Quality check

(a) Quality assessment on Loan_Amount_000 and Respondent_Name

We define those loans whose loan amount greater than (75% quantile + 3*inter quantile range of all loans) in its' own MSA or MD area are the statistical outlier, these loans are regarded as large loan and could be of poor quality. (saved in the data frame 'large_loan_byMSA').

MSA_MD_Description	Loan_Amount_000	Respondent_Name_TS	As_of_Year
<fctr>	<dbl>	<fctr>	<int>
CHARLESTON, WV	540	WESBANCO BANK, INC.	2012
CHARLESTON, WV	595	WESBANCO BANK, INC.	2012
CHARLESTON, WV	622	WESBANCO BANK, INC.	2012
CHARLESTON, WV	990	WESBANCO BANK, INC.	2012
CHARLESTON, WV	542	PIONEER WV FCU	2012
CHARLESTON, WV	1210	THE HUNTINGTON NATIONAL BANK	2012
CHARLESTON, WV	945	THE HUNTINGTON NATIONAL BANK	2012
CHARLESTON, WV	485	THE HUNTINGTON NATIONAL BANK	2012
CHARLESTON, WV	652	THE HUNTINGTON NATIONAL BANK	2012
CHARLESTON, WV	621	THE HUNTINGTON NATIONAL BANK	2012

1-10 of 11,854 rows Previous 1 2 3 4 5 6 ... 100 Next

The data frame 'large_loan_byRespondent' then shows the total number of those large loans owned by each respondent.

Respondent_Name_TS	count
<fctr>	<int>
WELLS FARGO BANK, NA	1589
SUNTRUST MORTGAGE, INC	1235
MEGACHANGE MORTGAGE	900
BANK OF AMERICA, N.A.	704
JPMORGAN CHASE BANK, NA	529
CITIBANK, N.A.	477
BRANCH BANKING AND TRUST CO	439
FIRST SAVINGS MORTGAGE CORPORA	261
PNC BANK N.A.	216
GEORGE MASON MORTGAGE LLC	213

1-10 of 497 rows Previous 1 2 3 4 5 6 ... 50 Next

Loan to income ratio (Loan_Amount_000/Applicant_Income_000) would also be an important column because this reflects applicant's ability to pay back.

The critical value for statistical outliers for the loan to income ratio equals to 8.69 (75% quantile + 3*inter quantile range). Which means the applicant will not be able to pay back the principal in almost nine years' time, we assume these loans as risky loans. We found those loans and saved in the data frame ‘risky_loan’. These loans are regarded as the risky loans.

Loan_Amount_000 <dbl>	Respondent_Name_TS <fctr>	loan_to_income_ratio <dbl>
505	PUTNAM COUNTY BANK	12.62
158	JPMORGAN CHASE BANK, NA	8.78
387	CNB BANK INC.	12.48
78	CNB BANK INC.	11.14
254	CITIBANK, N.A.	11.04
169	BANK OF AMERICA, N.A.	8.89
282	PNC BANK N.A.	9.72
174	GENERATION MORTGAGE COMPANY	8.70
151	ADVANTAGE BANK	9.44
119	HUNTINGTON FEDERAL SAVINGS	9.15

1-10 of 4,621 rows Previous 1 2 3 4 5 6 ... 100 Next

The data frame ‘risky_loan_byRespondent’ then shows the total number of those risky loans owned by each respondent.

Respondent_Name_TS <fctr>	count <int>
BANK OF AMERICA, N.A.	769
MEGACHANGE MORTGAGE	607
JPMORGAN CHASE BANK, NA	501
QUICKEN LOANS	422
NATIONSTAR MORTGAGE LLC	277
PNC BANK N.A.	250
CITIBANK, N.A.	184
SUNTRUST MORTGAGE, INC	161
WELLS FARGO BANK, NA	120
ALLY BANK	106

1-10 of 266 rows Previous 1 2 3 4 5 6 ... 27 Next

Combine (a) and (b), we then found those loans that are both large and risky, saved in data frame ‘large_risky_loan’. These loans are regarded as poor quality loans.

Loan_Amount_000 <dbl>	Respondent_Name_TS <fctr>	loan_to_income_ratio <dbl>
1500	THE BANK OF ROMNEY	14.85
1100	CITIBANK, N.A.	68.75
880	BANK OF AMERICA, N.A.	10.35
2345	JPMORGAN CHASE BANK, NA	9.20
1575	JPMORGAN CHASE BANK, NA	20.72
2300	GOLDMAN SACHS BANK USA	39.66
2300	PHH MORTGAGE CORPORATION	39.66
1350	THE BANK OF GLEN BURNIE	10.55
691	THE COLUMBIA BANK	19.74
1082	SUSQUEHANNA BANK	37.31

1-10 of 403 rows Previous 1 2 3 4 5 6 ... 41 Next

We thought there should be a big intersection between large loans and risky loans. However, there are 11,854 large loans and 4,621 risky loans, and only 403 risky large loans. Which means only around 10% of risky loans are from large loans. The data frame 'risky_loan_distribution' shows the distribution of risky loans in different loan groups:

low_loan_pct <dbl>	middle_loan_pct <dbl>	high_loan_pct <dbl>	extremly_high_loan_pct <dbl>
14.13	58.15	12.92	14.8

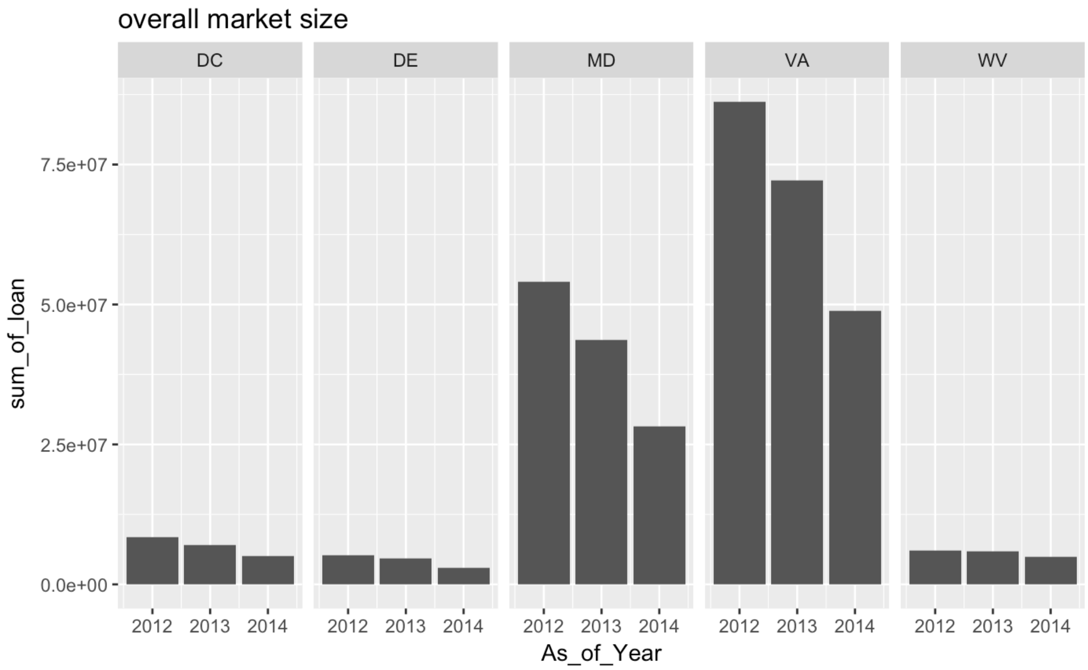
1 row

The table shows that almost 60% of risky loans come from the middle group of loan! These statistics demonstrate that loan amount is not always correlated with risk. For example, jumbo loans are intended for excellent borrowers with excellent credit looking to finance loan amounts greater than the amount allowed by Fannie Mae or Freddie Mac, according to our findings, jumbo loans are not having to be more risker than conforming loans.

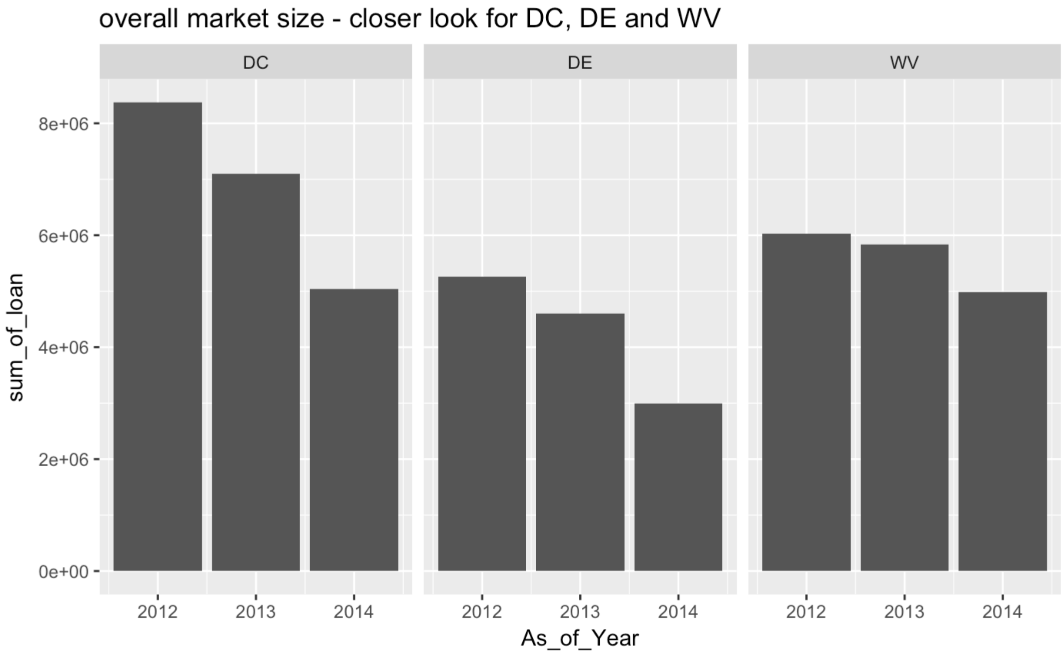
The market size for the jumbo market is the only market that shrinks slowly, and our quality assessment shows that most of the risky loans are coming from the middle group of loans. Therefore, my recommendation for any company that want to enter this home loan market could focus its business on the jumbo market.

Hypothesis: It is risky to enter this home loan market.

Visualization:
We can see from graph 1 and graph 2 that the market size decreased sharply from 2012 to 2014, except for WV, the market size in which only decreased slowly.



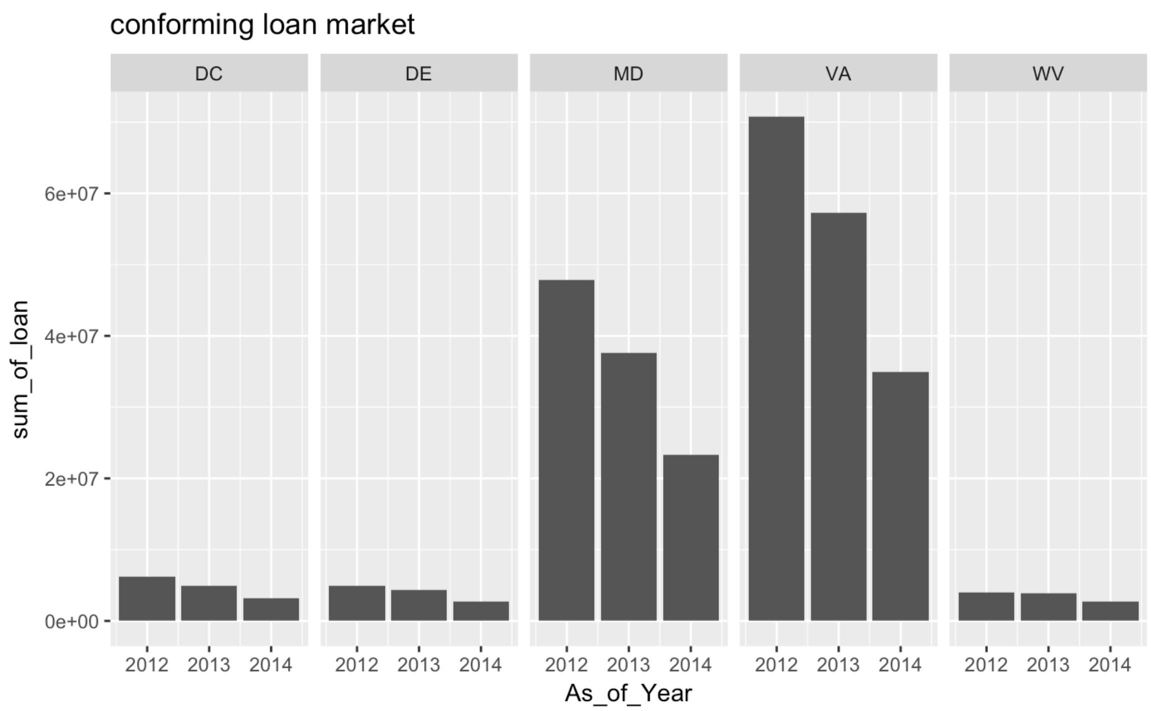
graph 1



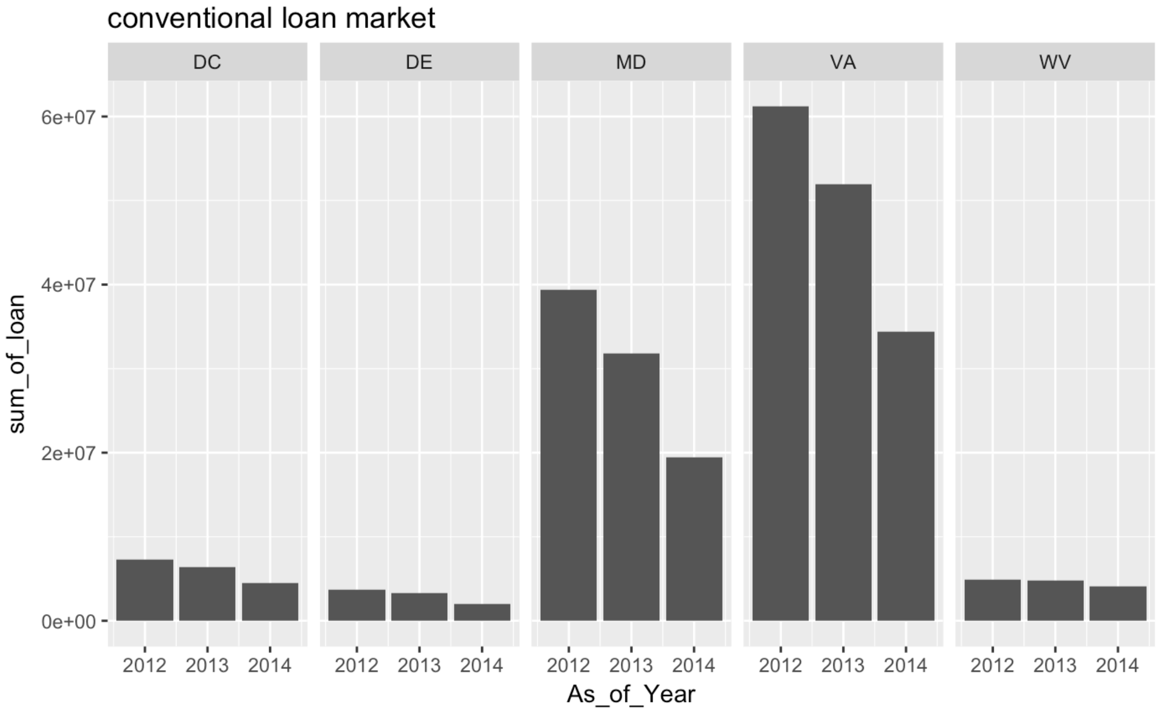
graph 2

Let’s take a look at the market size in both conventional and conforming market:

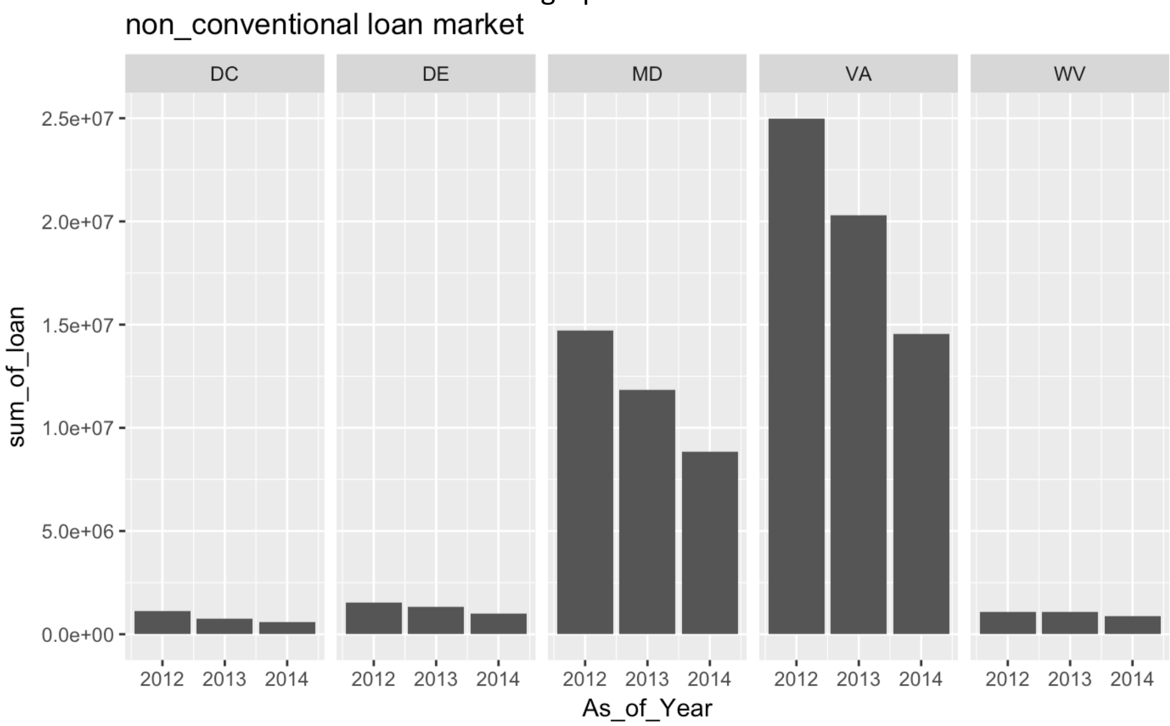
Graph 3,4,5 are the market size for conforming, conventional and non-conventional market; we can see a similar trend of decreasing as the overall market size from 2012 to 2014.



graph 3

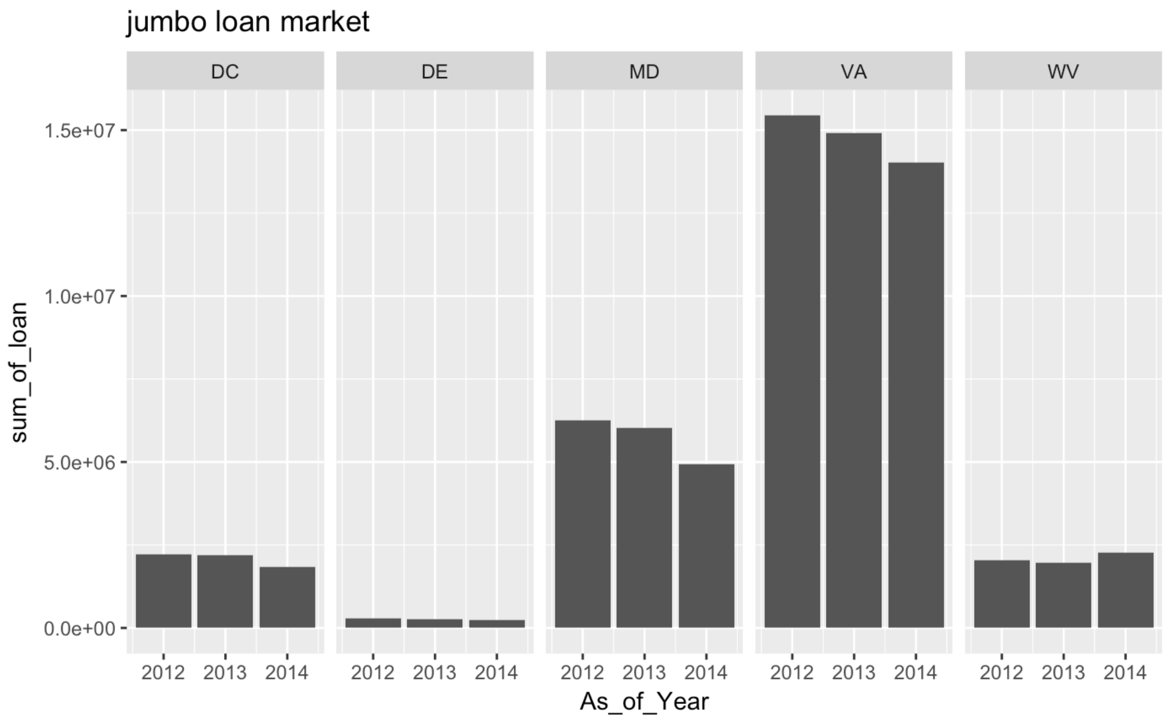


graph 4

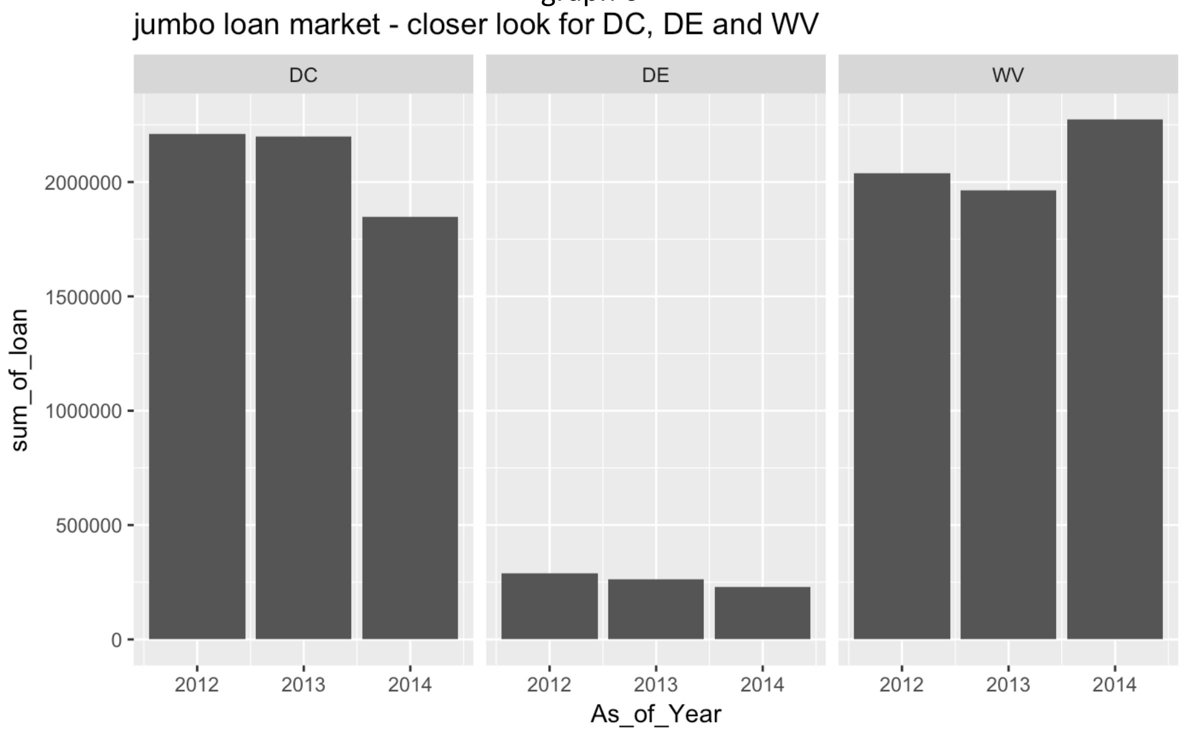


graph 5

However, when we take a look at the jumbo market (Graph 6 and 7), the market size for jumbo market shrink slightly and even increased in WV.



graph 6



graph 7

To prove or disprove our hypothesis we need more information: Although the overall market size is shrinking, the jumbo loan market decreases slowly compared to other markets, and even increase in WV. Therefore, we need more information about what type of loan Change financial want to focus on for its business and take the profitability and risk of jumbo loans into consideration before we can prove whether it is a good time for Change Financial to enter a new geography.

(c) Summarization of key insights:
Rank of the market size: VA > MD > DC > WV > DE.
Overall, the demand for the home loan decreased sharply from 2012 to 2014, except for WV, the market size in which only decreased slowly. However, in the jumbo loan market, the trend of decreasing is much more gentle. However, It’s worth noting that jumbo market size is relatively small when compared to both conforming and conventional markets.