

HMDA Bank Year Panel Regressions

Yiyang Han

June 23, 2024

1 Key Takeaways

1. Compared to small banks, midsize and large banks' total mortgage and retained mortgage grow slower.
2. Compared to the period before 2010, banks on average have a higher growth rate on retained mortgage, but a lower growth rate on total mortgage.
3. Compared to small banks during 2010 - 2018, midsize banks tend to have higher growth rate on both total mortgage and retained mortgage. We can observe a similar trend in large banks, although the statistical significance and magnitude are lower.
4. Compared to small banks after 2018, midsize banks have a higher growth rate on both total mortgage and retained mortgage. Such correlation cannot be observed on large banks.
5. When we use total loan as the weights in regression with same specification, those correlations above still hold (most of them becomes stronger, both in terms of statistical significance and magnitude). In addition, large banks \times (After 2018) becomes significant. Large banks experience higher growth in total and retained mortgage, when compared to small banks during the same period.

2 Panel Data Construction

We consider credit unions as non-banks. Therefore, for this bank year regression, we first exclude credit unions from the panel data. We note two main things involved in our panel data construction:

2.1 Duplicated lenders

In the process of panel data construction, we noticed a problem of duplicated lenders. Basically, in certain years, there will be two or more entities (banks) under the same HMDA lender identifier.

2.1.1 Cause

In the Avery file, some HMDA lender identifier will be linked to two or more banks in certain years.

2.1.2 Impact

We need the total assets data for size classification. If there are two or more total assets data (coming from those linked entities), we can't decide which total asset data we should use for size classification.

2.1.3 Solution

There are 688 such duplicated year \times lender identifier combinations.

1. In 685 cases, those lenders would be classified as small banks, regardless of which total asset data we use. This means all linked banks for those lender identifiers are small banks. In this case, we just classify those banks as small banks.

2. In 3 duplicated cases, let's say an identifier X is linked to two banks A and B. In those 3 cases, we have bank A being a small bank and bank B being a midsize bank. To tell whose total asset we should use for size classification, I checked lender X's total retained loan. In all 3 cases, the retained loan is greater than the total asset of the smaller bank (bank A - small bank), but less than the total asset of the larger bank (bank B - midsize bank). Therefore, I classify those 3 lenders as midsize banks. In other word, it is impossible that those 3 lenders are smaller banks (bank A), since their retained loan in that year exceeds the matched smaller banks' assets. Therefore, they must be the larger bank (bank B), which makes them being classified as Midsize banks.

2.2 Bank-level growth rate construction

We calculate the annual mortgage growth rate for an individual lender as

$$\frac{\text{Mortgage Lending This Year}}{\text{Mortgage Lending Last Year}} - 1$$

There are two special cases (both of they are very rare):

1. A bank that does not lend in the middle year of a 3 year interval. For example, bank A lends in 2007, 2009, but not 2008 (total originated mortgage in 2008 = 0). In this case, its total mortgage lending growth rate would be -1 in 2008, and ∞ in 2009. In this case, we drop the growth rate for both years.
2. A bank that does not retain any mortgage on their balance sheet for certain years. For similar reason explained above, for some years, their retained mortgage lending growth rate would be -1 or ∞ . We also drop the growth rate for those cases.

Then, we winsorize both growth rates at Bank size group \times Year level at [0.05, 0.95].

After winsorization, we have the following summary statistics on average total loan growth rate, average retained loan growth rate, and average assets (in Billion):

3 Summary Statistics

Period	Bank Size	Total Loan Growth	Retained Loan Growth	Assets (in Billion)
Before 2010	Large	0.09	0.07	750.17
Before 2010	Midsize	-0.08	-0.10	122.14
Before 2010	Small	0.27	0.18	0.89
2010-2018	Large	0.15	0.16	1078.35
2010-2018	Midsize	0.07	0.16	111.13
2010-2018	Small	0.15	0.19	0.99
After 2018	Large	0.16	0.14	989.32
After 2018	Midsize	0.14	0.18	122.05
After 2018	Small	0.21	0.23	2.24

4 Regression Results

4.1 Regression without weights

	Annual Not Sold Mortgage Lending Growth Rate of Banks (1)	Annual Total Mortgage Lending Growth Rate of Banks (2)
Midsize	-0.275*** (0.057)	-0.336*** (0.053)
Large	-0.108** (0.034)	-0.175*** (0.032)
(2010-2018)	0.008 (0.008)	-0.119*** (0.007)
(After 2018)	0.045*** (0.010)	-0.053*** (0.009)
(2010-2018) x Midsize	0.244*** (0.071)	0.265*** (0.066)
(2010-2018) x Large	0.083+ (0.048)	0.176*** (0.045)
(After 2018) x Midsize	0.232** (0.081)	0.258*** (0.075)
(After 2018) x Large	0.026 (0.099)	0.124 (0.093)
Constant	0.180*** (0.007)	0.267*** (0.007)
Observations	49554	49692
R-squared	0.001	0.006

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

4.2 Weighted by total loan

	Annual Not Sold Mortgage Lending Growth Rate of Banks (1)	Annual Total Mortgage Lending Growth Rate of Banks (2)
Midsized	-0.445*** (0.031)	-0.439*** (0.030)
Large	-0.234*** (0.026)	-0.187*** (0.025)
(2010-2018)	0.066** (0.021)	-0.373*** (0.020)
(After 2018)	0.072*** (0.021)	-0.286*** (0.021)
(2010-2018) x Midsized	0.349*** (0.037)	0.316*** (0.035)
(2010-2018) x Large	0.669*** (0.029)	0.602*** (0.028)
(After 2018) x Midsized	0.423*** (0.037)	0.342*** (0.036)
(After 2018) x Large	0.103*** (0.030)	0.071* (0.029)
Constant	0.199*** (0.018)	0.567*** (0.018)
Observations	49554	49692
R-squared	0.052	0.036

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

In these regressions, larger lenders would carry more weights. [A note on number of observations: column (2) have more observations because a small amount of banks do not retain mortgage on their balance sheet in certain years. When they started to retain mortgages, they would have a retained mortgage growth rate of ∞ . We drop out those growth rates. However, those banks do have a reasonable growth rate on total mortgage in those years. Therefore, they will only enter regressions for column (2) on total loan growth.]