**Delay Your Rivals: Vertical Integration in Securitization and Lending Competition**

José D. Salas[†] & Brandon Zborowski[‡]

## Abstract

We study the effects of vertical integration in the securitization chain on lending competition in the commercial mortgage backed securities (CMBS) market. We show that lenders that are vertically integrated (VI) with the investment bank structuring the CMBS originate loans that have rate spreads that are 8bps lower and have a 10% shorter time from origination to securitization, conditional on observables. VI lenders also have larger market shares, consistent with their relatively lower spreads. To shed light on one mechanism, we show evidence that VI loans are prioritized over non-VI loans when constructing pools, which we call the "prioritization" channel, and this leads to shorter times to securitization. This difference in time to securitization gets passed through to higher rates, and we estimate that this explains about 12% of the difference in spreads. The spread and time to securitization results are stronger in quarters with low loan origination, which is exactly when we would expect this prioritization result to have stronger effect. Additionally, we show that prioritization channel impacts credit allocation pool diversification. By prioritizing their own loans, VI securitizers forgo pool diversification, which could have financial stability implications. We also show that investors value diversification, and so prioritization leads to higher yields on securities through lower diversification. Finally, we construct a model of vertical integration in securitization and lending competition that highlights the problem the securitizer faces. The VI lender balances the benefit of including their rivals' loans when constructing the CMBS pool, which increases pool diversification and therefore securitization securitization profits, with the benefit of prioritizing their own loans, which lowers their own costs compared to their non-VI rivals, due to relatively shorter time from origination to securitization.

[†]Department of Economics, Northwestern University - Email: jdsalas@u.northwestern.edu
[‡]Kellogg School of Management, Northwestern University - Email: brandon.zborowski@kellogg.northwestern.edu

# 1 Introduction

Securitization has become an increasingly prominent aspect of various lending markets, including auto loans, credit cards, residential mortgages, and commercial mortgages. Policymakers and regulators have viewed these markets with increased scrutiny and implemented a variety of reforms, specifically the Dodd-Frank Act in an attempt to better align incentives of various players in these markets. But one aspect of securitization that is often overlooked is how the securitization chain itself is organized in terms of the participants involved and the interactions between them.

Much of what is understood about securitization comes from residential mortgage market, where the majority of securities are sponsored by Government Sponsored Entities (GSEs) such as Fannie Mae and Freddie Mac. These entities purchase mortgages from originators, guarantee the mortgages, and sell securities to investors. This results in investors facing no default risk and lenders facing no consequences for poor performing mortgages, as long as the underwriting of the loans met the GSE standards. The agency residential mortgage backed security market is special and most securitization markets differ in an important way. Specifically, most securitization markets require investment banks to acquire assets, construct securities, and sell the securities to investors. In this paper we explore the prevalence of vertical integration of the participants in the conduit Commercial Mortgage Backed Securities (CMBS) market[1]. When a lender is securitizing with an affiliated investment bank, which we define as vertically integrated (VI), their loans will be treated differentially in the structuring of the security, leading to different loan rates and market share, relative to lenders not affiliated with the banks structuring the CMBS issuance.

In our paper, we first document a few stylized facts related to pricing and market shares of VI lenders, and then provide evidence for one mechanism driving these results, which we call the "prioritization" channel. We explore implications of this "prioritization" channel for credit outcomes and CMBS security construction and the profitability of CMBS securitization. Then we outline a model of vertical integration in securitization and lending competition to show the tradeoffs at play, with the ultimate goal of structurally estimating the model to conduct welfare analyses in various counterfactuals. Lastly, we address concerns to our main results by exploring alternative explanations.

[1] GSEs are also active in a different segment of the CMBS market, specifically CMBS full of multifamily properties

First, we show that that VI loans have interest rate spreads that are 8bps lower than non-VI loans, conditional on observables. One explanation for this is that VI lenders have greater control of the securitization process relative to non-VI lenders, which is one of the main benefits of vertical integration to the lender. They have the final say as to whether a loan will be placed in a specific pool and they use it to prioritize their own loans, ensuring their loans makes it into the upcoming pool instead of having to wait for a future pool. This in turn lowers their warehousing costs, which is the risk that loans could become non-performing or that the interest rate environment could move against the lender before the loans have been securitized, relative to their non-VI competitors. These lower costs are passed through into lower spreads. Consistent with lower spreads, VI lenders also have higher market shares. There are other reasons that also explain the difference in rates, such as greater efficiency of VI lenders and double marginalization[2] of non-VI lenders , but the focus of our paper is to highlight the importance of the "prioritization" channel, which we interpret as evidence of vertical foreclosure.

Second, we show that VI loans have a shorter length from origination to securitization of about 7 days, or 10% of the average, relative to non-VI loans. As mentioned earlier, non-VI lenders have less control over whether their loans end up in any particular pool, so if a non-VI lender gets one of their loans rejected from a pool, they will have to hold it in their balance sheet for longer until a new pool accepts it, leading to higher number of days to securitization (DTS).

To understand why VI lenders will have shorter DTS than non-VI lenders, we show evidence for securitizers prioritizing VI loans over non-VI loans. In particular, we show that VI loans are more likely to end up in pools that close earlier, compared to the potential pool options available for their loan, relative to non-VI loans, and provide evidence that securitizers delay pool formation to increase the VI share of loans included in the CMBS pool.

To understand the strength of pass through of DTS onto rates, we use an IV regression to estimate the effect. We show an effect of about 18bps per 100 DTS, which is economically meaningful.

[2] Double marginalization applied to our setting states that VI lenders will internalize that they will be earning money on both the origination and the securitization of the loan, and so are willing to charge a lower price to earn a higher market share. Non-VI lenders need to sell their loans to an investment bank to become securitized, so both the non-VI lender and investment bank will charge a markup and price will be higher than that of a VI lender

Using a back-of-the-envelope calculation, we calculate a lower bound of the impact of prioritization on the spread differential between VI and non-VI of 1bp, which explains 12% of the spread differential.

We explore how the results change across the credit cycle by looking at the differences in spreads and DTS in periods of lower origination volume, which we define as "cool" periods. We show that the difference in spreads and DTS are larger in quarters in "cool" periods. In cool periods, there are fewer security issuances, and so the cost of being delayed by a pool increases, because the time until the next pool is higher. This makes the prioritization force stronger in these cool periods.

We show the implications of the "prioritization" incentive by looking at credit allocation. There is some market segmentation in lending by VI and non-VI lenders, where loans originated by non-VI lenders are more likely to be high-risk and smaller than loans originated by VI lenders. The prioritization incentive will therefore lead to these non-VI borrowers to be have higher rates or leave the market entirely when the impact of the prioritization is large. We show evidence consistent with this, where VI market share is higher during times of lower origination volume.

In addition to credit allocation, the prioritization incentive also impacts CMBS pool construction and profitability. We show that the required yield to maturity on securities issued is higher when diversification is poorer. The share of non-VI loans is positively related to measures of diversification of the pool, so when VI securitizers prioritize their own loans, they end up hurting the diversification of the pool, which increase the required yield by investors and therefore reduces the amount that investors are willing to pay for the securities produced through securitization.

Lastly, we build a dynamic model of vertical integration in securitization and lending competition. In each period, VI and non-VI lenders compete in the mortgage market for borrowers that have logit demand. Then, the securitizer decides whether to issue a security, and if so, which non-VI loans to include in the pool for the security issuance. The securitizer has an incentive to include more of their rivals' loans to increase diversification and therefore increase the value of the securities sold and securitization profits. At the same time, they are balancing the incentive to prioritize their loans to lower their costs, relative to their rivals, which will allow them to gain a competitive advantage in origination in future periods. The ultimate goal of the model is to be parameterized and empirically estimated to conduct counterfactuals that change the market structure. For example, we want to understand how pool diversification, loan rates, and market shares would change if we removed prioritization, if there was GSE guarantee of loans, or if a centralized GSE pools these securities, as in the RMBS market. In particular, by understanding the equilibrium pool diversification across various counterfactuals, we can assess financial stability implications of the current market structure.

## 2 Literature Review

This paper is most closely connected to two strands of literature. First, the paper contributes to the literature on lender incentives in securitization. Many papers finds that securitization creates issues of adverse selection in the residential mortgage market (Keys et al. (2010); Agarwal et al. (2012)). Other papers find that this is not true in other markets, such as CLOs and commercial mortgages, because these markets incorporate features designed to mitigate adverse selection, such as retention of parts of the loan by the originator in CLO market and loan kickout rights for the purchaser of the most junior tranche in CMBS (Benmelech et al. (2012); Ghent and Valkanov (2016)). Other papers structurally model the decision of banks to hold their loan on the balance sheet or to securitize their loans (Buchak et al. (2018a); Buchak et al. (2018b)). We contribute to this literature by exploiting the market structure of the CMBS market to determine how vertical integration impacts different types of lenders.

This paper also relates to a long strand of literature regarding vertical integration (Grossman and Hart (1986); Hart et al. (1990); Crawford et al. (2018a); Jiang (2019)). The main forces highlighted in this literature are twofold. On one hand vertical integration eliminates double marginalization by making the input and final good producers to internalize the joint profits of the integrated firms as opposed to each one trying to maximize their profits separately. On the other hand, new incentives arise for the integrated firms to stop buying/selling inputs (or raise their prices) to limit competition in the final goods' market, which is known as vertical foreclosure. The main contribution of this project is to understand how vertical integration affects lending in a context where the vertically integrated firm directly benefits from their rivals' loans through diversification and its impact on profitability of securitization. We also provide suggestive evidence for vertical foreclosure in the commercial mortgage lending industry by showing the VI lenders prioritize their own loans when constructing pools, which raises the non-VI lenders costs.

## 3 Institutional Background

The CMBS market connects commercial real estate borrowers to institutional investors through securitization. The securitization process begins with a loan originator. Originators underwrite and issue mortgages backed by commercial property, such as retail establishments, office buildings, and industrial buildings. During the underwriting process, originators determine the loan terms and whether to originate the loan based on property characteristics, borrower characteristics, and market conditions. The originators price their loan by using the swap rate as a reference rate, and adding a spread above that, which is based on the factors mentioned earlier. When originating, the lender and borrower knows whether the loan is intended to become securitized. In this study, we will be focusing exclusively on commercial mortgages that become securitized.

The originator then partners with an investment bank(s) by contributing their mortgages to a pool, which then becomes the collateral backing securities that are structured and sold by the investment bank(s) to investors[3]. Usually, there are numerous originators contributing to the same pool, to decrease the amount of time required to accumulate the mortgages necessary to create a CMBS issuance, and to provide greater diversity of mortgage types. The investment bank, also referred to as a underwriter or securitizer, attempts to structure the pool and securities optimally sell the securities to investors at the highest price. These investment banks also originate loans themselves through their lending affiliate and they determine which loans will be included in the pool to ensure proper diversification and the level of risk of the pool is appropriate for investors[4]. The structuring of the deal focuses on tranching the default risk, which leads to securities for investors with a variety of risk appetite. Overall, the objective of the investment banks is to construct securities that can be sold for an amount greater than the cost to originate the loans used in the securities, net of transaction costs.

CMBS investors are repaid from the cash flows of the underlying mortgages in the pool. These investors have different risk profiles and will purchase securities for their risk profile. For example, some investors will only buy the most senior, AAA-rated bonds, and these investors tend to be institutional investors who want safe and liquid securities. Other investors may purchase the riskiest, first-loss securities,

[3] An originator who sells loans to a CMBS pool is called a sponsor. A sponsor can also contribute loans they purchased from other originators

[4] The most junior bondholder, also known as the B-piece buyer, gets to kick out any particular loan that they do not want included in the pool

known as the B-piece, and these investors tend to be high-yield investors with commercial real estate experts. By providing securities with different risk-profiles, their pool of investors is larger. While all of these investors face different risk-return tradeoffs, they all want to acquire these securities at the lowest price, which suggests that these prices are fair for their risk. The prices of these securities are intimately linked to the risk of the underlying loans within the pool, the diversification of the pool, and the current market conditions.

## 4 Data

The primary source of data come from SEC filing ABS-EE, a monthly filing required for issuers of certain types of asset backed securities, starting in November of 2016. These filings contain loan and property level data on all commercial mortgages within the CMBS pool and contain details such as the originator, origination date, size, interest rate, repayment status, property characteristics, and property financials. For a description of the variables used in our analysis, see table 14.

We supplement the ABS-EE data with data from prospectus supplements from a complete set of non-agency, multiborrower commercial mortgage backed securities (CMBS) that settled after January 1, 2012[5]. For each deal where data from its prospectus supplement were available on Bloomberg, information on both the underlying loans and the underlying collateral properties were collected. The loan data provide details of each individual loan being securitized including its originator, origination date, size, interest rate, LTV, and DSCR. We merge the loan data with information from the property data, which provides the location and type of the property serving as collateral for each loan[6]. We filter our sample to only include whole loans, which are loans that are not tranched or split up into pari-passu pieces. This removes loans that are packaged into multiple CMBS issuances. Additionally, we use data from CMBS issuances up through September 2023, but limit our sample to loans originated through 2022, which ameliorates concerns of right-censoring in our data. During the sample periods, we have 17,678 loans from 484 separate CMBS issuances, which span the universe of conduit, non-agency, whole CMBS loans originated from 2012-2022.

[5]The vast majority of these data are from issuances that settle before November 2016, because after this time, they show up in the SEC data. There are exceptions where the , which are Rule 144A issuances, which have fewer disclosure requirements and are therefore not in the SEC data

[6]This data was generously provided to us by Craig Furfine from Furfine (2020)

We obtain borrower identities from Real Capital Analystics (RCA), which we merge onto the property level data using the property name and address. We aggregate to the loan level by keeping the borrower information for the largest property included in that loan, which is not a problem because it is always the same borrower on every property in the loan. While we are not able to merge on borrower information for every property, this does allow us to include borrower fixed effects to help control for unobserved borrower characteristics in our analyses[7]. Lastly, this dataset also provides a few extra pieces of information, such as price per square foot or whether the loan was for purchasing or whether it was for refinancing.

RCA also provides rich data on commercial real estate transactions generally, aggregated to the market level. This rich data allows for sampling of observations choosing the outside good, which in our case, is conducting a commercial real estate transaction without a CMBS loan. See section B for details.

We merge information on security underwriter identities from Commercial Mortgage Alert, a commercial real estate trade finance publication. In order to check if underwriters are associated with originators on the issuance, we manually match identities of originators and underwriters to their ultimate parent. If originators and underwriter share an ultimate parent, we define this originator as vertically integrated. We show a simplified visual representation of vertical integration (VI) and the market structure in figure 1. In this example, we show a security with J.P. Morgan Securities as the underwriter, which has two lenders contributing to the pool, J.P. Morgan and Starwood Mortgage [8]. Since J.P. Morgan is affiliated with J.P. Morgan Securities, loans originated by J.P. Morgan are considered VI, whereas loans contributed by Starwood Mortgage are considered non-VI.

We get data on pool level characteristics and the price of the securities issued from Green Street's Commercial Mortgage Alert. The full set of pool level characteristics used throughout the paper can be found in table 14. The database provides yield to maturities (YTM), prices, notional amounts, and subordination amount on securities issued for each CMBS deal, which allow us to compute the total amount of capital raised from investors. For some securities, such as securities held for risk retention purposes, we often do not observe yields or prices. To address this, we make a few assumptions. First, we construct a panel of corporate bond yields for each credit rating from AAA to CCC using the

[7]The RCA data available to us does not include data on self-storage properties and mobile home parks. We have information on 14,881 loans, which is 84% of observations

[8]Lenders that contribute to the pool are known as sellers or sponsors

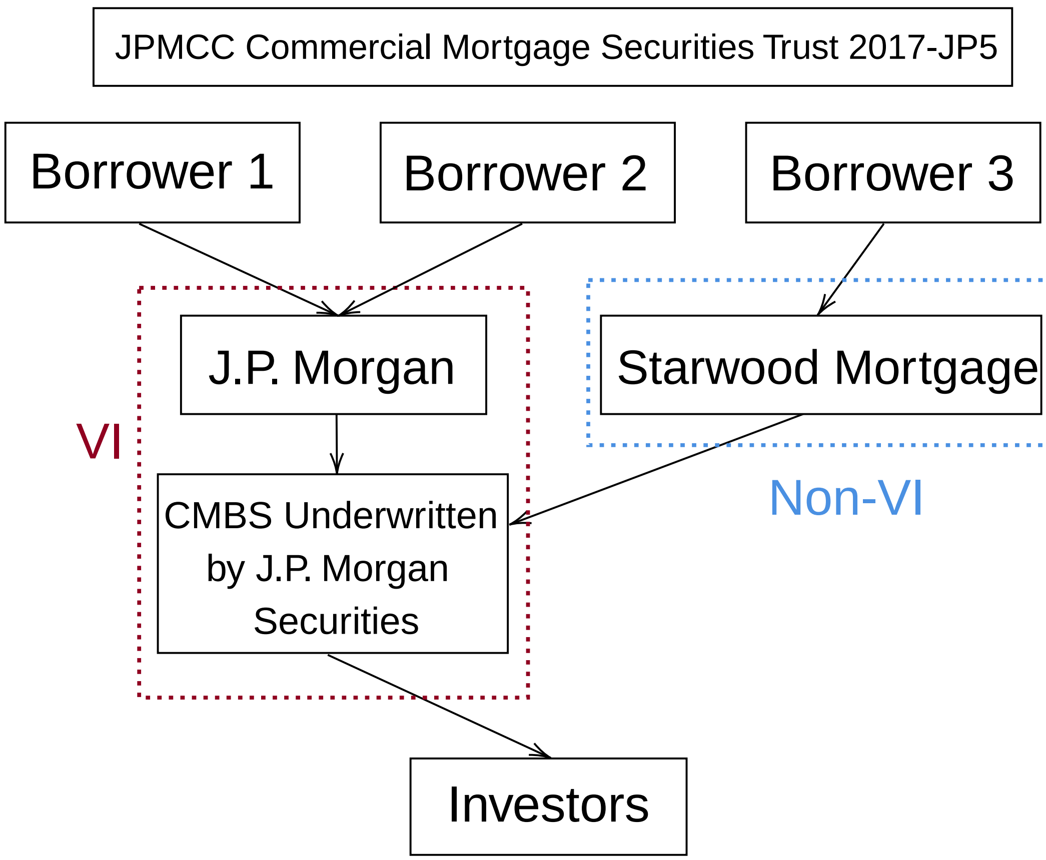


**Figure 1:** Simplified Market Diagram

Bank of America, Merill Lynch bond series on FRED. Then, when bond ratings are available for the CMBS securities, we merge on the corresponding corporate bond yields on the pricing date of the CMBS securities. We compute a "CMBS risk premium" which we calculate as the average difference in yields for CMBS bonds for which we know the yields and corresponding corporate bond yields, for both AAA securities and for non-AAA securities. For CMBS securities for which we do not observe a yield, we impute a yield by taking the corresponding corporate bond yield and adding the "CMBS risk premium" calculated previously.

Then, we compute the yield of all securities with no subordination and no YTM information, such as the risk retention pieces, at the CCC bond yield + 8.89% [9]. Secondly, we ignore interest only (IO) securities and calculate the yield and bond amounts only on non-IO securities. Thirdly, we interpolate the yield to maturity of any security that we are still missing by using the yields of securities we have and their subordination amount. Lastly, if we cannot interpolate the yield, we will set it equal to the average yield of that issuance and we will flag that security. If the notional amount of flagged securities exceeds 1% of the sum of all face value of the securities, we do not use the issuance in our analysis to ensure the total amount raised is accurate for observations used in our analyses.

We show the largest originators and their VI status in table 11, which shows the largest 25 lenders, along with the proportion of loans that they originated for which they are also underwriters. Notice that most lenders are either always VI or never VI, but there are a few lenders that have only a fraction of their loans as VI[10].

To show the differences in loans originated by VI and non-VI lenders, we show the means, as well as differences[11], for all outcome variables and controls used in the analyses in Table 12. On average, VI loans have lower loan spreads, have a shorter time to securitization, have lower LTV, are larger, have

[9] Generally, the yield and price of risk retention securities are not publicly available information, though sometimes it can found in the prospectus for the issuance. We get the 8.89% premium by looking at how much the horizontal RR piece was priced for the CSAIL 2021-C20 was priced over the CCC corporate bond yield on that date

[10] This can occur for three reasons. The first is that they switch to becoming VI lenders. The second is that they sell their loan to another originator, who then contributes that loan to a pool for which the original originator is not an issuer. Lastly, it could be because the lender contributes the loan to a pool for which they are not an underwriter. All of these situations are infrequent and so we will often abstract from this and refer to the lenders themselves as VI

[11] When calculating the differences of loans characteristics, we exclude multifamily loans. This is due to multifamily loans being very low risk and non-VI lenders being over-represented in this segment. The full summary stats include multifamily loans

higher debt service coverage ratio, have similar debt yield, and have a higher price per square foot. We discuss this further in section 6.1. For a full table of summary statistics for all loans variables, as well as issuance variables, see table 13

# 5 Vertical Integration and Lending

In this section we document empirical facts about differences in loan spreads and time to securitization in VI lending, we discuss synergies and incentives that explain these empirical patterns, and show evidence for VI prioritization and vertical foreclosure.

## 5.1 Loan Spreads

Vertically integrated loans have lower loan spreads than non-VI loans, and they have higher market shares. Since loan spreads depend on the riskiness of the individual loan as well as market conditions, we look at residualized spreads across VI and non-VI loans to ensure we are appropriately controlling for loan risk. Formally we use the following specification to calculate residualized spreads:

$$r_{i,j,p,t,g} = \omega Z_i + \psi_{pt} + \psi_g + \epsilon_{i,j,p,t,g} \tag{1}$$

The unit of observation is loan $i$, lender $j$, property type $p$, time $t$, and MSA $g$. Our control characteristics $Z_i$ include loan-to-value, debt yield, and debt service coverage ratio, which lenders use to price loans. Additionally, we control for the log loan size, log of price per square foot, and capitalization rate, which is a measure of risk for the property[12]. Our fixed effects include a property type x month fixed effect, which is controlling for the market conditions for each property type, and MSA fixed effects, which control for the overall risk level of each MSA. We define the $\widehat{\epsilon_{i,j,p,t,g}}$ as the residualized spread, and this can be interpreted for the price of the loan, conditional on loan characteristics and market conditions.

In figure 2 we plot annual market shares by lender and the residualized loan spreads, and split up each point by VI and non-VI lenders. We can see a striking pattern, which is that VI lenders (orange) tend

[12] Some observations are missing controls. We still include these observations in the specifications by setting their values to -1 and including dummies for the missing controls

to have lower average residualize spreads than their non-VI (blue) counterparts[13]. Consistent with lower spreads, we also see VI lenders capture higher market shares than non-VI lenders.

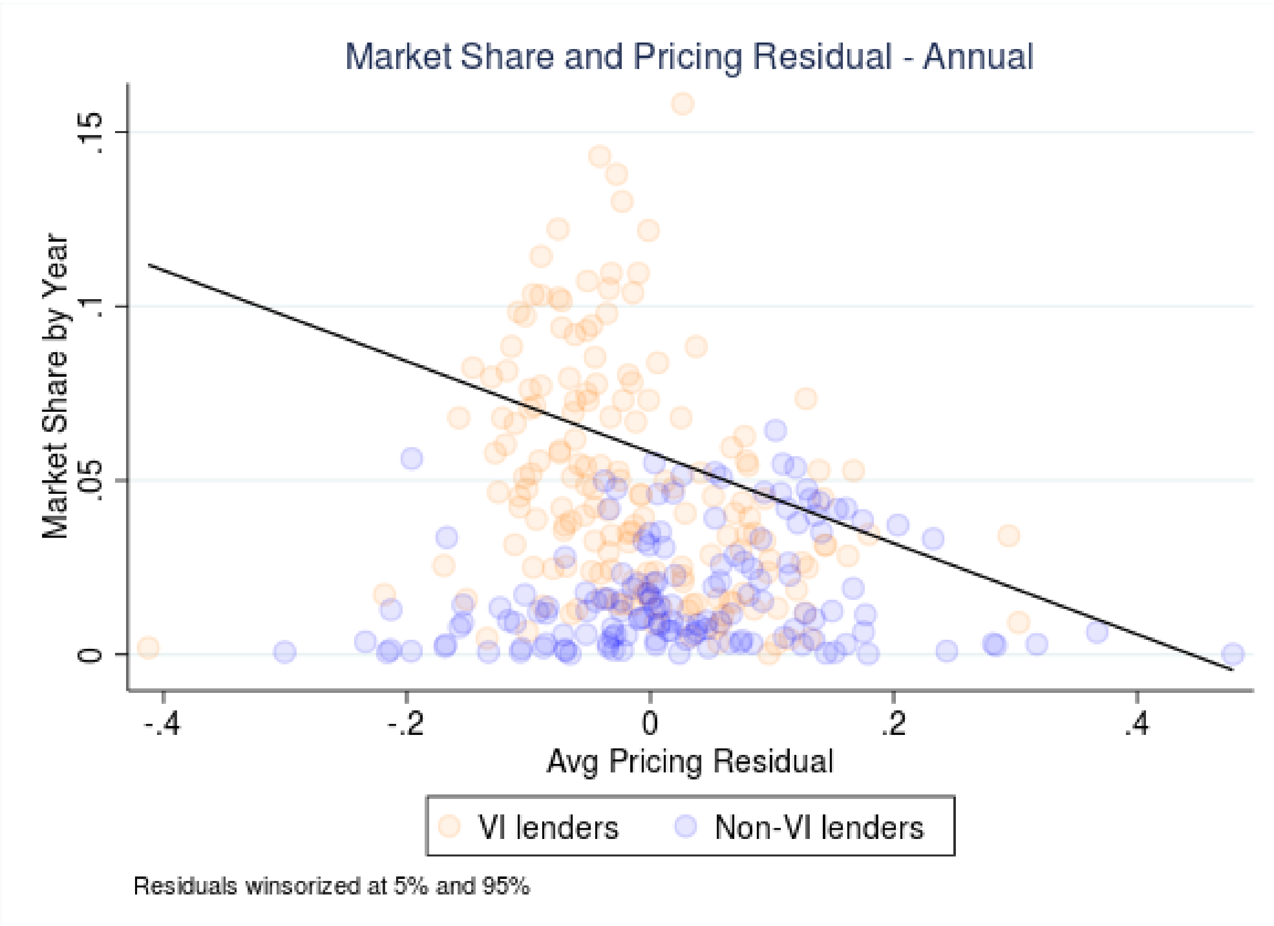


**Figure 2:** Average Residualized Spreads and Market Shares

We also look at this pricing result in the regression framework. The empirical approach compares spreads for loans financed by a vertically integrated lender to an observationally similar loan by a non-vertically integrated lender. First we regress our dependent variable $o_{i,j,p,t,g}$ on the dummy variable for whether the loan is from a vertically integrated lender $VI_{j,t}$, controls, and the same fixed effects as before. We also include borrower fixed effects $\psi_b$ in some specifications to control for unobservable borrower quality. Formally, the specification is below:

$$o_{i,j,p,t,g} = \alpha * VI_{j,t} + \omega Z_i + \psi_{pt} + \psi_g + \psi_b + \epsilon_{i,j,p,t,g} \tag{2}$$

[13] To see the distribution of raw and residualized spreads, see figure 11

We report the results of specification 2 in table 1, columns 1 and 2[14]. In column 1, we see that VI loans are originated with a spread that is 10.3bps lower on average. In column 2, we include borrower fixed effects and see that the difference in spreads declines to 8.5bps. This is still economically significant as it translates to $12,500 in annual savings on the average sized loan. By including borrower fixed effects, we are looking within borrower, meaning we are identifying the VI coefficient from borrowers that obtain a loan from both a VI lender and non-VI lender and therefore we are controlling for borrower quality. Since the coefficient decreases after including borrower fixed effects, this suggests that the unobserved borrower quality is lower for non-VI loans and therefore column (2) more accurately identifies the difference in spreads [15]

**Table 1:** Spreads and Days to Securitization for VI and Non-VI

| | (1)<br>SPRD | (2)<br>SPRD | (3)<br>DTS | (4)<br>DTS |
|---|---|---|---|---|
| VI | -0.103*** | -0.085*** | -6.713*** | -7.077*** |
| | (0.006) | (0.011) | (0.838) | (1.508) |
| Observations | 17561 | 8269 | 17592 | 8277 |
| $R^2$ | 0.712 | 0.854 | 0.242 | 0.533 |
| Fixed Effects | PxT,G | PxT,G,B | PxT,G | PxT,G,B |
| Controls | Y | Y | Y | Y |

Standard errors in parentheses
Standard errors clustered at the MSA level
Fixed effects codes: P=Prop Type,T=Orig Month,G=Prop MSA,B=Borrower
Controls: LTV, DSCR, Log Size, Debt Yield, Log Price/Sq Ft, and Cap Rate
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## 5.2 Days to Securitization

One reason that vertically integrated loans have lower rates is that VI loans have a significantly shorter time from origination to securitization relative to non-VI loans. When loans are expected to stay on the balance sheet for a short period of time, this lowers the warehousing cost of the loan, and this lower cost can be passed onto the loan spread. In this section, we show that non-VI loans take longer to go from origination to securitization than their VI counterparts[16].

[14]We cluster standard errors at the MSA level, following DeFusco et al. (2020), to account for serial correlation and MSA-specific random shocks in all specifications

[15]The observations decline significantly because most borrowers only borrow once in our sample. Additionally, borrower information is not available for our entire sample as described in section 4

[16]To see the distribution of raw spreads and residualized spreads, see figure 12

We show this result in a regression framework using specification 2 in table 1, columns 3 and 4. In column 3, we see that across all time periods, VI loans are originated 6.7 days faster than non-VI loans, which is virtually unchanged in column 4 . This is economically significant as the mean DTS is around 62 days.

To understand what is driving this result, we discussed with industry participants and one senior security underwriter described an important benefit of being both an originator and an investment bank (e.g. VI) is control over the CMBS issuance. One way this control materializes is that they are much more likely to include their own loans in the upcoming pool instead of non-VI lenders' loans, on the margin, which leads to a greater DTS on average for non-VI lenders.

## 5.3 Evidence of VI Prioritization

One important impact of vertical integration generally is vertical foreclosure, which is the concept that vertically integrated firms are able to raise the costs of their rivals to give themselves a competitive advantage. One way this can be seem in our setting is the prioritization of the VI lenders own loans when constructing pools, which would increase the time to securitization for the non-VI lenders, on average. In this section, we will show evidence that VI lenders are prioritizing their own loans over their competitors when forming CMBS pools.

In our first analysis, we will look at all of the possibilities of pools for where loans could end up and compare it to the pool where they ultimately end up. To define the set of pools for which a loan can enter, we look at all of the issuers[17] with which that loan's originator has worked. We assume that a loan can go into any pool created by that issuer if they have contributed a loan to that issuer in the past and the pool is open[18] when that loan is originated. Additionally we allow for originators to stop working with issuers by excluding pools created by issuers for which the lender has discontinued

[17] We standardize issuers based on patterns in the pool names. For example, any pool with "BNK" in it's name, we classify as part of the "BANK" issuer, which is an issuer jointly run by Bank of America, Morgan Stanley, and Wells Fargo

[18] We define a pool as open at a particular date if at least one loan that ultimately ends up in the pool has been originated and if the pool is at least 30 days from closing. We assume 30 days because practitioners told us it takes about one month for investors to do due diligence on the loans in the pool and for the legal contracts to close

contributing loans for over a year[19]. From there, we rank the potential pools it can enter by closing date, meaning a rank of 1 is the first possible pool that the loan can enter[20]. To illustrate, we show a simple diagram in figure 3. Here we show four pools with pool opening dates $T_{1,o}, T_{2,o}, T_{3,o}, T_{4,o}$ and pool end dates $T_{1,c}, T_{2,c}, T_{3,c}, T_{4,c}$. Loan $i$ was originated at time $t_i$, which occurs after pools 1,2, and 3 are open but before pool 4 is open. So loan $i$ has three pools in its choice set, and the earliest pool will be pool 1, which has the earliest pool closing date, followed by pool 2, and lastly pool 3.

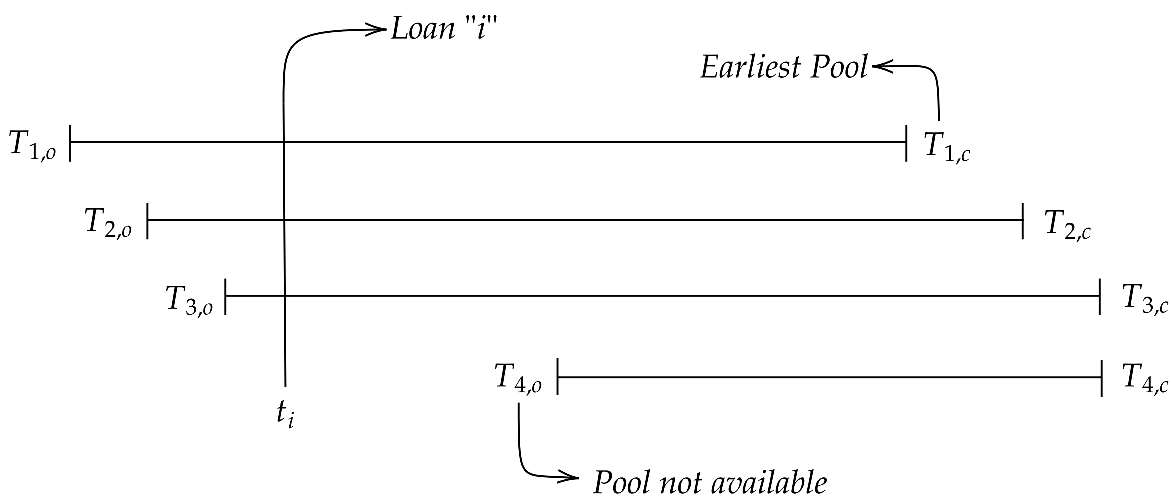


**Figure 3:** Pool Ranking Diagram

To show that VI loans are prioritized when pool building, we look at the probability that VI and non-VI loans end up in the first possible pool. Formally, we use the following specification:

$$o_{i,j,p,t,g} = \alpha * VI_{j,t} + \omega Z_i + \psi + \epsilon_{i,j,p,t,g} \tag{3}$$

Here $o_{i,j,p,t,g}$ is either an indicator for whether the the loan ends up in the first possible pool, or the pool number that the loan ultimately ends up in, and $\psi$ is a choice set fixed effect. The specification with choice set fixed effects is comparing the resulting pool number for two loans that have the exact

[19]For example, if the last time lender $a$ worked with issuer $B$ was 01/01/2018. We will exclude pools created by issuer $B$ in originator $a$'s choice set after 01/01/2019

[20]Alternatively, we could have ranked it by opening date, but we believe this would less realistic, because lenders know the relative pool closing ordering of pools available to them. Whereas the pool opening date is imprecisely measured because we don't truly know when the securitizer began building the pool

same set of possible pools to end up in, and so it addresses the concerns of any mechanical relationship between VI and pool number[21]. In table 2 we show the results of specification 3 for the earliest pool indicator in columns 1 and 2, and the pool number in columns 3 and 4. In columns 1 show that unconditionally, the probability that a VI loan ends up in the earliest pool is about 2.7 percentage points higher than a non-VI, though this is not statistically significant. In column 2, our preferred specification, we show that VI loans are 5.4pp more likely to end up in the earliest pool relative to a non-VI loan with the exact same choice set and set of observable characteristics, which is economically significant as the mean probability is about 62%. In column 3, we see that the average VI loan pool number rank is 0.08 smaller than the average non-VI pool rank, though statistically insignificant. When including choice set fixed effects, we see that average difference in pool number rank is .16, which is economically significant given the mean pool number rank is 1.6.

**Table 2:** Probability of Being the Earliest Pool and VI

| | (1)<br>Earliest Pool | (2)<br>Earliest Pool | (3)<br>Pool Num | (4)<br>Pool Num |
|---|---|---|---|---|
| VI | 0.027 | 0.054$^{**}$ | -0.080 | -0.164$^{***}$ |
| | (0.026) | (0.026) | (0.053) | (0.051) |
| Observations | 17377 | 16500 | 17377 | 16500 |
| $R^2$ | 0.001 | 0.594 | 0.002 | 0.581 |
| Fixed Effects | No | Choice Set | No | Choice Set |
| Controls | N | Y | N | Y |

Standard errors in parentheses
Standard errors clustered at the MSA level
Mean probability of being in the earliest pool: .624
Mean pool number of: 1.608
Controls: LTV, DSCR, Log Size, Debt Yield, Log Price/Sq Ft,Cap Rate, and Prop Type FE
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Since the result becomes stronger from column 3 to column 4, this result suggests that non-VI lenders have more pool choices on average. This is indeed true, on average VI loans have 7.8 pool possibilities whereas non-VI loans have 8 pool possibilities (see figure 7). This is consistent with a strategic response by non-VI lenders to increase the number of issuers they work with to reduce the impact of vertical foreclosure. Additionally, the fact that non-VI lenders tend to work with a greater number of issuers actually works against our DTS results presented in the previous section. In other words,

[21] In an extreme example, if a lender only has possible choice in the choice set, they will mechanically end up in their first pool

all else equal, having a greater number of pool choices to include your loan will decrease your DTS, and given that we see non-VI lenders have higher DTS and a larger number of pool choices, this suggests the impact vertical foreclosure force is stronger than the impact of having more pool choices and that the impacts of vertical foreclosure are greater than what we would estimate in the reduced form. For this reason, we believe our reduced form results represents a lower bound on the impacts of prioritization on DTS.

In a separate analysis, we look at how long securitizers have delayed their pools and what they may gain. In this exercise, we look at all available loans that an issuer could choose from when creating a particular pool and ask how much sooner could they have made that same pool, in terms of loan amounts in each separate property type[22]. Here we make a strong assumption that any loan made by an originator that has worked with the issuer in the past is available for the issuer's pool[23]. We also assume that the first date an issuer could possibly form the pool is the day after their most recent previous pool has closed, or 152 days[24] prior to closing if it is their first pool. Lastly, we assume that a loan becomes unavailable for a pool if, at that date, the loan ends up in another pool that closes within 30 days.

To better illustrate our exercise, we create a simple example in figure 5. On the top, we see the pool begin date outcome ($T_{j,o}$), pool end date ($T_{j,c}$) outcome, and the pool characteristics, which are amount of each property type that ends up in the pool ($P_j$). We observe the issuer pipeline, which are all the loans that are originated by any lender that contributes to that issuer. We then define the earliest time at which the same pool (in terms of characteristics) can be constructed as $T_{\tilde{j},c}$. In the figure, it would be the first time the origination pipeline has at least 400 million in retail property loans, 150 million of hotel loans, 250 million in office loans, and 200 million of multifamily loans.

In figure 5 we plot the distribution of pool delay. Here we see a few things. First is that most pools are not delayed for long, the median pool is delayed by 15 days. Second that there is a long right tail in the delay, which we can also see in terms of the mean of 29 days. It is not obvious that a securitizer

[22]Formally, Pool Delay = days between true closing date and earliest possible pool date, minus 30

[23]This is not completely realistic as an originator may have committed their loan to another issuer they work with

[24]The decision to use 152, or 5 months, was informed by our conversation with practitioners, where they told us that loans are originated approximately 1-5 prior to closing

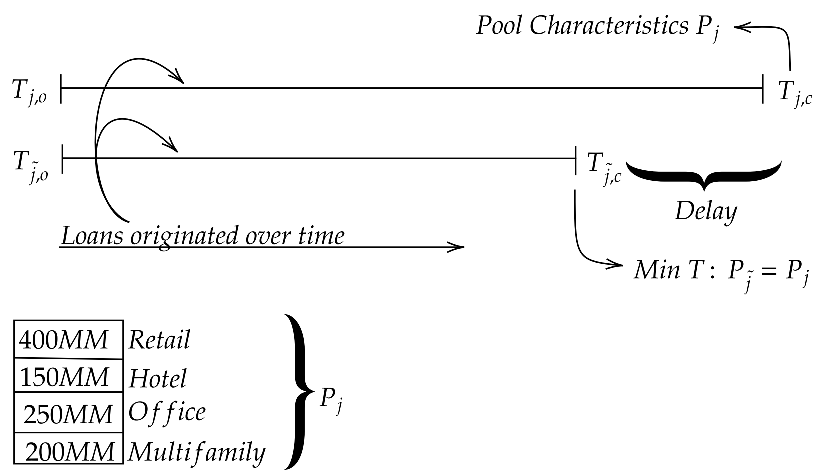


**Figure 4:** Pool Delay Diagram

would want to delay a pool, as it is costly to keep their own loans on the balance sheet longer, which is likely why we see only a small delay.

One potential reason to delay a pool is to ensure the securitizer can place more of their own loans in the pool, such as those that are soon to be originated. To test this, we calculate the difference in the share of loans that are originated by VI lenders in the true pool outcome versus the VI share of loans in the earliest pool possible[25]. We then plot a binscatter of the relationship between the difference in VI share and amount of time the pool is delayed in figure 6. This figure shows that there is a positive relationship between the difference in VI share and pool delay. While not causal, this plot suggests that securitizers are delaying the pools to get more of their own loans into the pool.

To understand how this prioritization impacts the time from origination for both types of lenders, we plot the days until the next possible pool[26] in figure 7 for each loan in our sample. First, we notice the days until next pool for VI (36.3 days) is slightly longer than non-VI (34.8 days), which is consistent with non-VI lenders having a larger number of possible pool choices. We calculate the effect

[25]VI share for the hypothetical pool is calculated as the weighted average of VI shares in each property type as of the earliest pool date, weighted by the property type shares in the true pool

[26]We calculate days until next pool as the number of days from the closing date of the pool it ended up in and the closing date of the next possible pool

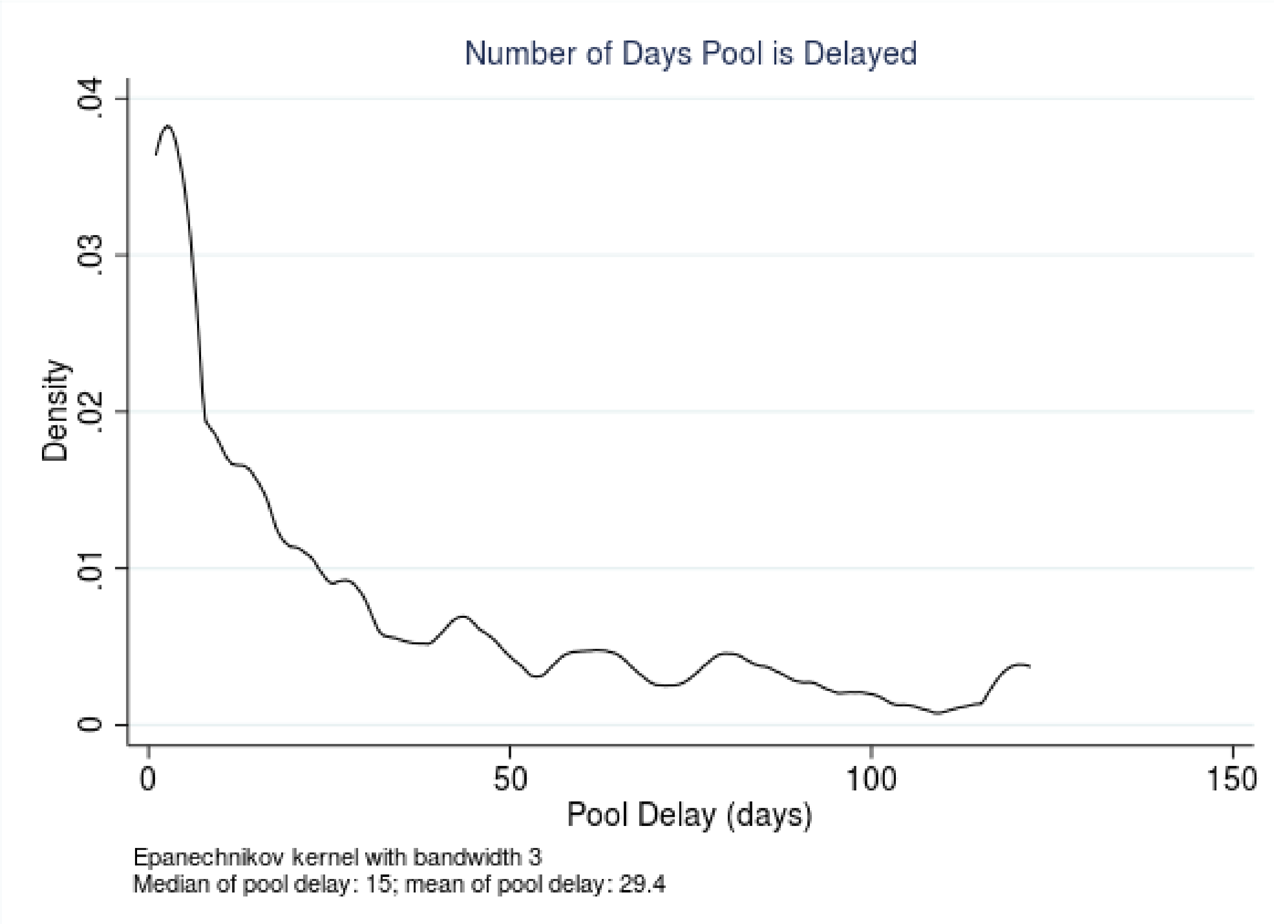


**Figure 5:** Pool Delay Kernel Density

of prioritization using a back-of-the-envelope calculation by combining our estimate of the differential pool number in column 4 of table 2 (.164) and the average number of days until next pool for non-VI loans (34.8 days), which yields an effect of about 5.7 days. This is remarkably similar to our baseline estimate of a difference of 7 DTS in column 4 of table 1, which suggests that prioritization is explaining most of the difference in DTS between VI and non-VI lenders.

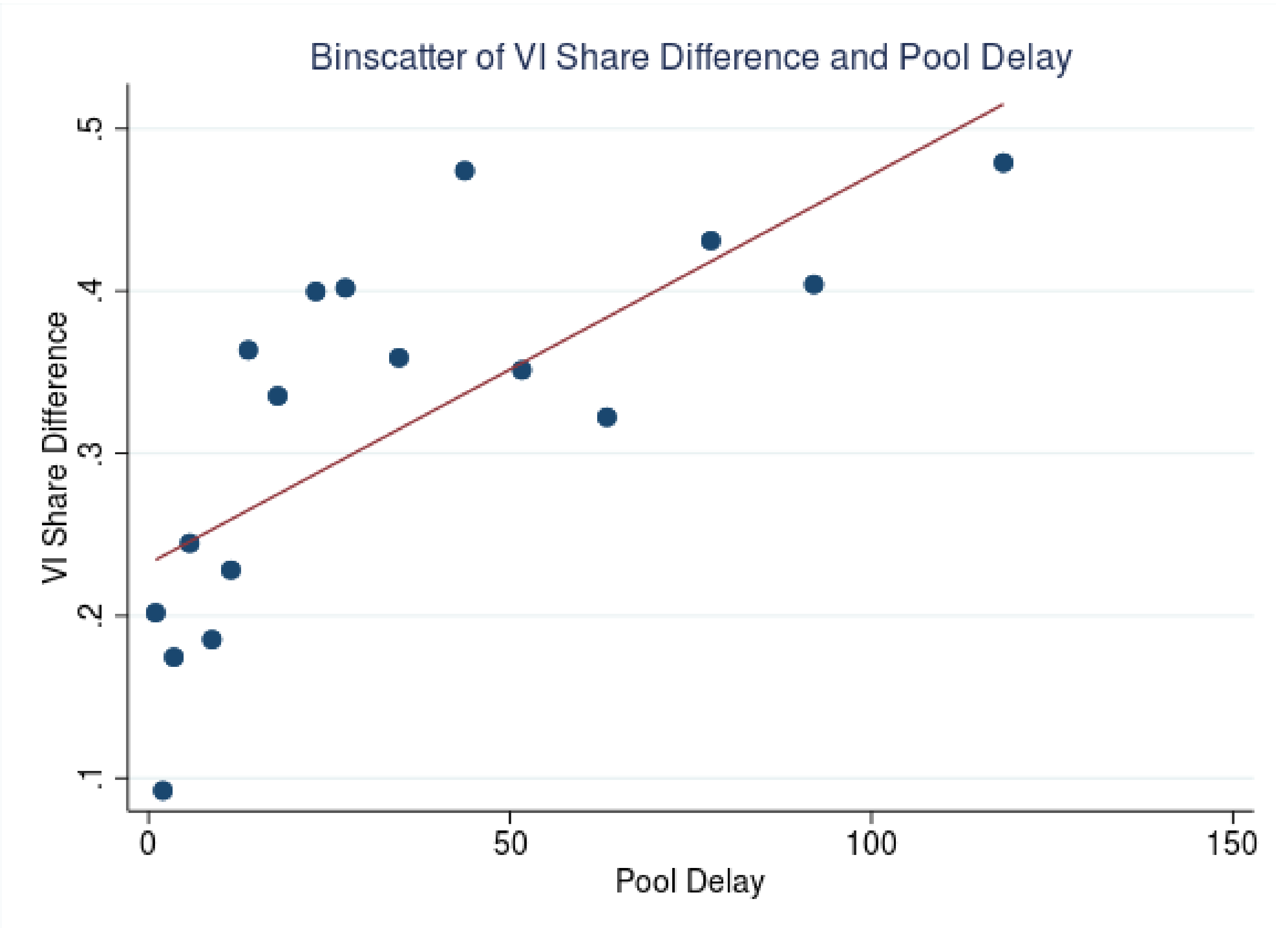


**Figure 6:** VI Share Difference and Pool Delay

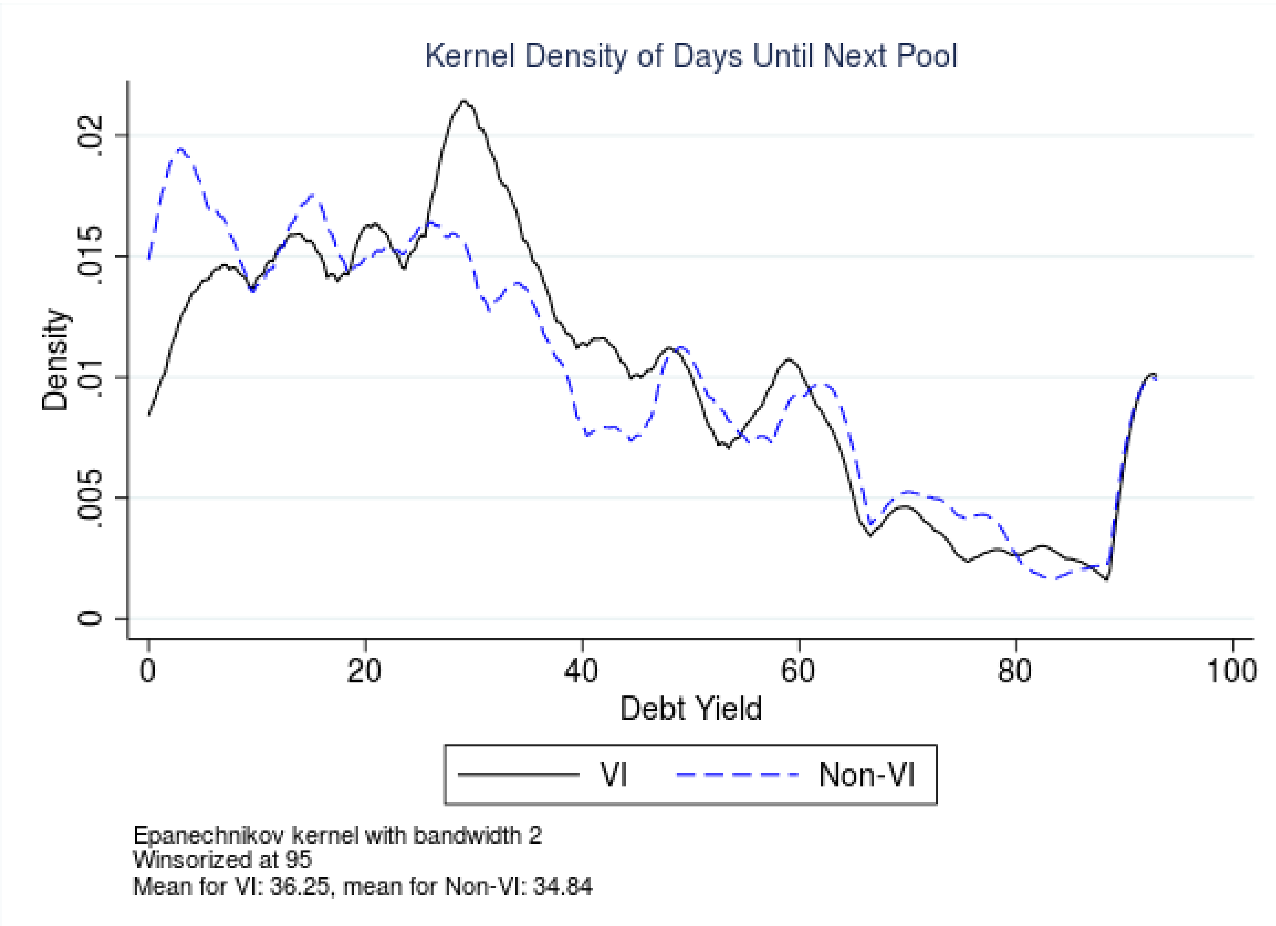


**Figure 7:** Cost of Delaying for VI and non-VI Loans

### 5.4 Pass Through of DTS to Loan Spreads

In this section, we provide evidence that the difference in rates between VI and non-VI is partially explained by difference in the days to securitization. As mentioned earlier, when a lender expects a loan to be on the balance sheet longer, this raises warehousing costs, which can then be passed onto the borrower through higher spreads.

To show this result, we use an IV regression framework. A simple OLS regression with DTS as an independent variable will suffer from endogeneity concerns because DTS itself is an outcome that is determined after the loan spread has been determined. For example, a loan with an unusually low spread may be seen as undesirable by the securitizer structuring the pool and therefore they may push the loan to a future pool. In this case, we would see a negative relationship between the DTS and loan spreads. Therefore, we need an instrument that shifts rates only through DTS and is not related to other unobservable factors related to pricing.

We instrument for DTS by constructing new measures from the pipeline analysis described earlier. In particular we instrument for DTS by using the days since most recent pool closing[27] and a piecewise linear function function of how subscribed the earliest pool for is the property type of the loan being originated. As we have seen in the previous section, most loans end up in the earliest possible pool. Therefore, the days since the most recent possible pool will shift the DTS if pools are being created at regular intervals, which means it meets the relevance criterion. For example, if a loan was originated just after a possible pool had closed, it will need to wait until the next pool closes, which can be a long time. Compare that to a loan originated a couple months after the most recent pool closes, which will only have to wait a short time for the next pool to close. Additionally, it is reasonable to assume that the time since the most recent pool closing is unrelated to unobserved pricing characteristics of the loan, which means that the instrument meets the exclusion restriction.

We also use a piecewise linear function of how subscribed the particular property type is in the earliest pool is when the loan was originated. When a loan of property type $p$ is originated, we calculate the total loan amount of type $p$ in the earliest pool's issuer's pipeline as of the day before that loan was

[27] We calculate this measure by taking the difference between the closing date of latest possible pool before origination and the origination date

originated. Then we divide that amount by that total loans of type $p$ that end up in that ultimately end up in the earliest pool and call this the subscription ratio[28]. The idea behind this instrument is simple; if a lender has originated a loan when the issuer's pipeline is empty for that property type, it will take longer to securitize because it will take time to originate of that type. As it becomes closer to being "full" of that particular property type, the time to securitization will decline. But, if the property is oversubscribed, meaning that the pipeline has more of that property than is possible to put into the pool, that loan will likely have to wait until a future pool and will therefore have a higher DTS. Because of the way the level of subscription would impact DTS, we cannot simply use the subscription ratio, but instead use the subscription ratio, an indicator for the subscription ratio being greater than 1, and in interaction between the subscription ratio and the indicator for the ratio being greater than 1. The exclusion restriction for this instrument is met if the subscription ratio is not related to pricing factors (excluding DTS), conditional on controls and fixed effects. We believe this is a reasonable assumption because we include property type x origination month fixed effects, which control for property specific market conditions.

The regression specification that we use is below:

$$r_{i,j,p,t,g} = \alpha * VI_{j,t} + \beta * DTS_i + \omega Z_i + \psi_{pt} + \psi_g + \psi_b + \epsilon_{i,j,p,t,g} \tag{4}$$

In table 3, we show the IV estimates of specification 4. In column 1, the IV specification[29] without borrower fixed effects, the effect is positive at 15bps per 100 DTS. In column 2, the IV specification with borrower fixed effects, shows an effect of 18bps per 100 DTS.

Our preferred specification, column 2, both addresses soft information through borrower fixed effects, and the endogeneity of other omitted variables correlated with both DTS and spread, and so this is most accurate estimate of the impact of of DTS on loan spreads. An estimate of 18bps per 100 DTS is sensible and economically meaningful. If a non-VI lender expects to be foreclosed and is not able to place their loan into the upcoming pool, we can use a back-of-the-envelope calculation to see how this

[28] The implicit assumption here is that the lenders have information about what is in the issuers pipeline when they are originating the loan and that they have information about what loan volume each property type will end up in the earliest pool

[29] For both columns 1 and 2, the first stage F-statistic is greater than 30 and therefore suggests that these are not weak instruments

**Table 3:** Rates and DTS - Various Instruments

| | (1) SPRD | (2) SPRD |
|---|---|---|
| DTS (100s) | 0.149** | 0.183** |
| | (0.060) | (0.091) |
| Observations | 16587 | 7749 |
| Fixed Effects | PxT,G | PxT,G,B |
| Controls | Y | Y |
| IV | YES | YES |
| 1st Stg F-stat | 210.18 | 97.41 |

Standard errors in parentheses

Fixed effects codes: P=Prop Type,T=Orig Month,G=Prop MSA,B=Borrower

Controls: LTV, DSCR, Log Size, Debt Yield, Log Price/Sq Ft, and Cap Rate

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

would impact their pricing. If we look at the average impact of VI prioritization on non-VI pricing with our prior back-of-the-envelope effect of 5.7 days, this would raise the spreads of non-VI lenders by 1 basis point. Comparing this to our estimated difference in spreads from column (2) in table 1, this suggests that vertical foreclosure explains about 12% of the difference in loan spreads between VI and non-VI lenders.

## 5.5 Prioritization Amplified During Credit Crunches

In a well functioning CMBS securitization market, commercial mortgages loans are originated and securitized rather quickly. During these normal times, the costs of delaying until a later pool are relatively small, because pools are being created frequently. When the CMBS securitization market is disrupted and fewer pools and securities are being sold to investors, this is not true, because the time until the next pool is large. Therefore, it is during these "cool" times where we would expect this vertical foreclosure effect to have a stronger effect.

To test this, we first define "cool periods" as quarters that are in the bottom quartile of total origination volume. Figure 14 shows total origination volume by quarter and whether the quarter is a cool period. We then run the following specification, which is similar to our baseline for it includes an interaction with a "cool" dummy.

$$o_{i,j,p,t,g} = \alpha * VI_{j,t} + \beta * Cool_t x VI_{j,t} + \omega Z_i + \psi_{pt} + \psi_g + \psi_b + \epsilon_{i,j,p,t,g} \quad (5)$$

In table 4, we report the results of we report the results of specification 5 for the spread and days to securitization. In column 1, we see that in non-cool periods, the gap in spread between VI and non-VI loans is 10bps, which grows to 13bps in cool periods, though this increase is not statistically significant. In column 2, we include borrower fixed effects and results are mostly unchanged. In columns 3 and 4, we see a more striking gap in DTS across non-cool and cool periods. In column 3, we see that in non-cool periods the gap is 5 days, which is about 8% of the mean DTS. In cool periods, this gap in DTS grows to 21 days, which is about 33% of the mean DTS. Results in column (4) remain similar.

**Table 4:** Spreads and Days to Securitization for VI and Non-VI in Different Market Conditions

| | (1)<br>SPRD | (2)<br>SPRD | (3)<br>DTS | (4)<br>DTS |
|---|---|---|---|---|
| VI | -0.100*** | -0.081*** | -4.990*** | -6.033*** |
| | (0.006) | (0.011) | (0.880) | (1.570) |
| Cool x VI | -0.026 | -0.056 | -16.261*** | -13.648** |
| | (0.018) | (0.034) | (2.422) | (5.879) |
| Observations | 17561 | 8269 | 17592 | 8277 |
| $R^2$ | 0.712 | 0.854 | 0.244 | 0.534 |
| Fixed Effects | PxT,G | PxT,G,B | PxT,G | PxT,G,B |
| Controls | Y | Y | Y | Y |

Standard errors in parentheses
Standard errors clustered at the MSA level
Fixed effects codes: P=Prop Type,T=Orig Month,G=Prop MSA,B=Borrower
Controls: LTV, DSCR, Log Size, Debt Yield, Log Price/Sq Ft, and Cap Rate
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

These results are consistent with the prioritization channel. During cool periods, the impact of the vertical foreclosure incentive increase, because fewer pools are created. On average, 5.7 pools are created in a cool quarter vs 12 pools are created in a non-cool quarter. Since the amount of time between the next pool increases and because non-VI loans are less likely to get into the earliest pool, we see a larger gap in the DTS in these cool periods. This translates to a higher cost, as discussed earlier, which is why we also see a greater gap in spreads.

# 6 Implications of Vertical Integration in Securitization

## 6.1 Market Segmentation and Market Shares

The market for CMBS lending is partially segmented, with VI lenders and non-VI lenders serving different borrowers in terms of loan characteristics. In figure 8 we plot the kernel densities of our six loan characteristics that are used as controls in the regression specifications, by VI status[30]. The figures show that VI lenders originate loans that are larger and less risky. Non-VI loans have much higher loan-to-value, smaller debt service coverage, higher capitalization rates (a measure of property risk/return), similar debt yield, and lower price per square foot. It is also important to note that while the loans originated are different, there is a large overlap in the densities for these loan characteristics.

We also look at the share of loans that are originated by VI lenders over time, along with total origination volume in the CMBS market, in figure 9. Here we see that there is an initial downward trend, followed by a general upwards trend in the share of VI loans and that the share is inversely related with total origination volume. These results suggest that in times with lower overall origination volume, or cool periods, credit allocation to riskier borrowers that tend to borrow from non-VI lenders will decline. This pattern is consistent with a natural consequence of our rate results that showed that VI lenders have lower rates in cool periods, because borrowers are sensitive to the interest rate charged and some borrowers who may have chosen a non-VI lender may instead exit the market or choose a VI lender.

[30] As in table 12, we exclude multifamily properties from the sample as these are very low-risk loans and non-VI lenders specialize in this lending segment

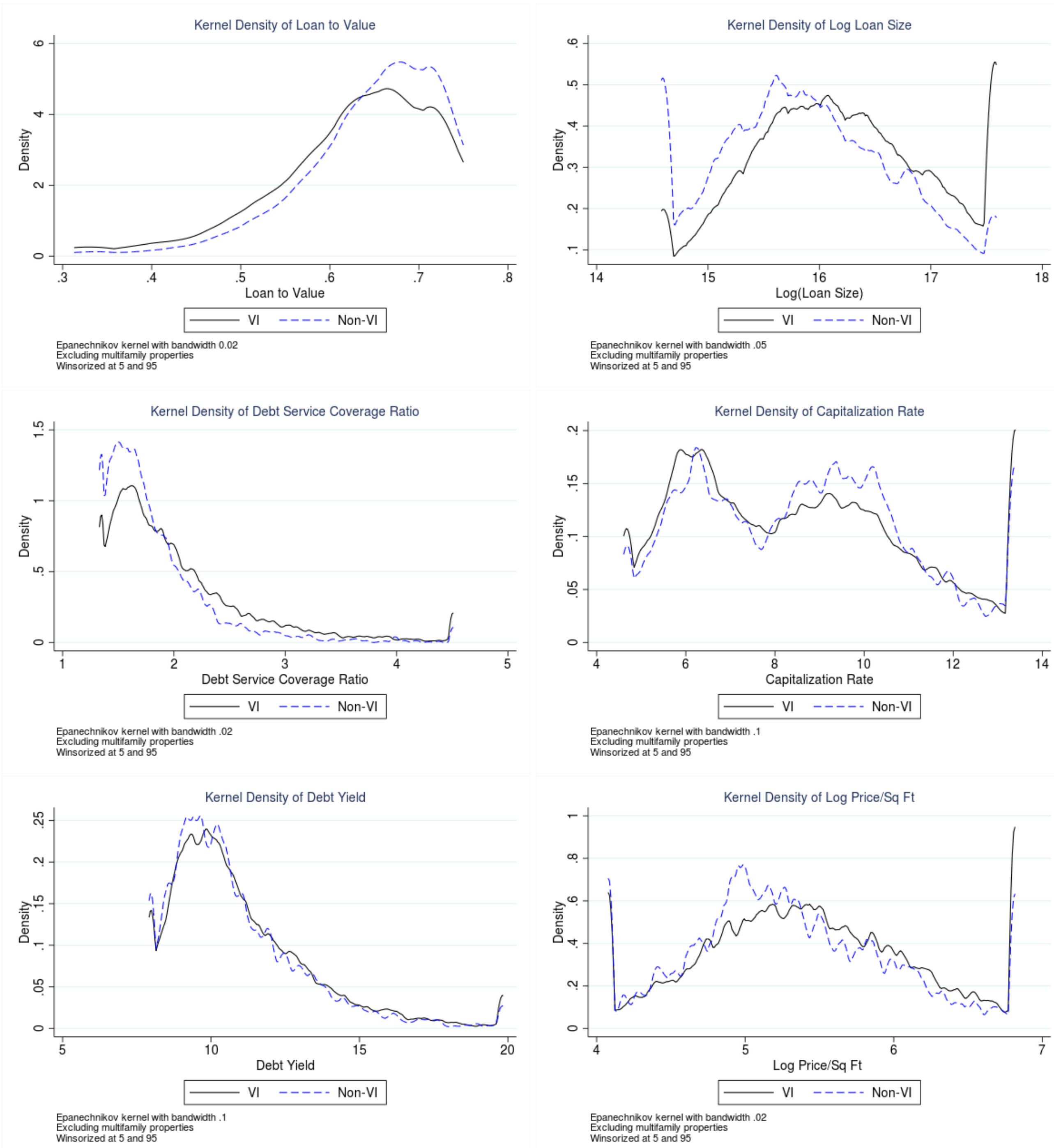


**Figure 8:** Loan Characteristics of VI vs Non-VI

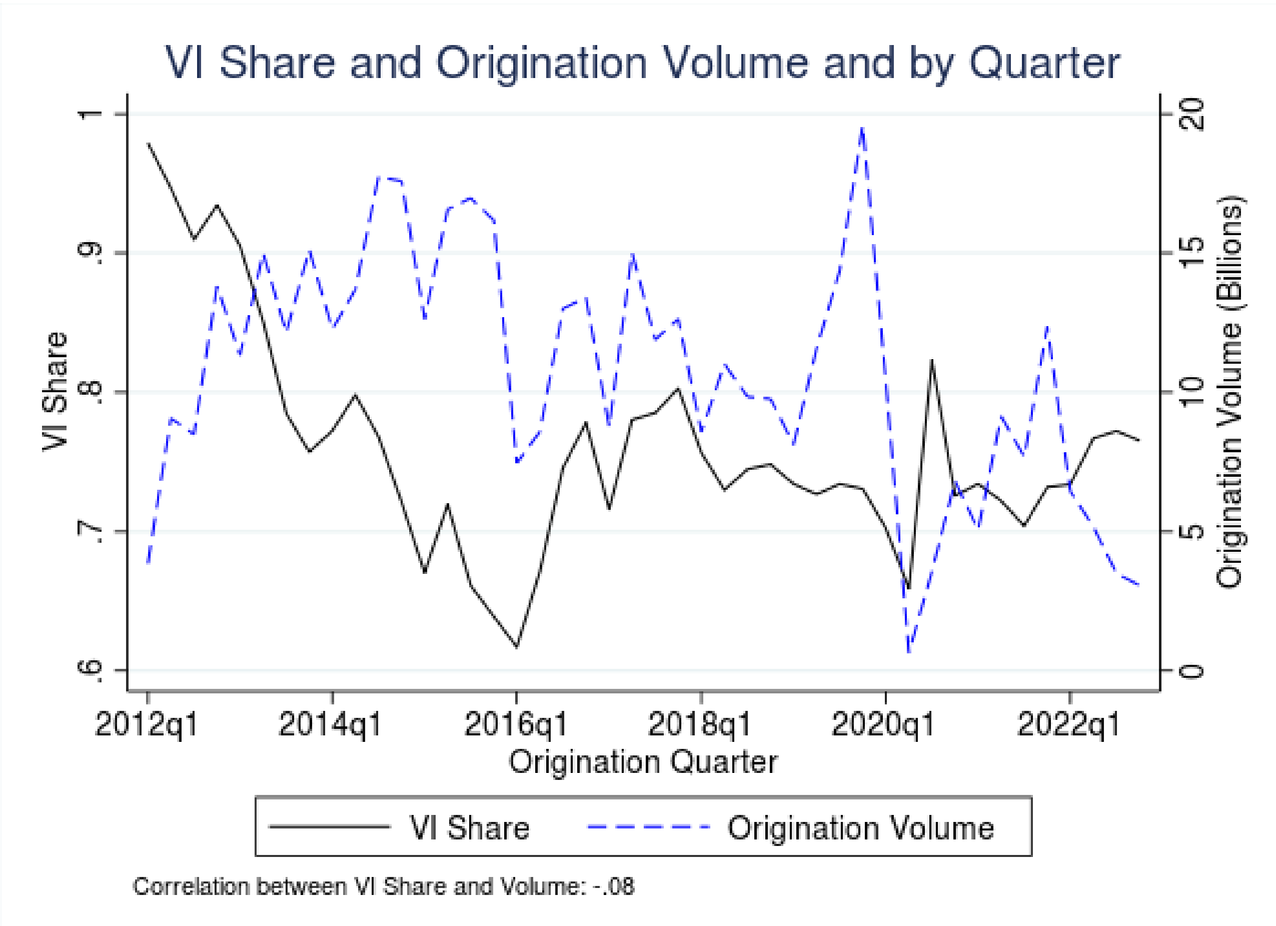


**Figure 9:** VI Share and Origination Volume

## 6.2 Securitization Profitability

Securitizers face an important choice in deciding which loans are included in the pool that will back the securities issued. Securitizers choose the loan pools in a way to maximize the profitability of the issuance, which is directly tied to the risk of the underlying pool through the bond ratings the CMBS securities receive from the ratings agencies. To achieve this, securitizers attempt to diversify the pool as much as possible along a variety of dimensions, which is difficult given the relatively small number of loans, which is 57 loans per pool on average. VI securitizers can increase the diversification of their pools by working with non-VI lenders, and this is particularly true if the non-VI lenders specialize in different markets than the VI securitizer. We show that non-VI share is related to other measures of diversification that we can compute, which are property type HHI, property geography HHI, and loan HHI of the pool. In figure 10, we plot binscatters of the non-VI share of a pool against these three measures of pool diversification. Across all three measures, we see that as non-VI share increases, so does the measure of diversification.

We explore how these measures of diversification are related to the required yield on the securities. As our left hand side measure, we compute a yield spread by taking the difference between the weighted average YTM on the bonds in the issuance and the risk-free rate, which can be interpreted as a risk-premium required for holding that issuance of CMBS. We use this as our measure because participants told us that this is what they attempt to maximize the difference between weighted average loan rates and weighted average YTM. For a fixed weighted average loan rate and risk-free rate, the lower the YTM, the higher the securitization profits [31]. We use the following specification to show the relationship between the profitability and the diversification measures:

$$y_{i,j,t} - rf_t = \beta * Div_j + \omega Z_j + \psi_i + \psi_t + \epsilon_{i,j,t} \tag{6}$$

The unit of observation is the issuance $j$ from issuer $i$ at time $t$. $y_{i,j,t}$ is the weighted average yield to maturity (YTM) on the securities from the deal, $rf_t$ is the 10-year swap rate on the date the security was price, and $Div_j$ are measures of loan pool diversification. Our controls $Z_j$ include risk controls: weighted-average loan rate, weighted-average maturity, weighted-average DSCR, weighted-average LTV, interest-only loan share, whole loan share, and dummies for risk retention type; pool

[31] This can easily converted into a lump sum profit at securitization by calculating the present value of an annuity, where the number of periods is the maturity of the loans

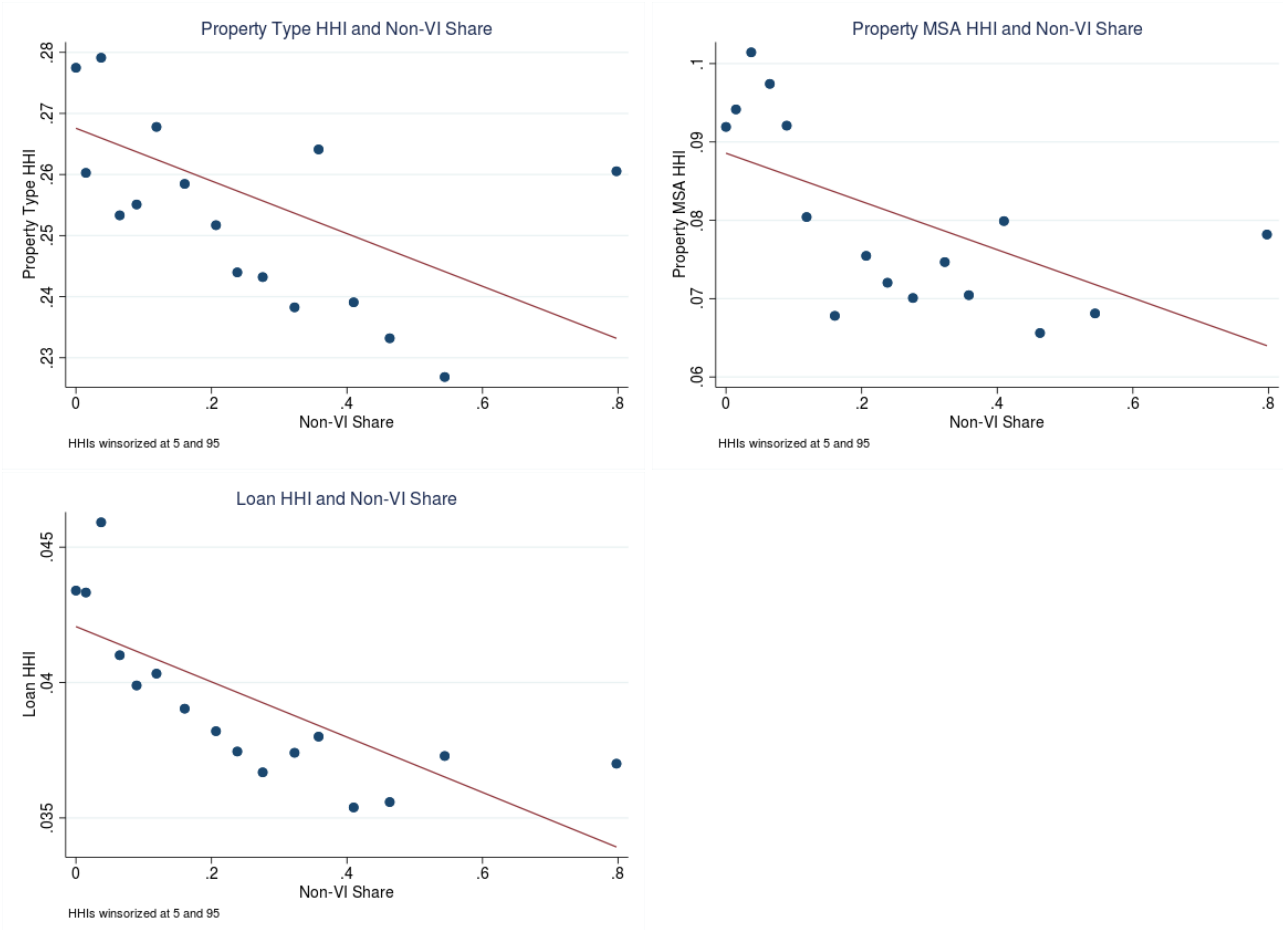


**Figure 10:** Non-VI Share and Diversification

size controls: log total loan amount[32]. We include closing quarter fixed effects to control for market conditions at securitization and issuer fixed effects to control for differences in issuer quality.

In table 5 we show the results of specification 6. In column 1, we see that loan amount HHI is significantly correlated with the yield spread, after controlling for risk controls, pool size controls, and issuer and time fixed effects[33]. In column 2, we include other measures of pool diversification, and while the coefficients point in the correct direction, they are statistically insignificant, while the coefficient on loan HHI remains statistically significant.

The results in table 5 are both economically meaningful and assuring. 1 SD decrease in loan HHI

[32] See table 14 for definitions of all of these variables

[33] An issuer, or a "shelf" is a group of loan contributors that repeatedly collaborate to issue CMBS. One example is "Benchmark Mortgage Trust" which consists of Citigroup, Deutsche Bank, Goldman Sachs, and JP Morgan

**Table 5:** Yield to Maturity and Diversification

| | (1)<br>YTM-Rf | (2)<br>YTM-Rf |
|---|---|---|
| Loan Amount HHI | 6.250*** | 5.044** |
| | (2.390) | (2.453) |
| Property Type HHI | | 0.550 |
| | | (0.391) |
| Property Region HHI | | 0.304 |
| | | (0.482) |
| Observations | 446 | 446 |
| $R^2$ | 0.581 | 0.583 |
| Fixed Effects | I,T | I,T |
| Risk Controls | Y | Y |
| Pool Size Controls | Y | Y |

Standard errors in parentheses
Fixed Effect Codes: I=Issuer,T=Quarter
Robust Standard Errors Reported
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

(.0085), combined with the estimate from column 1, will lead to a lower yield of 5.3bps. This translates to an average increase in securitization profits of \$3.43M [34]. In column 2, we see that loan amount HHI is still the most relevant diversification measure. This is assuring because the securitizer tries to maximize the amount of highly rated (AAA) securities allowed by the rating agencies. When looking through Moody's CMBS rating methodology, the measure of diversification they discuss most is "Herf score" which is defined as $\frac{1}{LoanAmountHHI}$, and we observe an empirical connection between yield spread and the stated measure of diversification by the rating agencies. It is also important to keep in mind that the sample size is small and many coefficients to estimate, so statistical significance is difficult to achieve.

## 7 An Empirical Model of the CMBS Market

Being a vertically integrated lender in this market has the advantage that they have control over which loans to put in each pool. This means that they get the final say regarding which loans will

[34] A decrease in YTM of 0.053, combined this with the average loan maturity of 113 months, the average pool amount is \$986M, and monthly payments, and securitizer cost of capital of 8%, leading to an increase in securitization profit of $986M * \frac{.00053}{12} * \frac{1-(1+\frac{.08}{12})^{-112}}{\frac{.08}{12}} = \$3.43M$

go in a given pool. In contrast, non vertically integrated lenders are the residual space claimants of the issuance. In short, non-VI lenders can only securitize loans if the securitizer has some space left for these loans. One of the main direct benefits of this is that the warehousing costs of vertically integrated lenders should be reduced on average as they can always prioritize their loans if needed, even if this comes at the expense of lower securitization profits per loan. Given that all of the loans originated in our sample are conduit loans, or loans made with the intention to securitize, all lenders are always trying to minimize the risk associated with keeping the loan on their balance sheet until securitization, or warehousing risk.

The timing of the model is as follows:

1. Projects/Borrowers arrive to the market to obtain funding each period
   - Discrete choice over lenders (posted prices) within a market and outside option (other funding sources)
   - VI and non-VI lenders compete via posted prices
2. Loans originated and lenders' balance sheet evolve accordingly
3. VI lenders decide whether to securitize or not at the end of the current period
   - They choose among the available bundles that include the non-VI loans
   - Loans are packaged and sold to secondary market investors (competitive market)
4. Loans securitized leave balance sheets at the end of the period and next period begins

## 7.1 Borrower Side

Each loan is indexed by $i$ and its purpose is to finance the purchase of a property by borrower $b$. Loans arrive to the market randomly, which generates lumpiness. We define a market $m$ as a geographic region of the US $g$, a property type $p$, and a period of time $t$. A loan arrives to market $m$ with a probability given by $\lambda_m$. Once a loan arrives to the market, borrowers shop for loans between the available lenders in order to get financing for their project. Preferences for a borrower whose loan is $i$ from bank $j$:

$$u_{ijm} = -\alpha r_{ijm} + \xi_{jm} + Z_i\beta + \epsilon_i^A + \epsilon_i^S + \epsilon_{ijm}$$

where the term $r_{ijm}$ is the interest rate that loan $i$ is charged by lender $j$ in market $m$. The term $\xi_{jm}$ captures quality or service characteristics that bank $j$ offers in market $m$ that all borrowers agree on and observe, but that are not observed by us. $Z_i$ are loan-specific characteristics that shift the utility of the inside good. These are analogous to demographics in the discrete choice literature[35]. $\epsilon_i^A$ is a shifter of inside utility that is not observed to either the lender or the econometrician, which we will call asymmetric information following Crawford et al. (2018b). $\epsilon_i^S$ is soft information that shifts the utility of the inside good, which is observed to the lender but not the econometrician. Lastly, the term $\epsilon_{ijm}$ captures borrower specific preferences for bank $j$ in market $m$. This last term captures horizontal differentiation between banks in a market since not all borrowers agree on.

Given this preferences, and the rate of arrival of new projects to the market, the ex-ante probability that lender $j$ originates loan $i$ in market $m$ is given by:

$$D_{ijm} = \lambda_{im}\mathbb{P}_{ijm}$$

where the term $\mathbb{P}_{ijm}$ is characterized in section 8. We will assume that the utility of not obtaining a loan $u_{i0m}$ is 0. We denote by $N_{jmt}$ the total number of loans from market $m$ that lender $j$ has originated in period $t$, and we remove the $t$ subindex when we refer to all loans originated up to that point that have not been securitized yet.

## 7.2 Lender Side

In each market $m$ there are $\mathcal{J}_m$ lenders. Each lender can be categorized as either VI or nVI for Vertically-Integrated or non-Vertically-Integrated respectively. In every market, each of the $\mathcal{J}_m$ lenders submit bids $r_{ijm}$ to originate loans. Every period, the VI lenders decide whether or not to securitize. If they decide to securitize, they will empty their balance sheet, and they have to decide how much of each type of loan from the nVI lenders to include in the pool. If they do not securitize, all originated loans up to that point that have not been securitized, including those originated in the current period, are kept in the balance sheet for the next period.

[35] In discrete choice, $X_{jm}$ refers to product specific characteristics, which in our setting would be lender characteristics. In our model, we have no lender characteristics and so we don't include $X_{jm}$

The flow payoffs of loans that are not securitized yet (and therefore on the balance sheet) are given by:

$$\pi_j^b = \sum_{k \in N_j} n_k(r_{k,j} - c_{j,m}(B_{j,t}))$$

where basically the per-dollar flow profits of the originated loans is given by the mark-up term $r_{k,j} - c_{j,m}(B_{j,t})$, and this is multiplied by the loan amount $n_k$. $c_{j,m}(B_{j,t})$ is a bank-market specific per-dollar balance sheet cost that captures the warehousing risks and the funding costs for this loans while they are kept in the balance sheet of the originating bank. The marginal costs is allowed to vary at the bank-market level, and it directly depends on $B_{j,t}$ which summarizes the state of the bank's balance-sheet. The flow payoff that bank $j$ gets when it decides to securitize is given by:

$$\pi_j^S(y(\chi_{j,t})) = \sum_{k \in N_j} n_k(r_{k,j} - y(\chi_{j,t}))\Phi$$

Note that the profits take a similar form than the profits from origination, except that the markup term now includes the cost of funds $y(\chi_t)$ which is interpreted as the yield that investors demand to hold the bonds made of from the underlying loans in the security, and $\Phi$, an annuity factor that transforms this per-dollar mark-up into a lump sum transfer equivalent to its present value. $\chi_{j,t}$ corresponds to the choice of loans included in the securitization. The choice of $\chi_{j,t}$ affects the the cost of funds $y(\chi_t)$ through the correlation structure of loan performances, so the securitizer can reduce the rate demanded by investors through diversification of the loan pool. We use the following asset pricing model whereby the yield required is a mapping from the variance of the underlying returns of the loan pool into a single interest rate that investors demand in order to hold the bond. The expression is below.

$$y(\chi_{j,t}) - r_f = f(\chi_{j,t})$$

Where $r_f$ is the risk free rate, and $f(\chi_{j,t})$ is a function of the characterstics of the loan pool. One example is $f(\chi_{j,t}) = \beta * HHI_{loan}(\chi_{j,t})$, which would assume that the only factor relevant for security pricing is loan HHI[36]. The objective function of the securitizer or VI lender can be written as follows:

$$V_j(\mathbb{X}_t) = \max_{\{r_{jt}>0, \chi_{j,t}\in\mathbb{X}_t\}} \mathbb{E}\left\{ \sum_{t=\tau}^{\infty} \beta^{\tau-t} \left[\pi_j^b(r_{\tau,j}) + \pi_j^S(\chi_{j,t})\right] \middle| \mathbb{X}_t \right\}$$

and the evolution of the balance sheet is as follows:

$$B_{j,t} = B_{j,t-1} + N_{j,t} - \mathbf{1}_{sec_{j,t}}(B_{j,t-1} + N_{j,t})$$

were $\mathbf{1}_{sec_{j,t}}$ is an indicator equal to 1 if the lender decided to securitize the loans it has on their balance sheet, and 0 otherwise. $\mathbb{X}$ summarizes the state space which essentially consists of the loans originated by all actors up to this point in time. Note that we have made the assumption that when a VI lender decides to securitize, they empty their balance sheet and start the next period with zero loans. In case they decide to securitize, they get the securitization profits $\pi_j^S(y(\chi_{j,t}))$. Recall that this term involves choosing the optimal bundle of loans to be included in the pool as detailed above. The corresponding Bellman equation is given by:

$$V_j(\mathbb{X}) = \max_{\{r_j>0, \chi_j\in\mathbb{X}\}} \left[\pi_j^b(r_j) + \pi_j^S(\chi_j)\right] + \beta V_j(\mathbb{X}')$$

The optimal strategies of the VI lender are given by the following first order conditions or differences. For pricing we have:

$$\frac{\partial V_j(\mathbb{X})}{\partial r_j} = \underbrace{\frac{\partial \pi_j^o(r_j)}{\partial r_j}}_{\text{Standard Static Profits}} + \underbrace{\beta \frac{\partial V_j(\mathbb{X}')}{\partial \mathbb{X}'} \frac{\partial \mathbb{X}'}{\partial r_j}}_{\text{Change in Continuation Value}}$$

where the lender is choosing $r_{jm}$ changing its origination profits today, which is captured by the first term, but also affecting its future profits by changing their balance sheet costs tomorrow, which is captured by the second term. This is what introduces dynamic linkages in pricing incentives of the lender, and allows them to be more or less aggressive in their pricing strategies depending on what they have already originated up to this period and what they expect to be able to originate next period, conditional on their balance sheet costs. For securitizing the lender chooses the bundle that provides the highest bundle. When comparing two bundles, they compare:

[36] Under certain assumptions, this expression is connected to a classic asset pricing expression derived from investors with CARA utility, because loan HHI can be mapped into a variance of the pool of loans

$$\Delta_{\chi,\chi'} = \pi_j^S(y(\chi)) - \pi_j^S(y(\chi^{'})) + \beta \left[V(\mathbb{X}) - V(\mathbb{X}^{'})\right]$$

which essentially a comparison of profits in the two scenarios. Two bundles have a difference in securitization profits $\pi_j^S(y(\chi)) - \pi_j^S(y(\chi^{'}))$ today and difference in continuation values $V(\mathbb{X}) - V(\mathbb{X}^{'})$ next period. Embedded in the difference in continuation values is the effect of vertical foreclosure. If the VI lender chooses a bundle that does not include many of nVI loans, they will raise the nVI costs and rates, which will soften competition and therefore increase the continuation value, compared to a bundle of loans that includes many of the nVI loans.

## 8 Estimation

Since this is a market in which we only observe the prices of the signed contracts, we need to compute the prices that were available in the choice set of each loan that was originated in order to estimate demand with a model of posted prices. Following Crawford et al. (2018b) we leverage the existence of multiple loans per borrower in order to take into account that there could be soft information observed by the lenders, but not the econometrician. Additionally, since our model has dynamic pricing incentives, our model for constructing these prices has to take into account that at different points in time, the same lender could be more or less aggressive with its pricing policy. We compute prices with the following three step procedure. First, we need to build the choice set for each loan. We do this by assuming that all lenders that have made loans in that market are available. Second, we regress prices, as measured by their spread with respect to the corresponding swap rate, onto loan characteristics and adding a rich battery of fixed effects. We define a market $m$ as a geographic region of the US $g$, a property type $p$, and a period of time $t$. Lenders are indexed by $j$, while borrowers are indexed by $b$ and loans by $i$. The empirical specification is as follows:

$$\begin{aligned} r_{ijm} &= \tilde{r}_{ijm} + \tilde{\nu}_{ijm} \\ &= \tilde{\psi}_{jm} + \tilde{\omega} Z_i + \tilde{\psi}_b + \tilde{\nu}_{ijm} \\ &= \tilde{r}_{jm} + \tilde{\omega} Z_i + \tilde{\psi}_b + \tilde{\nu}_{ijm} \end{aligned}$$

where $Z_i$ are the same controls used in the pricing regressions. Note that the fixed effect $\psi_{jm}$ is a lender-property type-geography-period fixed effect. This fixed effect is essentially what we refer to as the "posted price of a lender", which is why we will relabel it as $\tilde{r}_{jm}$. It only needs to be adjusted by the specific loan characteristics contained in $Z_i$ and the borrower fixed-effect $\psi_b$ for us to be able to construct the counterfactual prices. The borrower fixed effect captures soft information about the borrower that might be observed by the lenders, but not by us, and it might influence pricing decisions. Once we take into account that we have to estimate the counterfactual prices, the utilities take the following form:

$$\begin{aligned}
u_{ijm} =& -\alpha r_{ijm} + \xi_{jm} + \epsilon_i^A + Z_i\beta + \epsilon_i^S + \epsilon_{ijm} \\
& -\alpha r_{ijm} + \xi_{jm} + \epsilon_i^A + Z_i\beta + \omega\hat{\psi}_b + \epsilon_{ijm} \\
=& -\alpha\left(\tilde{r}_{jm} + \tilde{\gamma}Z_i + \tilde{\psi}_b + \tilde{\nu}_{ijm}\right) + \xi_{jm} + \epsilon_i^A + Z_i\beta + \omega\hat{\psi}_b + \epsilon_{ijm} \\
=& \underbrace{-\alpha\tilde{r}_{jm} + \xi_{jm}}_{\tilde{\delta}_{jm}} + \underbrace{\overbrace{(-\alpha\tilde{\gamma} + \beta)}^{\nu_1} Z_i + \overbrace{(\omega - \alpha)}^{\nu_2}\tilde{\psi}_b + \epsilon_i^A}_{V_{ijm}} + \underbrace{(-\alpha\tilde{\nu}_{ijm} + \epsilon_{ijm})}_{\varepsilon_{ijm}} \\
=& \underbrace{-\alpha\tilde{r}_{jm} + \xi_{jm}}_{\tilde{\delta}_{jm}} + \underbrace{\nu_1 Z_i + \nu_2\tilde{\psi}_b + \epsilon_i^A}_{V_{ijm}} + \underbrace{(-\alpha\tilde{\nu}_{ijm} + \epsilon_{ijm})}_{\varepsilon_{ijm}} \\
=& \tilde{\delta}_{jm} + V_{ijm} + \varepsilon_{ijm}
\end{aligned}$$

where $\epsilon_i^A$ is asymmetric information that is related to demand for credit, meaning neither the lenders or the econometrician can observe it. We will assume $\epsilon_i^A \sim N(0, \sigma_{\epsilon^A})$ . This is a random coefficient on the constant in the terminology of the discrete choice literature. $\epsilon_i^S$ is soft information that is related to the demand for credit, meaning it is observed by the lenders but not the econometrician. Following Crawford et al. (2018b), we will approximate this with the estimate of the borrower fixed effect, multiplied by a parameter: $\epsilon_i^S = \omega\hat{\psi}_b$.

Now we can recover the estimates for for $\{\tilde{\delta}_{jm}, \theta\} = \{\tilde{\delta}_{jm}, \nu_1, \nu_2, \sigma_{\epsilon^A}\}$ using SMLE, following an approach of Goolsbee and Petrin (2004). Under the assumption that $\varepsilon_{ijm}$ follows a type I extreme value distribution, the probability of choosing lender $j$ in market $m$ is given by the mixed logit expression:

$$\mathbb{P}_{ijm}(\tilde{\delta}_{jm}, \nu_1, \nu_2, \sigma_{\epsilon^A}) = \int \underbrace{\left[\frac{exp(\tilde{\delta}_{jm} + V_{ijm})}{1 + \sum_l exp(\tilde{\delta}_{lm} + V_{ilm})}\right]}_{L_{ijm}(\theta, \sigma_{\epsilon_i^A})} f(\epsilon_i^A | \sigma_{\epsilon^A})\, d\alpha_i = \int L_{ijm}(\theta, \epsilon_i^A) f(\epsilon_i^A | \sigma_{\epsilon_i^A})\, d\epsilon_i^A \tag{7}$$

After recovering estimates of $\tilde{\delta}$, we can use the following specification, along with a valid instrument, to recover $\alpha$, which governs the price elasticity of demand.

$$\hat{\tilde{\delta}}_{jm} = -\alpha * \widehat{r_{jm}} + \epsilon_j + \epsilon_m + \Delta\epsilon_{jm} \tag{8}$$

For a complete description of the estimation, see section B.

# 9 Alternative Explanations

In this section of the paper, we will attempt to show that the main results are robust and not solely driven by alternative explanations.

## 9.1 Unobserved Lender Quality

One natural concern about our findings is that VI lenders are fundamentally different from non-VI lenders and that these differences could be entirely driving differences in spreads. For example, VI lenders are often deposit taking institutions and so they are able to access a lower cost of funds through deposits and therefore are able to charge lower spreads on their loans. To address these concerns, we can look at variation of spreads and DTS within originator to show that when the same originator is VI, they have lower spreads and lower DTS than when they are non-VI[37].

In table 6, we show the baseline results with originator fixed effects. In our preferred specification, column (2), we see that VI loans have spreads that are 7bps smaller than non-VI loans from the same lender, which is remarkably similar to our baseline results, though they are only marginally statistically significant. Our estimates for DTS are slightly larger in magnitude. It is important to keep in mind that in our sample, eight originators have both VI and non-VI loans, and they tend to mostly be VI

[37] From conversations with industry participants, a lender can become an underwriter if they are contributing enough loans to the pool. Lenders tend to be VI when they originate larger number of loans

lenders. The coefficient on VI is therefore determined by these eight lenders and therefore precision of the estimate may be low.

**Table 6:** Spreads and Days to Securitization for VI and Non-VI

| | (1) SPRD | (2) SPRD | (3) DTS | (4) DTS |
|---|---|---|---|---|
| VI | -0.037 | -0.071* | -19.066*** | -11.870* |
| | (0.027) | (0.039) | (4.280) | (6.176) |
| Observations | 17559 | 8265 | 17590 | 8273 |
| $R^2$ | 0.738 | 0.862 | 0.391 | 0.601 |
| Fixed Effects | PxT,G,O | PxT,G,O,B | PxT,G,O | PxT,G,O,B |
| Controls | Y | Y | Y | Y |

Standard errors in parentheses
Standard errors clustered at the MSA level
Fixed effects codes: P=Prop. Type,T=Orig. Month,G=Prop. MSA,B=Borrower, O=Originator
Controls: LTV, DSCR, Log Size, Debt Yield, Log Price/Sq Ft, and Cap Rate
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## 9.2 Unobserved Borrower Quality

One potential explanation for our results is that our pricing model is not capturing unobservable (soft) information that is observed by the lenders and priced in. If this unobservable quality differs by lenders, this may be driving the difference in spreads. For example, if non-VI lenders make loans for projects that have more unobservable (to the econometrician) risk than VI lenders, we would expect them to be charging higher interest rates to compensate for this unobservable risk. We will test for this in two ways. First, we will test whether our pricing model is missing unobservable information by conducting an information test. Second, we will explore differences in ex-post outcomes and how these ex-post outcomes are related to pricing (this will be done in section 9.3).

We conduct a soft information test, following (Chiappori and Salanié, 2000; Crawford et al., 2018b; Stillerman, 2021) . This tests whether the pricing residual for loans, which contains the unobservable quality, has predictive power of the ex-post outcomes. If our pricing model is not omitting any relevant variable for loan risk (i.e. the true pricing model only includes hard information), then the residuals will have no predictive power of ex-post outcomes.

Formally, the information test is a two stage procedure where in the first stage we predict prices on our risk controls and fixed effects. In the second stage, we use that residual in a regression of ex-post outcomes (i.e. default). The coefficient on the residual will capture whether non-VI loans are pricing based on soft information. If we see statistically positive coefficient on the residual, this would imply our pricing model is missing unobservable quality characteristics.

$$r_{i,j,t} = \underbrace{\omega * Z_i + \psi_{pt} + \psi_g + \psi_b}_{Hard\ Info.} + \underbrace{\epsilon_{i,j,t}}_{Soft\ Info.\ +\ Noise} \tag{9}$$

$$o_{i,j,t} = \tilde{\omega} Z_i + \tilde{\psi_{pt}} + \tilde{\psi_g} + \tilde{\psi_b} + \underbrace{\delta \widehat{\epsilon_{i,j,t}}}_{Soft\ Information} + \upsilon_{i,j,t} \tag{10}$$

Here the outcome of interest are indicators variables for ex-post outcomes. In particular, we use whether the loan was ever late (Late) or whether the loan was ever at least 90 days late (Sev. Del)

In table 7 we show the results of specification 10, with and without borrower fixed effects[38]. We only observe ex-post outcomes for the subset of the sample obtained from ABS-EE, which begin in November 2016. Additionally, the inclusion of borrower fixed effects significantly reduces the number of observations, so we include a specification without borrower FE, but for the sample for which we include borrower FE so we can make conclusions about the addition of borrower FEs, rather than just differences in sample.

The first thing to notice is that in specifications without borrower fixed effects, we do find evidence that our pricing model is missing measures of unobservable loan quality. For example, in columns 1 and 2 for Late, we see that the coefficient on the residual is positive and statistically significant, implying that our pricing model is missing some soft information that the lenders are using when pricing the loans. In column 3, where we control for borrower risk with borrower fixed effects, the coefficient on the residual shrinks and becomes statistically insignificant, which implies that we cannot reject the hypothesis that the residuals have no effect on default. Results are similar for the severe delinquency.

[38] We only have ex-post outcome data on data from November 2016 onwards due to data limitations in the Bloomberg data. We denote this sample in regression tables as "2017-2022"

This analysis suggests that our baseline pricing model without borrower fixed effects suffer from omitted soft information. When including borrower fixed effects, this issues is vastly reduced and we cannot reject the hypotheses that our pricing models are not systematically missing soft information related to delinquency. With this in mind, it is important to look at our baseline results with borrower fixed effects to more precisely look at the differences in rates.

**Table 7:** Soft Information Test

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Late | Late | Late | Sev. Del. | Sev. Del. | Sev. Del. |
| Residual | 0.102*** | 0.082*** | 0.027 | 0.062*** | 0.064*** | 0.011 |
| | (0.014) | (0.028) | (0.040) | (0.008) | (0.019) | (0.022) |
| Observations | 9068 | 2770 | 2770 | 9068 | 2770 | 2770 |
| $R^2$ | 0.276 | 0.375 | 0.685 | 0.223 | 0.320 | 0.660 |
| Fixed Effects | PxT,G | PxT,G | PxT,G,B | PxT,G | PxT,G | PxT,G,B |
| Controls | YES | YES | YES | YES | YES | YES |
| Sample | 2017-22 | 2017-22,B FE | 2017-22,B FE | 2017-22 | 2017-22,B FE | 2017-22,B FE |

Standard errors in parentheses

Standard errors clustered at the MSA level

Fixed effects codes: P=Prop Type,T=Orig Month,G=Prop MSA,B=Borrower

$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## 9.3 Differential Unobserved Borrower Quality Across Lenders

As discussed earlier, a natural concern to our empirical results are that unobserved loan quality differs by VI and non-VI lenders, conditional on our controls and fixed effects. To test this, we are first going to directly look at delinquency rates, conditional on observables by using specification 2 with ex-post outcomes as the left hand side variable. In table 8, we show the results

As before, we include both specifications with and without borrower FE, and also an intermediate specification without borrower FE but limited to the sample that includes borrower FE. Here we see that across all specifications, there is no statistically significant difference in ex-post outcomes across VI and non-VI, conditional on obervables.

Another way show that there is little difference in unobservable loan quality between VI and non-VI loans is by including ex-post outcomes as a control in the pricing regression (equation 2) and see how the coefficient on the VI dummy changes. If ex-post outcomes are correlated with unobservable loan

**Table 8:** Delinquency and VI Status

| | (1) Late | (2) Late | (3) Late | (4) Sev. Del. | (5) Sev. Del. | (6) Sev. Del. |
|---|---|---|---|---|---|---|
| VI | -0.004 | -0.011 | -0.008 | -0.007 | -0.014 | -0.019 |
| | (0.013) | (0.026) | (0.032) | (0.006) | (0.011) | (0.018) |
| Observations | 9068 | 2770 | 2770 | 9068 | 2770 | 2770 |
| $R^2$ | 0.272 | 0.373 | 0.685 | 0.215 | 0.315 | 0.660 |
| Fixed Effects | PxT,G | PxT,G | PxT,G,B | PxT,G | PxT,G | PxT,G,B |
| Sample | 2017-22 | 2017-22,B FE | 2017-22,B FE | 2017-22 | 2017-22,B FE | 2017-22,B FE |

Standard errors in parentheses

Standard errors clustered at the city level

Fixed effects codes: P=Property Type,T=Origination Month,G=Property MSA,B=Borrower

$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

quality, then we would expect loans that end up defaulting are priced with higher spreads. Additionally, if VI loans have higher unobservable quality than non-VI loans and the default indicators are capturing this, then we would expect the coefficient on the VI indicator to decrease after including the indicators for default. We show the results in table 9.

**Table 9:** Rates, Delinquency, and VI

| | (1) IR | (2) IR | (3) IR | (4) IR | (5) IR | (6) IR |
|---|---|---|---|---|---|---|
| VI | -0.146$^{***}$ | -0.146$^{***}$ | -0.145$^{***}$ | -0.097$^{***}$ | -0.097$^{***}$ | -0.097$^{***}$ |
| | (0.009) | (0.009) | (0.009) | (0.022) | (0.022) | (0.022) |
| Late | | 0.053$^{***}$ | | | 0.013 | |
| | | (0.008) | | | (0.019) | |
| Sev. Del. | | | 0.145$^{***}$ | | | 0.000 |
| | | | (0.020) | | | (0.040) |
| Observations | 9068 | 9068 | 9068 | 2770 | 2770 | 2770 |
| $R^2$ | 0.862 | 0.863 | 0.864 | 0.946 | 0.946 | 0.946 |
| Fixed Effects | PxT,G, | PxT,G, | PxT,G, | PxT,G,,B | PxT,G,,B | PxT,G,,B |
| Controls | YES | YES | YES | YES | YES | YES |

Standard errors in parentheses

Standard errors clustered at the city level

Fixed effects codes: P= Property Type,T=Origination Month

$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

In column 1 we see the baseline result with the sample of firms for which we observe ex-post outcomes. In column 2 we include an indicator for whether the loan was ever late in our sample. The coefficient on the late indicator is 5bps, which means that loans that end up being late on their payment have spreads that are 5bps higher than what would have otherwise predicted, which is evidence that unobservable quality is priced into the loan. Importantly, we see the coefficient on VI remains virtually unchanged, suggesting there is no difference in unobservable loan quality between VI and non-VI lenders. In column 3 we instead include an indicator for the loan ever being 90 days late and we see the coefficient on this dummy is about 15bps, which suggest that these loans are of very low unobservable quality and that is priced in. As in specification 2, the coefficient on VI remains mostly unchanged. In specifications 4-6, we repeat the analysis with borrower FEs. After including borrower fixed effects, the coefficients on the default dummies become statistically indistinguishable from zero, which suggests that after including borrower fixed effects, unobservable loan quality is not relevant. This is consistent with the results from table 7, which suggests that the pricing model does not suffer from unobservable information after controlling for borrower fixed effects. As before, we also see the coefficient on VI remain virtually unchanged.

## 9.4 Alternative Definitions of VI

In our last robustness, we address a potential concern about our definition of VI. As mentioned earlier, we define VI as any lender that is affiliated with an underwriter on the CMBS security issuance. Within underwriters, there are three different roles defined in our dataset, which are lead bookrunner, bookrunner, and co-manager. While all underwriters are listed on the CMBS prospectus offered to investors, some roles have more responsibilities than others.

Lead bookrunner is the "lead-left" underwriter on the transaction. They have the final say over constructing pools and structuring the securities and are usually responsible with placing a majority of the bonds to investors, and therefore receiving the largest league table credit, though this is not always the case [39]. The bookrunner has a smaller say over the pool construction and they sell some of the CMBS securities with investors. Co-managers play the smallest role with constructing the pool and securities and do not sell any securities to investors.

In table 10 we show the preferred baseline specification for three different definitions of VI, where

[39] Some issuers always have the same lead bookrunner such as Wells Fargo Commercial Mortgage Trust (WFCM), which always has Wells Fargo as lead bookrunner. Others rotate which bank is the lead bookrunner, such as BANK, which rotates Wells Fargo, Morgan Stanley, and Bank of America as lead bookrunners. For issuers that rotate, the only distinction between lead bookrunner and bookrunner is the title

we include borrower fixed effects. These three include our standard definition, an originator and any type of underwriter; a slightly stricter definition, an originator and bookrunner/lead bookrunner; and the strictest definition, an originator and lead bookrunner. Across all definitions, our main results for spreads and DTS remain the same, though their magnitudes vary. In particular, when defining VI with lead bookrunner, the results have smaller magnitudes because the non-VI lender group now occasionally contains large lenders with investment banks, such as JP Morgan, Bank of America, Citigroup, Wells Fargo, Goldman Sachs, and Morgan Stanley.

**Table 10:** Spreads and Days to Securitization for VI and Non-VI

| | (1) SPRD | (2) SPRD | (3) SPRD | (4) DTS | (5) DTS | (6) DTS |
|---|---|---|---|---|---|---|
| VI - Underwriter | -0.085*** | | | -7.077*** | | |
| | (0.011) | | | (1.508) | | |
| VI - Bookrunner | | -0.077*** | | | -11.484*** | |
| | | (0.010) | | | (1.644) | |
| VI - Lead Bookrunner | | | -0.041*** | | | -1.975 |
| | | | (0.009) | | | (1.376) |
| Observations | 8269 | 8269 | 8269 | 8277 | 8277 | 8277 |
| $R^2$ | 0.854 | 0.853 | 0.852 | 0.533 | 0.537 | 0.531 |
| Fixed Effects | PxT,G,B | PxT,G,B | PxT,G,B | PxT,G,B | PxT,G,B | PxT,G,B |
| Controls | Y | Y | Y | Y | Y | Y |

Standard errors in parentheses
Standard errors clustered at the MSA level
Fixed effects codes: P=Property Type,T=Origination Month,G=Property MSA,B=Borrower
Controls: LTV, DSCR, Log Size, Debt Yield, Log Price/Sq Ft, and Cap Rate
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## 10 Conclusion and Future Directions

In this paper, we provide insight of the importance of market structure in securitization on credit, by exploring the CMBS market. We explore how the incentives of the investment bank securitizing the CMBS deal differentially impact the lending the decisions of lenders affiliated with the investment bank, which we call vertically integrated (VI), and lenders not affiliated with the securitizer.

Our main results show that VI loans have lower rate spreads and shorter time from origination to securitization, conditional on observables, and that this result is stronger in cool times, which are periods with low origination volume. We provide evidence that this difference is partially driven by VI lenders prioritizing their own loans, which we call the prioritization channel. We explore implications of the prioritization channel on credit allocation. We also show that the non-VI lenders provide diversification of loan pools, and that this diversification is priced by investors, and so therefore prioritization is reducing pool diversification and securitization profits.

To rationalize the results, we build a dynamic model of vertical integration in securitization and competition. In each period, VI and non-VI lenders compete in the mortgage market for borrowers. Then, the securitizer decides whether to issue a security, and if so, how many non-VI loans to include in the pool for the security issuance. The securitizer balances the benefits of including more of their rivals loans (higher securitization profit), with the costs (lowering their rivals' future lending costs). We plan to empirically estimate the parameters of the model to quantify the impacts of different forces and conduct counterfactuals that change the market structure. First, we would shut down the prioritization incentives and see how rates, quantities, and pool diversification would differ. Then we would conduct counterfactuals introducing GSE involvement. The first would be to introduce a GSE that guarantees the commercial mortgages, as in the agency CMBS market. This would remove the diversification benefit of including more rival loans and might push securitizers to prioritize their own loans more. The second would explore what would happen when introducing a centralized GSE pooling and securitizing these commercial mortgages, as in the RMBS market. This would remove the prioritization incentive and reduce all lenders balance sheet costs, though it may introduce moral hazard for lenders.

## References

Agarwal, S., Chang, Y., and Yavas, A. (2012). Adverse selection in mortgage securitization. *Journal of Financial Economics*, 105(3):640–660.

Allen, F., Carletti, E., Goldstein, I., and Leonello, A. (2015). Moral hazard and government guarantees in the banking industry. *Journal of Financial Regulation*, 1(1):30–50.

Benmelech, E., Dlugosz, J., and Ivashina, V. (2012). Securitization without adverse selection: The case of clos. *Journal of Financial Economics*, 106(1):91–113.

Berry, S. T., Levinsohn, J. A., and Pakes, A. (1993). Automobile prices in market equilibrium: Part i and ii.

Buchak, G., Matvos, G., Piskorski, T., and Seru, A. (2018a). Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy. Technical report, National Bureau of Economic Research.

Buchak, G., Matvos, G., Piskorski, T., and Seru, A. (2018b). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of financial economics*, 130(3):453–483.

Chiappori, P.-A. and Salanié, B. (2000). Testing for asymmetric information in insurance markets. *Journal of political Economy*, 108(1):56–78.

Crawford, G. S., Lee, R. S., Whinston, M. D., and Yurukoglu, A. (2018a). The welfare effects of vertical integration in multichannel television markets. *Econometrica*, 86(3):891–954.

Crawford, G. S., Pavanini, N., and Schivardi, F. (2018b). Asymmetric information and imperfect competition in lending markets. *American Economic Review*, 108(7):1659–1701.

DeFusco, A. A., Johnson, S., and Mondragon, J. (2020). Regulating household leverage. *The Review of Economic Studies*, 87(2):914–958.

Furfine, C. (2020). The impact of risk retention regulation on the underwriting of securitized mortgages. *Journal of Financial Services Research*, 58(2):91–114.

Furfine, C. H. (2014). Complexity and loan performance: Evidence from the securitization of commercial mortgages. *The Review of Corporate Finance Studies*, 2(2):154–187.

Ghent, A. and Valkanov, R. (2016). Comparing securitized and balance sheet loans: Size matters. *Management Science*, 62(10):2784–2803.

Goolsbee, A. and Petrin, A. (2004). The consumer gains from direct broadcast satellites and the competition with cable tv. *Econometrica*, 72(2):351–381.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of political economy*, 94(4):691–719.

Hart, O., Tirole, J., Carlton, D. W., and Williamson, O. E. (1990). Vertical integration and market foreclosure. *Brookings papers on economic activity. Microeconomics*, 1990:205–286.

Jiang, E. X. (2019). Financing competitors: shadow banks' funding and mortgage market competition. *USC Marshall School of Business Research Paper Sponsored by iORB, No. Forthcoming*.

Keys, B. J., Mukherjee, T., Seru, A., and Vig, V. (2010). Did securitization lead to lax screening? evidence from subprime loans. *The Quarterly journal of economics*, 125(1):307–362.

Manove, M., Padilla, A. J., and Pagano, M. (2001). Collateral versus project screening: A model of lazy banks. *Rand journal of economics*, pages 726–744.

Stillerman, D. (2021). Loan guarantees and incentives for information acquisition.

Stroebel, J. (2016). Asymmetric information about collateral values. *The Journal of Finance*, 71(3):1071–1112.

# A Appendix Figures and Tables

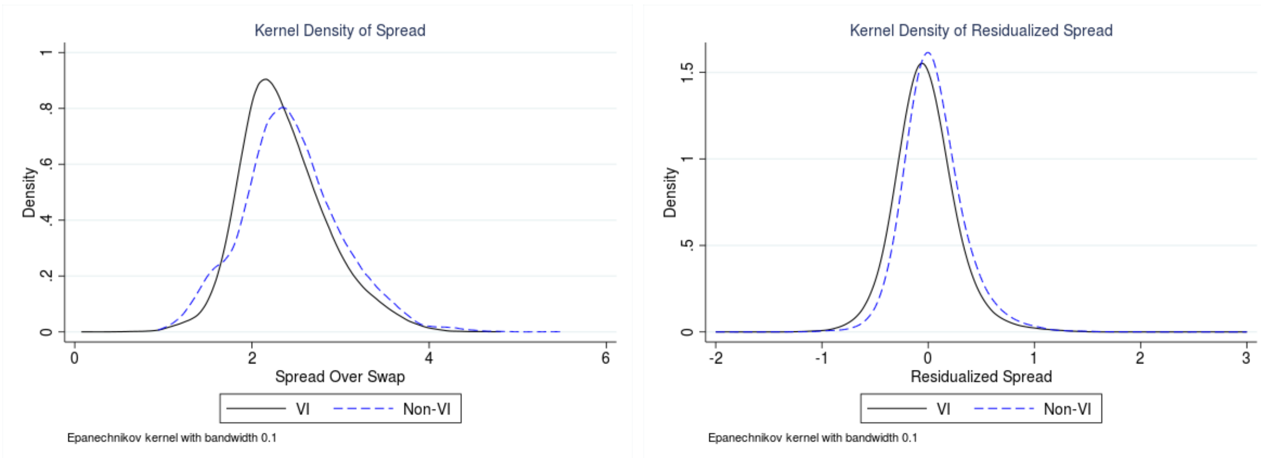


**Figure 11:** Spread and Residualized Spread Distributions

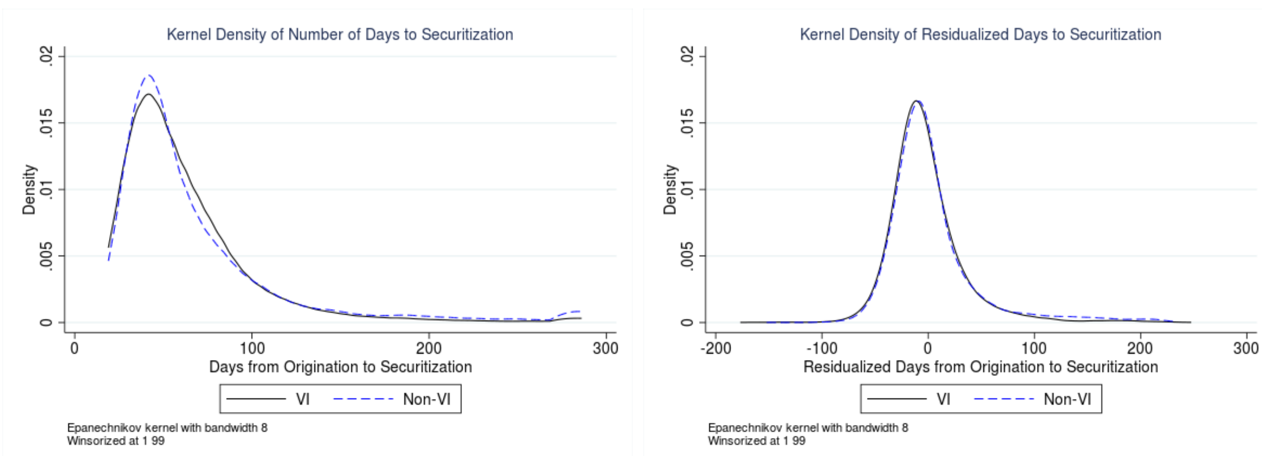


**Figure 12:** Days to Securitization and Residualized DTS Distributions

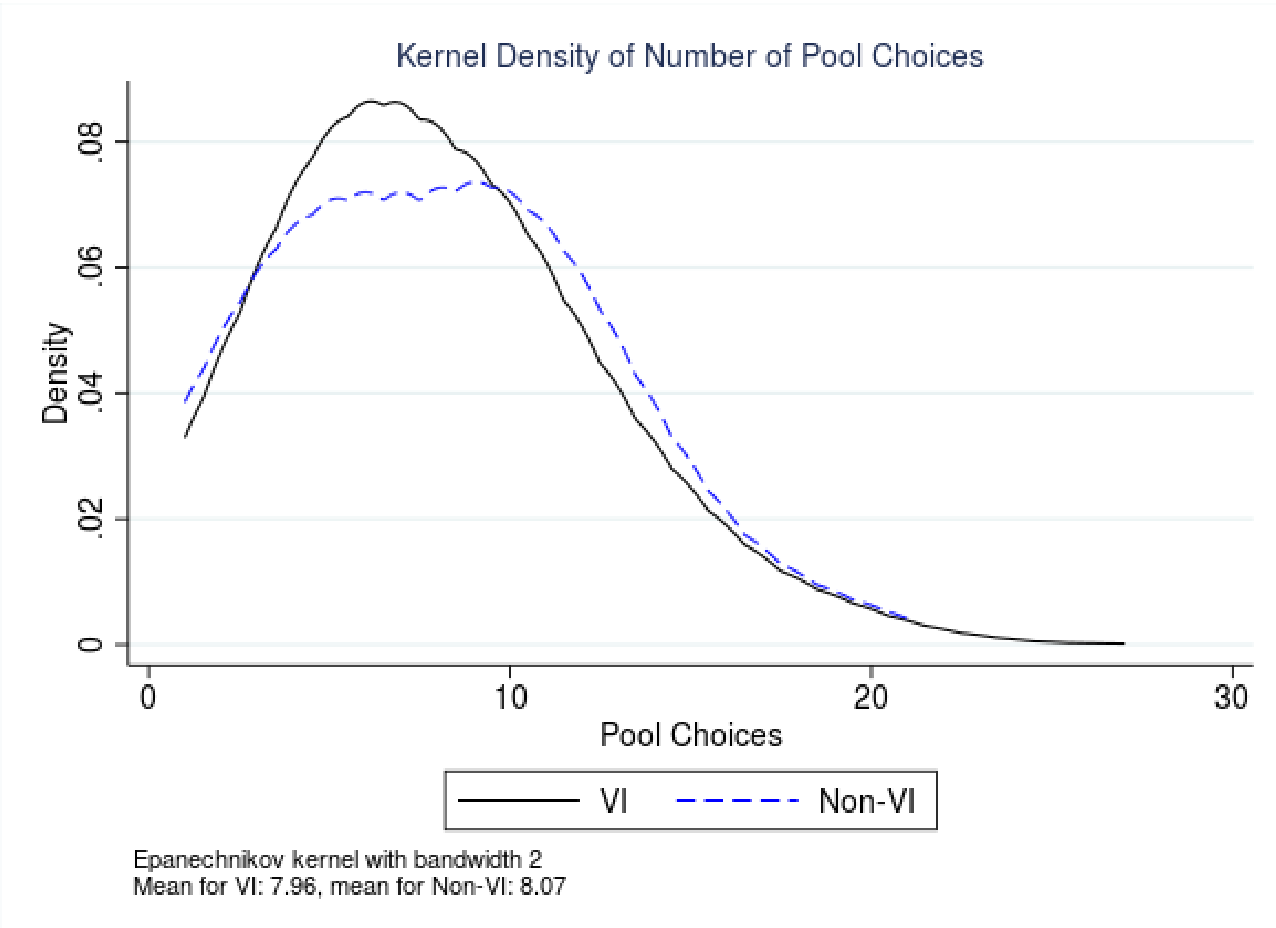


**Figure 13:** Number of Pool Choices for VI and non-VI Loans

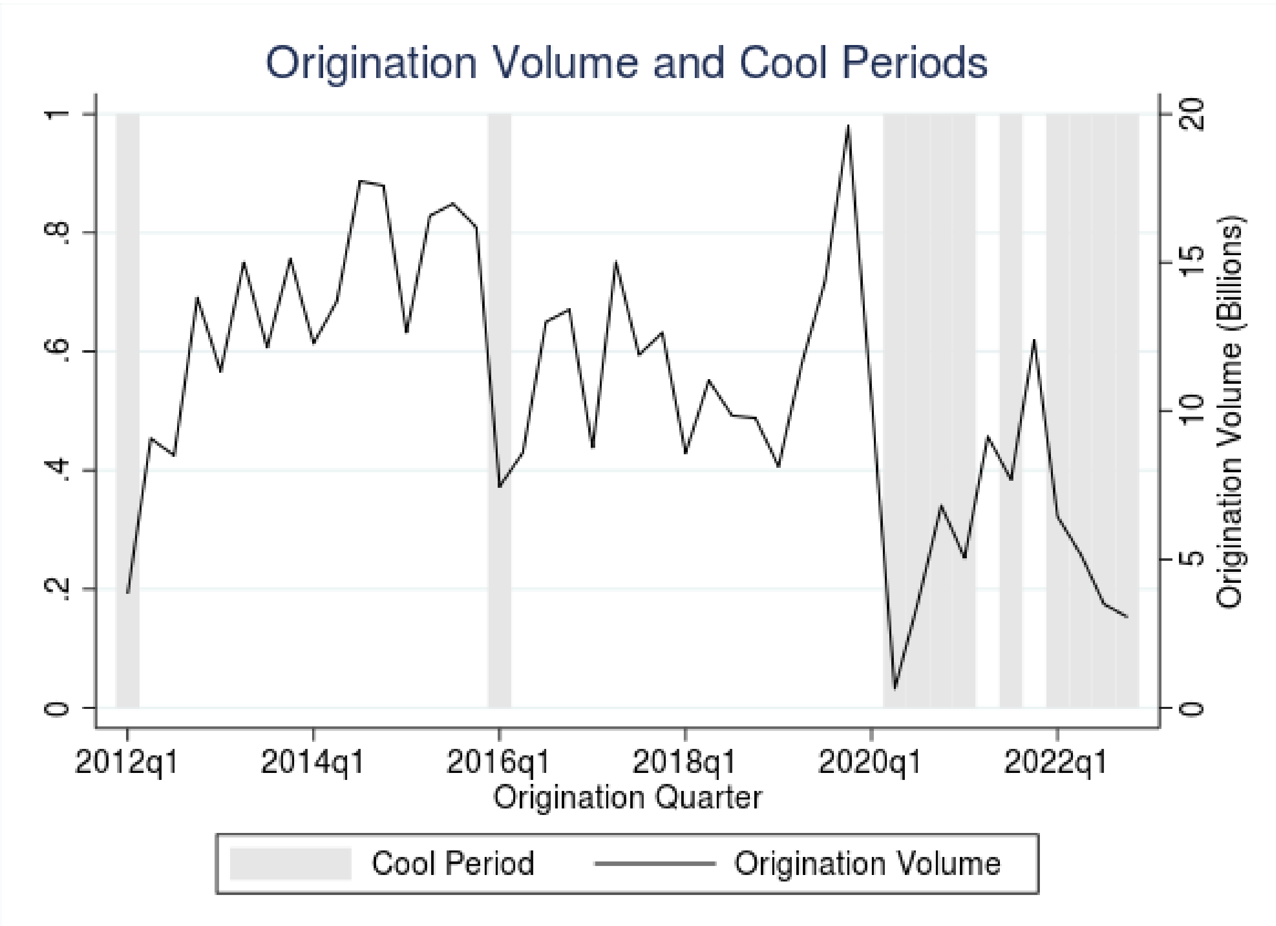


**Figure 14:** Origination Volume and Cool Periods

**Table 11:** Originators and Proportion of Vertically Integrated Loans

| Originator | Total Loans | Proportion VI loans |
|---|---|---|
| Wells Fargo | 1557 | 1.000 |
| Citigroup | 1310 | 1.000 |
| Morgan Stanley | 1215 | 1.000 |
| Deutsche Bank | 972 | 1.000 |
| Starwood Mortgage | 954 | 0.000 |
| Bank of America | 929 | 1.000 |
| CCRE | 894 | 0.944 |
| NCB | 797 | 0.000 |
| Barclays | 734 | 1.000 |
| Ladder Capital | 708 | 0.169 |
| Goldman Sachs | 699 | 1.000 |
| Rialto Mortgage Finance | 695 | 0.000 |
| J.P. Morgan | 686 | 1.000 |
| UBS | 660 | 0.980 |
| KeyBank | 486 | 1.000 |
| Argentic | 455 | 0.000 |
| C-III Commercial Mortgage | 409 | 0.000 |
| CIBC | 300 | 0.993 |
| Benefit Street Partners | 296 | 0.000 |
| Natixis | 291 | 0.976 |
| Societe Generale | 276 | 1.000 |
| LoanCore Capital | 251 | 1.000 |
| RBS | 191 | 0.953 |
| LMF Commercial | 188 | 0.000 |
| Basis Investment | 153 | 0.000 |

**Table 12:** Origination Level Characteristics

| | VI Loans | Non-VI Loans | Difference |
|---|---|---|---|
| Spread (SPRD) | 2.372 | 2.581 | 0.209*** |
| Days to Securitization (DTS) | 61.993 | 67.365 | 5.372*** |
| Loan to Value | 0.629 | 0.649 | 0.020*** |
| Loan Size (millions) | 16.355 | 11.155 | -5.200*** |
| Log(Loan Size) | 16.194 | 15.887 | -0.307*** |
| Value at Securitization (millions) | 26.784 | 17.644 | -9.140*** |
| Debt Service Coverage Ratio | 2.019 | 1.817 | -0.201*** |
| Capitalization Rate | 8.519 | 8.631 | 0.113* |
| Debt Yield | 11.000 | 10.777 | -0.223*** |
| Price/Sq Foot | 341.268 | 284.641 | -56.628*** |
| Log Price/Sq Foot | 5.426 | 5.306 | -0.120*** |
| Late | 0.321 | 0.369 | 0.048*** |
| Sev. Del. | 0.053 | 0.064 | 0.011 |
| Observations | 13936 | | |

**Table 13:** Summary Statistics for Loans and Security Issuances

| | p25 | p50 | p75 | mean | sd | count |
|---|---|---|---|---|---|---|
| Spread (SPRD) | 2.040 | 2.333 | 2.693 | 2.39 | 0.52 | 17618 |
| Days to Securitization (DTS) | 38.000 | 51.000 | 75.000 | 64.32 | 43.87 | 17652 |
| Loan to Value | 0.583 | 0.650 | 0.702 | 0.63 | 0.11 | 17658 |
| Loan Size (millions) | 5.150 | 9.150 | 17.250 | 14.18 | 15.31 | 17678 |
| Log(Loan Size) | 15.455 | 16.029 | 16.663 | 16.06 | 0.83 | 17678 |
| Value at Securitization (millions) | 8.900 | 15.402 | 29.000 | 24.75 | 30.44 | 17660 |
| Debt Service Coverage Ratio | 1.530 | 1.770 | 2.210 | 2.04 | 0.79 | 16553 |
| Capitalization Rate | 6.229 | 8.259 | 10.088 | 8.35 | 2.48 | 17527 |
| Debt Yield | 9.119 | 10.236 | 12.061 | 11.09 | 2.92 | 17558 |
| Price/Sq Foot | 124.180 | 199.368 | 344.930 | 315.32 | 457.23 | 11754 |
| Log Price/Sq Foot | 4.822 | 5.295 | 5.843 | 5.34 | 0.73 | 11754 |
| Late | 0.000 | 0.000 | 1.000 | 0.31 | 0.46 | 9249 |
| Sev. Del. | 0.000 | 0.000 | 0.000 | 0.05 | 0.21 | 9249 |
| YTM-Rf | 0.997 | 1.261 | 1.525 | 1.29 | 0.40 | 455 |
| Non-VI Share | 0.000 | 0.151 | 0.346 | 0.21 | 0.22 | 456 |
| Avg Maturity | 110.000 | 113.000 | 116.000 | 112.57 | 4.31 | 456 |
| Pool DSCR | 1.740 | 1.970 | 2.280 | 2.06 | 0.42 | 456 |
| Pool LTV | 56.900 | 59.950 | 63.900 | 60.17 | 4.49 | 456 |
| Indicator for Horizontal Risk Retention | 0.000 | 0.000 | 0.000 | 0.17 | 0.38 | 456 |
| Indicator for L-Shaped Risk Retention | 0.000 | 0.000 | 0.000 | 0.15 | 0.35 | 456 |
| Indicator for Vertical Risk Retention | 0.000 | 0.000 | 0.000 | 0.19 | 0.39 | 456 |
| Weighted Avg Loan Rate | 4.164 | 4.475 | 4.792 | 4.42 | 0.49 | 456 |
| Interest Only Share | 0.228 | 0.555 | 0.820 | 0.52 | 0.30 | 456 |
| Whole Loan Share | 0.440 | 0.539 | 0.638 | 0.54 | 0.13 | 456 |
| Log Pool Size | 20.509 | 20.683 | 20.863 | 20.68 | 0.24 | 456 |
| Number of Loans | 44.000 | 55.000 | 68.000 | 57.16 | 16.63 | 456 |
| Number of Properties | 75.000 | 98.000 | 128.500 | 106.62 | 41.85 | 456 |
| Property Type HHI | 0.224 | 0.248 | 0.283 | 0.26 | 0.05 | 456 |
| Property MSA HHI | 0.058 | 0.073 | 0.096 | 0.08 | 0.03 | 456 |
| Loan Amount HHI | 0.033 | 0.039 | 0.046 | 0.04 | 0.01 | 456 |

**Table 14:** Variable Definitions

| Variable | Description |
|---|---|
| SPRD | Spread over swap rate, at origination |
| DTS | Days between issuance closing date and origination date |
| Loan to Value | Origination amount / property value at closing |
| Origination amount (millions) | Origination amount in millions |
| Value at Securitization (millions) | Value of property at closing in millions |
| Log(Loan Size) | Log of the origination amount |
| Debt service coverage ratio | Net income / debt service amount at closing |
| Capitalization rate | Net income / value of the property at securitization |
| Debt Yield | Net income / origination amount |
| Log Price/Sq Ft | Log of the price per square foot |
| Late | Indicator for whether the loan was ever late on payment |
| Sev. Del. | Indicator for whether the loan was ever 90+ days late on payment |
| YTM-Rf | Weighted average YTM on securities sold minus 10-yr swap rate on pricing date |
| Non-VI Share | Total Non-VI loan amount / total loan amount in pool |
| Weighted Avg Loan Rate | Loan-size weighted average interest rate in pool |
| Avg Maturity | Loan-size weighted average of maturity of loans in pool, as of securitization date |
| Pool DSCR | Loan-sized weighted average debt service coverage ratio |
| Pool LTV | Loan-sized weighted average loan to value |
| Log Pool Size | Log of total loan amount in pool |
| Number of Loans | Number of loans in pool |
| Number of Properties | Number of properties in pool |
| Property Type HHI | Sum of squared property type shares in pool |
| Property MSA HHI | Sum of squared MSA shares in pool |
| Loan HHI | Sum of squared loan shares in pool |
| Interest Only Share | Total interest-only loan amount/ total loan amount in pool |

# B Demand Estimation Details

In the first stage of demand estimation, we estimate the parameters $\{\theta, \delta\}$ from equation 7 using a simulated maximum likelihood estimate (SMLE). To define the likelihood, we first introduce the variable $\chi_{ijm}$ for whether borrower $i$ chooses lender $j$ in market $m$

$$\chi_{ijm} = \begin{cases} 1 & borrower\ i\ chooses\ lender\ j \\ 0 & otherwise \end{cases}$$

The maximum likelihood is then given by:

$$L(\theta) = \prod_i \prod_j^{J_{m(i)}} \mathbb{P}_{ijm(i)}(\theta)^{\chi_{ijm(i)}} = \prod_i \mathbb{P}_{ij_i m(i)}(\theta)$$

Where $J_{m(i)}$ is the set of lenders in market $m$ for loan $i$ and $P_{ij_i m(i)}$ is the probability of the actual chosen lender for loan $i$. This is because all of the terms that aren't chosen are exponentiated by 0 and therefore become 1.

The log likelihood is then

$$LL(\theta) = \sum_i ln(\mathbb{P}_{ij_i m(i)}(\theta)$$

Since we will only be able to estimate each individual $\mathbb{P}_{ij_i m(i)})(\theta)$ with Monte Carlo, we will instead maximize the simulated log likelihood

$$SLL(\theta) = \sum_i ln(\hat{P}_{ij_i m(i)})(\theta)$$

Where we calculate each individual $ln(\hat{P}_{ij_i m(i)})(\theta)$ using the following (where R is the number of random draws)

$$\begin{aligned} \hat{P}_{ijm}(\theta, \delta_{ijm}) =& \frac{1}{R}\sum_r^R L_{ijm}(\theta, \epsilon^A_{r_i}) = \frac{1}{R}\sum_r^R \frac{exp(\tilde{\delta}_{jm} + V_{ijm})}{1 + \sum_l exp(\tilde{\delta}_{lm} + V_{ilm})} \\ =& \frac{1}{R}\sum_r^R \frac{exp(\tilde{\delta}_{jm} + \nu_1 Z_i + \nu_2 \tilde{\psi}_b + \epsilon^A_{ri})}{1 + \sum_l exp(\tilde{\delta}_{lm} + \nu_1 Z_i + \nu_2 \tilde{\psi}_b + \epsilon^A_{ri})} \end{aligned}$$

where, $\epsilon^A_{ri}$'s are draws from the distribution $f(\epsilon^A_i|\sigma_{\epsilon^A})$.

From here we used a nested optimization, as in Berry et al. (1993), where in the outer loop, we search over the parameters $\theta$, and in the inner loop, we force the $\tilde{\delta}$'s to be consistent with true market shares.

To perform the SMLE, we need a loan choice dataset with all loan choices, including ones choosing the outside good. To do this, we need to supplement our loan level dataset (inside good), with draws from potential loans choosing the outside good. We do this by supplementing our data with rich, market level, data on mean loan characteristics $Z_i$ from RCA, along with many assumptions. We will assume that, at the market level, each outside observation comes from a jointly normal distribution, and we will randomly sample loans from that joint normal distribution.

To calculate the number of loans to sample in market $m$ ($N_{om}$), we will use RCA data on the total transaction amount in that market. From here, we subtract that total inside transaction amount, which is the total amount of CMBS loans in that market, and divide this by the average transaction amount in that market, which is provided by RCA. We will round this down to make this an integer value and we will then have the number of outside loans to sample.

To get the mean of the normal distribution to sample from ($\mu_m$), we make use RCA's rich data. RCA directly provides mean values for each of our $Z_i$ at the market level[40]. We include the soft information[41] term as one of our loan characteristics. Motivated by the findings of Ghent and Valkanov (2016), we assume that the mean of the soft information term will be the same as the inside good, so we set it zero. We will be assuming that the mean and SD of soft information will be the same across inside and outside loans.

To compute the variance-covariance matrix ($\Sigma_m$) for each market $m$, we do the following. We first construct a vector $\sigma_m$ which contains the SD of each characteristic from the sample of inside goods in market $m$, along with the following adjustments. We make use summary statistics provided in Ghent and Valkanov (2016) for characteristics in their summary statistics table to scale the SD estimates in each market. In that paper, they compare balance sheet loans (which comprise the majority of the outside good) with CMBS loans and report SD's of a few of our characteristics. To account for the fact that the SD's may be different for the outside good relative to the inside good, we scale the SD's with the ratio of the paper's reported SD and our CMBS loan sample SD. For example, Ghent and Valkanov (2016) find the SD of LTV for balance sheet loans to be 0.51, whereas our sample of CMBS has the SD as 0.11, so we scale our market level estimates of SD of LTV by $\frac{0.51}{0.11} = 4.636$.

Next we compute the correlation matrix of our of our characteristics using the entirety of the inside sample used in demand estimation, which we call $\hat{C}$. By doing this, we are assuming that the correlation structure of the characteristics for loans choosing the outside good are exactly the same as the correlation structure for CMBS loans.

To construct a valid covariance matrix, we compute it as follows:

$$\Sigma_m = W_m \widehat{C} W_m$$

where $W_m$ is a diagonal weighting matrix whose elements are the elements of the vector $\sigma_m$. Here we are exploiting a well known result that any covariance matrix can be decomposed as the product of a diagonal weighting matrix, whose diagonal elements are the SD's, along with a correlation matrix, which itself is positive-semi definite, and finally the diagonal weighting matrix again ($\Sigma = WCW$). Therefore this covariance matrix is positive semi-definite and therefore a valid covariance matrix.

Finally, for each market $m$, we randomly sample $N_{O,m}$ loans from $N(\mu_m, \Sigma_m)$ and append these to our loan level dataset. This procedure allows for heterogeneity across markets for the distribution of loan characteristics and number of loans choosing the outside good. All of these distributions are completely driven by both data from the inside goods (CMBS loans) and data from the outside good, along with a few extra moments provided by Ghent and Valkanov (2016). This allows us to identify the parameters $\theta$ in our SMLE procedure.

[40] RCA provides the debt metrics (DSCR, Debt Yield, LTV) at the property type x period level, not property type x period x geography. We assume that all geographies have the same mean for the debt metrics. For apartments and hotels, RCA provides average price per unit instead of price per square foot. We back out price per square foot by using information of units / square foot information from the CMBS loans in each market. Lastly, some of our controls are the logarithm of characteristics provided by RCA. With the assumption that the distribution of the characteristic is log-normal, we can back out the mean and SD of the log of the characteristic with only knowing the mean and SD of the characteristic

[41] The soft information term is the borrower fixed effect estimate from the pricing regression, as in Crawford et al. (2018b)