be a majority, however its rare and seems to be trending downward post 2007 in the data. Only 2.8% of the originations post 2008 were classifed as having preapproval requested.

Using the fitted model built using data from 2004 to 2007, new data was scored, 2008 through 2016 with the prior probabilities adjusted. Scoring the data gives us the probability of a loan being classified as Subprime Institution. With this, we can evaulate the models ability to predict the subprime institution between 2008 and 2016.

*Plot 2: Subprime Instution Predictions*

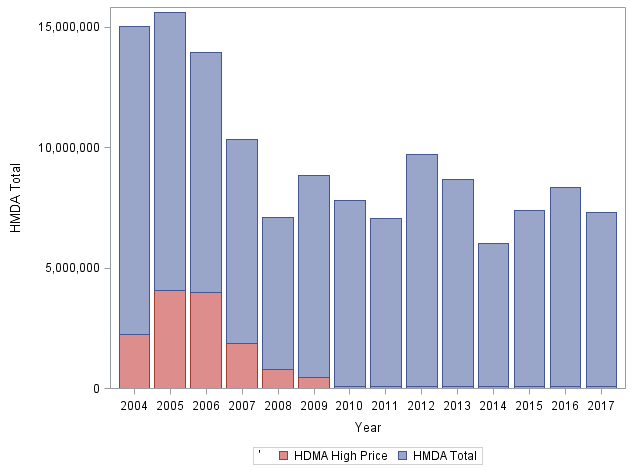


The model fit is lower than the actual Subprime Instituion events marked in the data. Between 2004 and 2007 we see that 2005 and 2006 had higher events than 2004 and 2007 which is consistent with the HMDA higher priced. In 2008, the steep decline is also present, again inline with the HMDA higher priced defintition.

More than likely the drop off in the subprime institution predictions is due to a substantial shift in the data represetend in the LAR’s files. From table 2, 46% of the defined subprime instution loans were marked in the Private Security which after 2008 loans, for the most part were no longer being sold to this category.

As mentioned, the HMDA higher priced loans are directionally consistent with the subprime instution definition throughout the time frame of 2004-2016.

*Plot 3: HMDA Higher Priced Loans*



Notes: Italics are reported p-values

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

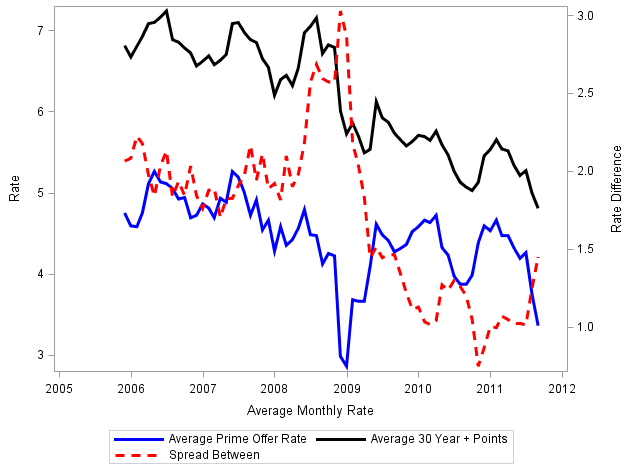
**Conclusion:**

Overall, the HMDA high priced and subprime failed institution definitions resulted in similar responses. Being pre 2007 there was a substantial increase in the amount of subprime lending and post 2008 subprime lending roughly disappeared. Primarily the disappearance is due to the tightening of lending standards and removal of riskier products after the financial recession.

However, it should be noted that the data in the LAR’s files is sporadic during the time window of this analysis. Concerning the HMDA high price definition, there were changes to the methodology for classifying higher priced loans. This work used 3 percentage point spread for first lien loans and 5 percentage point spread for second lien loans. Generally speaking, the spread is calculated using the Consumer Financial Protection Bureau (CFPB) Average Prime Offer Rate (APOR). The APOR is calculated as the term, example a 30 year fixed mortgage rate and the additional points associated to the loan. Additional points are fees associated to origination reflected as a percentage of the price (interest rate) of the loan. The CFPB’s definition at the time of this writing utilized an APOR derived from Freddie Mac (FHLMC), however the analysis in this work used the 30 year commitment rate. Additionally, the HMDA high priced thresholds are stricter with FHLMC’s definition, being 1.5 percentage points over the APOR would be classified as a HMDA high price for first lien loans, second lien are 3.5 percentage points over APOR ~~(~~[~~https://www.consumerfinance.gov/policy-compliance/rulemaking/regulations/1026/35/#a~~](https://www.consumerfinance.gov/policy-compliance/rulemaking/regulations/1026/35/#a)~~).~~

The rapid reduction of the HMDA high price population after 2008 is likely due to the lower rate enviorment, as Mayer and Pence described. One can make inferences on this by reviewing the CFPB’s APOR rate during the time window and comparing this to the 30 year commitment rate. Plot 4 depicts the APOR and 30 year commitment rate over a partial time window for this analysis.

*Plot 3: HMDA Higher Priced Loans*



The spread between the APOR and the 30 year commitment rate (includes the additional loan points) decreased by half after 2009. The decrease, on average would limits the amount of HMDA high priced loans in the housing market.

Given the definition change in the HMDA high priced loans (3 🡪 1.5percentage points) it would be beneficial to recalculate the definition. However, the LAR’s data does not provide the origination interest rate provided to the borrower. Table 4 is a depiction of populated and missing rate spreads in the LARs data from 2004 to 2017. It is good to note here that if the rate spread is over the HMDA high priced definition it will be populated. Table 4 shows the changes in the HMDA high price definition over time. In 2009 onward the spread was reduced to below 3 percentage points, however for the purpose of this paper HMDA high priced loans were forced to be 3 or 5 percentage points over the APOR.



**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)

For the APOR Spread Chart:

/\* https://www.ffiec.gov/ratespread/newcalchelp.aspx#4 \*/

<https://fred.stlouisfed.org/series/MORTPTS30US>

<https://fred.stlouisfed.org/series/MORTGAGE30US>