HMDA CLL Analysis

Yiyang Han

June 29, 2024

1 Key Takeaways

1. For both regressions: results might be a bit counter-intuitive when we include all region (or bank) $\times$ year combinations into regressions. This is because we have relatively few regions / banks that experienced CLL limit change / exposed to CLL limit change shock.

2. When we look at intensive margin, we found CLL increase is correlated with higher total mortgage growth in this region.

3. When we look at intensive margin, we found exposure to CLL increase shock is correlated with higher residential mortgage / asset ratio, and securities / asset ratios for banks.

2 Mortgage growth and CLL change regressions

2.1 Panel Data Construction

We gather loan amount, lender identifier, state code, county code, purchaser type variables from HMDA. Again, we only focus on originated loans in HMDA. We consider loans that were not sold (retained mortgage / balance sheet mortgage) as loans whose purchaser type is 0, 6, or 8.
We then merge HMDA data with Avery data (through year and lender identifier) and CLL data (through state code, county code, and year) to obtain bank indicator (for each individual lender) and CLL (for each region).
We then aggregate loan amount at region-year level to obtain the following variables:

1. Total mortgage

2. Bank total mortgage (sum of mortgage amount lent by bank lenders)

3. Bank retained mortgage (sum of retained mortgage amount lent by bank lenders)

4. Number of bank lenders

Using these variables, we are able to calculate the following variables needed for the regression:

1. $\Delta CLL = CLL_y - CLL_{y-1}$

2. Bank retained mortgage share change = $\frac{\text{Bank retained mortgage}_y}{\text{Bank total mortgage}_y}$ - $\frac{\text{Bank retained mortgage}_{y-1}}{\text{Bank total mortgage}_{y-1}}$

3. Bank retained mortgage growth = $\frac{\text{Bank retained mortgage}_y - \text{Bank retained mortgage}_{y-1}}{\text{Bank retained mortgage}_{y-1}}$ (Bank total mortgage growth and Total mortgage growth are calculated analogously).

2.2 Summary Statistics

2.2.1 Extensive Margin dataset

The 4 dependent variables at winsorized at [0.01, 0.99] level per year.

Variable	Average
Retained mortgage share change	0.001
Bank retained mortgage growth	0.054
Bank mortgage growth	0.016
Total mortgage growth	0.051
Change in CLL	674.115

2.2.2 Intensive Margin dataset

Variable	Average
Retained mortgage share	0.003
Bank retained mortgage growth	0.069
Bank mortgage growth	-0.003
Total mortgage growth	0.022
Change in CLL	5445.020

2.2.3 Propensity Score Matching

We match regions from the treatment group (with non-zero CLL change) and control group (with 0 CLL change) using the following characteristics:

1. Number of Banks
2. CLL
3. Total Mortgage

This is a summary statistics for both groups:

Treatment	Number of Regions	# Banks in the region	CLL	Total Mortgage
F	3584	33.03	436758.58	772386.86
T	3584	52.67	441566.18	956455.66

As we can see, the number of banks in the region and total mortgage are a bit off.

2.3 Regressions

We include two way fixed effects on region and year. Standard errors are clustered at Region level. In addition, we include control on change in number of banks in a specific region from last year to this year. We tried two regressions: one with extensive margin (in which all region $\times$ year combinations enter the regression), and the other with intensive margin (in which only those region $\times$ year combinations who experienced CLL change enter the regression).
The reason that we tried intensive margin is a majority of region $\times$ year combinations do not have a CLL variation.

2.3.1 With all region × year

	Bank retained mortgage share change	Bank retained mortgage growth	Bank total mortgage growth	Total mortgage growth
Δ CLL (in millions)	0.024	-0.141	-0.272*	-0.245*
	(0.045)	(0.176)	(0.107)	(0.100)
Δ Number of Banks	0.000***	0.008***	0.009***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	28872	28908	28930	28939
R-squared	0.163	0.164	0.305	0.385

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.3.2 When we only look at intensive margin

	Bank retained mortgage share change	Bank retained mortgage growth	Bank total mortgage growth	Total mortgage growth
Δ CLL (in millions)	-0.012	0.073	0.119	0.341*
	(0.065)	(0.191)	(0.175)	(0.164)
Δ Number of Banks	0.000	0.003*	0.002*	0.004***
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	3578	3578	3582	3584
R-squared	0.971	0.990	0.978	0.962

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.3.3 Propensity Score Matching

	Bank retained mortgage share change	Bank retained mortgage growth	Bank total mortgage growth	Total mortgage growth
Δ CLL in 1 M	-0.030	-0.395+	-0.429**	-0.230
	(0.057)	(0.226)	(0.163)	(0.153)
Δ Number of Banks	-0.001**	0.015***	0.014***	0.010***
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	7104	7140	7156	7161
R-squared	0.274	0.298	0.317	0.352

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.4 Takeaways

1. When we look at intensive margin, we can observe total mortgage growth positively correlated with increase in conforming loan limits (significant at 5%). The magnitude is relatively moderate: 1 million increase in CLL is correlated with 34.1% increase in total mortgage growth.

2. Under intensive margin regression, we can also see the share of retained mortgage negatively correlated with CLL increase (because loans are easier to securitize), and bank total mortgage lending growth positively correlated with CLL increase. Although those two correlations are not statistically significant.

3. Under intensive margin, the explanatory power of regression increased significantly, suggesting changes in CLL can account for a significant proportion of variations in bank retained mortgage share, bank total mortgage growth, and total mortgage.

4. The propensity score matching on regions works less well: the number of banks and total mortgage between two samples are not very well aligned. In the regression, we can no longer see the correlation between change in CLL and total mortgage growth. We start to observe a negative correlation between increase in CLL and bank total mortgage growth. This could due to more competition from the shadow bank.

3 Exposure Regressions

3.1 Panel Data Construction

We gather loan amount, lender identifier, state code, county code variables from HMDA. We only focus on originated loans in HMDA.
We then merge HMDA data with Avery data (through year and lender identifier) to select loans from bank lenders only. We also merge with CLL data (through state code, county code, and year) to obtain CLL (for each region x year).
We then aggregate individual loan data to bank-year level data by calculating the following variables:

1. Total mortgage: the sum of mortgages that each bank lent each year
2. Exposure: the bank's last year mortgage lending's exposure to CLL change

Specifically, the exposure variable is calculated by:

$$\sum_{i, i \in \text{Bank's originated mortgage last year}} (\text{Loan Amount}_i \times \text{I}_{\text{Conforming this year}} - \text{Loan Amount}_i \times \text{I}_{\text{Conforming last year}})$$

Where the summation occurs on all loans across different regions (with various CLLs) the bank lent **last year**. The underlying logic for this calculation is that we sum up bank's exposure to CLL change on individual loans (depending how the region's CLL change). For each loan, the exposure is $(\text{Loan Amount} \times \text{I}_{\text{Conforming this year}} - \text{Loan Amount} \times \text{I}_{\text{Conforming last year}})$:

Conforming This Year	Conforming Last Year	Caused by	Exposure	Implication for the Bank
T	T		0	
F	F		0	
T	F	Increase in CLL	+ Loan Amount	More conforming loans
F	T	Decrease in CLL	- Loan Amount	Less conforming loans

We then obtain the bank's exposure shock as

$$\frac{\text{Exposure calculated above}}{\text{Bank's total mortgage last year}}$$

We then merge this with bank call report, from which we derive the following ratios:

1. Securities / Assets: we did two versions, as we did in the call report data collection task.
   (a) Version 1: (Total securities: Held-to-maturity, Fair value + Total securities: Available-for-sale, Fair value) / Assets, this version better reflects the securities held by banks measured at the market value.
   (b) Version 2: (Total securities: Held-to-maturity, Amortized Cost + Total securities: Available-for-sale, Fair value) / Assets, this version values the securities held by banks better if banks hold those Held-to-maturity securities to maturity.
2. Loans / Assets: Total Loan / Total Assets
3. Mortgages / Assets: we calculate mortgage as the sum of Loans secured by real estate each bank hold.

4. Residential Mortgages / Assets: we calculate the residential mortgage as the sum of following items: Loan secured by 1-4 family residential properties (RCON1797 + RCON5367 + RCON5368)

5. Debt Securities / Assets: we consider debt securities that finance household and firm debts. It consists of the following items:

We use those ratios from Q4 each year.

3.2 Summary Statistics

3.2.1 Extensive Margin dataset

Variable	Average
Exposure Shock	0.00606
Mortgage/Asset	0.518
Residential Mortgage / Asset	0.212
Loan / Asset	0.665
Securities / Asset (Version 1)	0.194
Securities / Asset (Version 2)	0.194
Debt Securities / Asset (Version 1)	0.096
Debt Securities / Asset (Version 2)	0.092

3.2.2 Intensive Margin dataset

Variable	Average
Exposure Shock	0.02730
Mortgage/Asset	0.555
Residential Mortgage / Asset	0.236
Loan / Asset	0.701
Securities / Asset (Version 1)	0.171
Securities / Asset (Version 2)	0.172
Debt Securities / Asset (Version 1)	0.107
Debt Securities / Asset (Version 2)	0.100

3.2.3 Propensity Score Matching [Updated]

We match banks from the treatment group (with non-zero exposure) and control group (with 0 exposure) within the same year using the following characteristics:

1. Equity / Asset ratio
2. Deposit / Asset ratio
3. Residential loan / Total loan ratio
4. State

This is a summary statistics for propensity score matching for **2009-2017** (during this period, all treated banks are matched):

Treatment	Number of Banks	Residential Mortgage / Total Loan	Deposit / Asset	Equity / Asset
F	2056	0.372	0.802	0.111
T	2056	0.386	0.797	0.109

This is a summary statistics for propensity score matching for **the whole sample** (note that since 2018, not all treated banks are matched):

Treatment	Number of Banks	Residential Mortgage / Total Loan	Deposit / Asset	Equity / Asset
F	6406	0.342	0.832	0.111
T	6406	0.357	0.810	0.109

3.3 Regressions

We include two way fixed effects on state (where the bank's headquarter is located) and year. Standard errors are clustered at Bank level.
We tried two regressions: one with extensive margin (in which all bank × year combinations enter the regression), and the other with intensive margin (in which only those bank × year combinations who experienced non zero exposure shock change enter the regression).
The reason that we tried intensive margin is a majority of bank × year combinations do not have a exposure shock.

3.3.1 With all bank × year

	Mortgage Asset	Res Mortgage Asset	Loan Asset	Sec (1) Asset	Sec (2) Asset	Debt Sec (1) Asset	Debt Sec (2) Asset
Exposure Shock	0.195***	0.203***	0.138**	-0.125**	-0.121**	-0.009	0.006
	(0.048)	(0.042)	(0.043)	(0.039)	(0.040)	(0.033)	(0.032)
Observations	47773	47773	47773	47773	47773	47773	47773
R-squared	0.175	0.244	0.088	0.061	0.061	0.039	0.041

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3.2 When we only look at intensive margin

	Mortgage Asset	Res Mortgage Asset	Loan Asset	Sec (1) Asset	Sec (2) Asset	Debt Sec (1) Asset	Debt Sec (2) Asset
Exposure Shock	-0.047	0.252***	-0.213***	0.094*	0.099*	-0.021	-0.005
	(0.050)	(0.052)	(0.047)	(0.042)	(0.042)	(0.038)	(0.037)
Observations	10598	10598	10598	10598	10598	10598	10598
R-squared	0.198	0.234	0.129	0.078	0.080	0.065	0.077

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3.3 Propensity Score Matching [Updated, 2009-2017]

	Mortgage Asset	Res Mortgage Asset	Loan Asset	Sec (1) Asset	Sec (2) Asset	Debt Sec (1) Asset	Debt Sec (2) Asset
Exposure Shock	-0.074	0.083	-0.104*	0.121**	0.120**	0.062+	0.062+
	(0.063)	(0.053)	(0.053)	(0.046)	(0.046)	(0.037)	(0.037)
Observations	4112	4112	4112	4112	4112	4112	4112
R-squared	0.228	0.297	0.122	0.074	0.073	0.079	0.079

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3.4 Propensity Score Matching [Updated, whole sample]

	$\frac{\text{Mortgage}}{\text{Asset}}$	$\frac{\text{Res Mortgage}}{\text{Asset}}$	$\frac{\text{Loan}}{\text{Asset}}$	$\frac{\text{Sec (1)}}{\text{Asset}}$	$\frac{\text{Sec (2)}}{\text{Asset}}$	$\frac{\text{Debt Sec (1)}}{\text{Asset}}$	$\frac{\text{Debt Sec (2)}}{\text{Asset}}$
Exposure Shock	0.150**	0.119*	0.160***	-0.114**	-0.111**	0.032	0.032
	(0.053)	(0.048)	(0.047)	(0.043)	(0.043)	(0.033)	(0.033)
Observations	12812	12812	12812	12812	12812	12812	12812
R-squared	0.220	0.257	0.121	0.077	0.078	0.054	0.054

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.4 Takeaways

1. When we extend the sample to include observations from 2018-2022, we find a lot more banks experiencing CLL exposure shocks. This is because the CLL limit change during those years is more frequent.

2. When we look at the sample of all banks / propensity score matching for the whole sample, we see strong correlation between positive CLL shock exposure and higher mortgage / asset, residential mortgage / asset, and loan / asset ratios. However, it is a bit wired that there is a negative correlation between positive CLL exposure shock and securities / asset ratio.

3. When we look at intensive margin, we can still observe the strong positive correlation between CLL exposure shock and residential mortgage / asset ratio. We also see a positive correlation between CLL exposure shock and securities / asset ratio. It is a bit wired to observe a correlation between positive CLL exposure shock and lower loan / asset ratio.

4. When we look at propensity score matching for 2009-2017, we can observe a strong, positive correlation between CLL exposure shock and securities / asset ratio. We can also observe a weak correlation between CLL exposure shock and debt securities / asset ratio.