The Pennsylvania State University

The Graduate School

The Smeal College of Business

# ESSAYS ON THE ROLE OF SOFT DATA AND SPILLOVER EFFECTS IN REAL ESTATE

A Dissertation in
Business Administration
by
Thao Thi Thanh Le



Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2017

The dissertation of Thao Le was reviewed and approved* by the following:

Brent W. Ambrose
Smeal Professor of Risk Management, Graduate Program Chair
Dissertation Advisor
Co-Chair of Committee

Keith Crocker
The William Elliott Chaired Professor of Insurance and Risk Management
Co-Chair of Committee

Austin Jaffe
Philip H. Sieg Professor of Business Administration

Jiro Yoshida
Associate Professor of Business

Liang Peng
Associate Professor of Risk Management

Peter Illiev
Associate Professor of Finance

*Signatures are on file in the Graduate School.

# Abstract

This dissertation comprises three essays on the role of soft data and spillover effects in real estate.

In the first essay, tracking a sample of modified loans underlying private-label mortgage-backed securities, I compare the modification effectiveness of servicers who originated mortgages versus those who simply serviced them. The probability of re-default among loans modified by the former is over 6.9 percentage point lower than the latter. Further tests show that the differences in modification success likely come from the soft information acquired during the origination process. These findings suggest that the loss of soft information in mortgage securitization can impose a substantial cost on mortgage servicing, which raises important policy implications for government regulations in this market.

The second essay examines the effect of peer firm sentiment on firm investment decisions using data from public homebuilders in the U.S. over 2003Q1-2016Q3. Peer sentiment is measured by the NAHB/Wells Fargo Housing Market Index, derived from a monthly survey of homebuilders' perceptions about the conditions of the single-family housing market. I find that a one-standard-deviation increase in the peer sentiment index induces homebuilders to increase their land inventory by 8.4%-12.6%. In addition, big builders are just as prone to peer sentiment as small firms. Consistent with the catering theory, homebuilders held by more short term investors are more likely to follow their peers than those held mainly by institutional shareholders. Interestingly, firms that overbuild compared to their peers have lower stock returns in the next quarter while underbuilding is rewarded with higher stock prices, but this effect decreases as the magnitude of underbuilding increases.

Finally, the third essay investigates the effect of separating real estate from the Financials sector in the Global Industry Classification Standard. Since Sep 1, 2016, real estate became an independent sector instead of being an industry group under the Financials sector together with banks and insurance. Using Real Estate Investment Trusts to represent the new GICS Real Estate sector, I find that their correlation with the Financials sector fell from 0.568-0.775 to 0.338-0.581 after their departure. The reduction in their connection occurred first at announcement and again at implementation. In addition, REIT returns became as much as 60% less volatile than before.

However, becoming a separate sector did not affect trading activities in the REIT market, at least in the short post-implementation period covered in this paper.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgments

This dissertation is not an achievement of my own, but one made possible by all the wonderful people surrounding my life.

I am indebted to my advisor, Dr. Brent Ambrose, for his constant support and guidance along my journey. He is undoubtedly the best advisor anyone could ask for, always welcoming any student who knocks on his door.

I am also very lucky to join a department full of terrific faculty. Dr. Austin is a father figure to all students under his wings. Dr Jiro is amazingly kind, smart and humble. Dr Liang is always attentive and helpful. Dr Keith is such a wise, funny and friendly Department Head. Together they make the best department in the world!

My gratitude also goes to Dr. Joseph Ooi who encouraged me to pursue this journey. He is the reason I am where I am now.

Thanks to all my friends (you know who you are) who have made my life at Penn State a memorable time. I did not know PhD life could be so fun and enjoyable!

Last but not least, I want to thank my family for always having my back. I am thankful for my parents' constant nagging about my weight loss, my sisters' frequent gossip and Facebook updates, and my two little nieces' cute questions about my long absence from home. I dedicate all my accomplishments to them.

# Chapter 1

# Information Loss in Mortgage Securitization: Effect on Loan Modification

## 1.1 Introduction

Following the financial crisis, one of the most prominent concerns of mortgage securitization is the potential conflicts of interest resulting from the separation of a loan's originator and its ultimate owners. This concern arises from the lenders' lack of incentives to carefully screen borrowers in the initial loan underwriting process. Indeed, several studies have argued that the lax lending standards adopted by virtually all creditors during the pre-2007 credit boom should take the primary blame for the ensuing collapse of the subprime mortgage market (see, for example, Dell'Ariccia, Igan, and Laeven, 2008; Keys, Mukherjee, Seru, and Vig, 2010; Kruger, 2014; Mian and Sufi, 2009; Purnanandam, 2011).[1] Demiroglu and James (2012) show that retaining more "skin in the game", through becoming the sponsors or servicers of mortgage-backed securities (MBS), can help alleviate the incentive problem on the part of loan originators. In addition, since originators know the quality of the loans they originate, they can use this private information to cherry-pick better loans to keep in their portfolio while offloading subpar loans to MBS investors (Agarwal, Chang, and Yavas, 2012; Ambrose, Lacour-Little, and Sanders, 2005; Elul, 2016).[2]

---

[1] Dell'Ariccia et al. (2008) and Mian and Sufi (2009) provide empirical evidence that credit became significantly more available during the period leading to the crisis. Such a sharp increase in credit growth was followed by an unprecedented increase in mortgage defaults, as found in Mian and Sufi (2009) and Keys et al. (2010).

[2] Consistent with the adverse selection hypothesis, Elul (2016) finds that securitized loans have a higher default rate than portfolio loans. On the other hand, using data from the 1990s, Ambrose et al. (2005) show that the loans retained in portfolios have a higher default risk, likely due to capital arbitrage or concern for reputation. Interestingly, Agarwal et al. (2012) find that lenders retain loans with higher default risk but lower prepayment risk relative to the loans they sell to the secondary market.

The focus of research thus far is on the asymmetric information problem at the time of loan origination and securitization. This paper extends the existing literature by examining whether the transfer of mortgages to the security market also creates challenges for servicers in managing loan performance. In particular, I study the effectiveness of servicers in performing loan modification when a borrower defaults. Generally, the lender has several options when dealing with defaults (usually defined as missing at least three payments): forebear on the delinquency, modify the loan terms, allow a short sale (whereby borrower can sell the property at a price lower than the remaining loan balance), or foreclose on the property.[3] In the midst of the latest financial crisis, the unprecedented volume of foreclosures prompted pressing calls for greater use of the loan modification option, as it is often believed to be less costly for both borrowers and lenders.[4] However, loan modification is not without cost. The decision to modify a loan depends mostly on the ability of the lender to determine if the borrower can continue to make their payment after receiving assistance. This task requires the lender to carefully collect and evaluate costly information about the financial conditions of the borrower.[5] While it is in the interest and the responsibility of the loan originators to carry out these tasks if the loan is held on their balance sheets, in the case of securitized mortgages, the servicers are in charge of making such decisions on behalf of MBS investors.

Several possible explanations exist for why originators can be more successful than servicers in performing loan modification. Consider the informational advantages of originators as an example. While some information on the borrower's characteristics, such as FICO score, can

---

3 See Ambrose and Capone (1996) for an analysis of the different foreclosure alternatives.

4 Maturana (2014) estimates that an additional modification reduces loan losses by as much as 40% relative to the loan average loss of 40.5%.

5 Eggert (2007) estimates the cost of modification to be between \$500 and \$600 per loan. In addition to expensive monetary expenses, loan modification also brings about moral hazard problem on the part of borrowers (Mayer, Morrison, Piskorski, and Gupta, 2014; Riddiough and Wyatt, 1994; Wang et al., 2002).

be credibly documented and transferred to servicers after securitization, other important soft information, such as the loan officer's assessment of the borrower's personality, that is unverifiable and importable may be lost. Therefore, originators may have informational advantages over any other subsequent institutions involved in handling the mortgages. In other words, information obtained during loan origination can be valuable for servicers in making their decisions to modify troubled loans, but is not available to them. In this paper, I present several empirical tests to explore various hypotheses about what drives the performance of servicers in loan modification.

The empirical strategy in this paper makes use of the fact that in several MBS deals the loan originators retain the servicing rights to the securitized loans (I refer to these as originator-servicers, or OS). Hence, a comparison of the re-default probability of mortgages modified by OS to those modified by non-OS can shed light on this study's research questions.[6] Nonetheless, the challenge in estimating such effect is that retaining servicing rights is an endogenous choice for originators. Similar to the adverse selection problem arising when originators decide to keep safer loans on their books and securitize riskier ones, one can question whether originators cherry-pick mortgages of better quality to retain in their servicing portfolios. In support of this reasoning, Demiroglu and James (2012) compare the ex-post foreclosure rates of securitized mortgages and find a lower rate among those whose originators are affiliated with their servicers. Comparing the observable characteristics of OS and non-OS loans, I confirm that the former appear to have lower observable risk profiles, such as higher FICO scores and lower loan-to-value (LTV) ratios. Furthermore, they have lower default rates on average, even after controlling for observable loan attributes, which implies that originators also make their selection based on unobservable quality.

[6] Some loans may receive modification even before they are officially in default, because the servicers/lenders may have foreseen the borrowers' difficulties; thus, the term "re-default" might not appear appropriate in these cases. Nevertheless, for simplicity, throughout this paper re-default is defined as a default occurrence after a loan has been modified.

To control for this endogeneity problem, I select a sample of comparable loans that include only high-quality non-OS loans and low-quality OS loans, using loan age at securitization as a proxy for unobservable quality. Specifically, the sample includes only non-OS loans that were securitized more than two years after their origination, and OS loans securitized within six months from origination date. The rationale behind this identification strategy is that loans that have performed for at least two years before being securitized are likely high quality loans, while those sold quickly are often risky, conduit-type loans originated solely for the purpose of securitization. Comparing their post-securitization default rate, I confirm that the latter group indeed has higher default risk than the former, justifying my assumption that loan age at securitization can act as a signal of loan quality.

After controlling for their quality difference, OS loans are found to have a considerably lower probability of re-default. More specifically, the probability of re-default within 6 months after receiving modification for OS loans is 8.7 percentage points lower than non-OS loans. The differences are 6.9 and 7.5 percentage points when examining re-defaults within 12 months and 24 months, respectively. In addition, further tests provide insights about possible explanations behind this observation. I find that OS have certain advantages that allow them to better evaluate and select loans for modification. Such advantages likely come from soft information acquired during the origination process, as evidenced in the finding that OS are most effective in modifying loans whose success rate is more difficult to assess, and that their superior performance dissipates as the modification is done further from the origination date. The large difference in the re-default rates of OS and non-OS loans underscores the important advantages an originator has over other institutions in assessing their borrowers. It is particularly essential for practitioners as well as policy makers to recognize that the loss of soft information in mortgage securitization can impose a substantial cost on the effectiveness of servicers in performing such important tasks as loan modification.

## 1.2 Literature review

Although theoretical work on mortgage workouts has existed since Ambrose and Capone (1996), the majority of the empirical research was motivated only recently as a result of the foreclosure crisis in 2008. Securitization plays a central role in the massive development of the mortgage market since 2000. Given the high foreclosure rate relative to modification rate in the recent crisis, many papers have explored the theory that securitization presents a major barrier to loan modification.[7] Agarwal, Amromin, Ben-David, Chomsisengphet, and Evanoff (2011) document that bank-held loans are 26%-70% more likely to be renegotiated than comparable securitized mortgages, after controlling for servicer fixed effects. However, this type of study is often plagued with the selection issue concerning the possibility that originators can choose to sell loans of lower quality to investors. To overcome this problem, Piskorski, Seru, and Vig (2010) make use of the early pay default (EPD) clauses in many Pooling and Servicing Agreements to design a quasi-experiment, and find that the foreclosure rate of delinquent loans held in bank portfolios is 13% to 32% lower than similar securitized loans.[8] However, Adelino, Gerardi, and Willen (2013, 2014) challenge the above claim based on the fact that investors in reality do not

[7] As pointed out in Eggart (2007), there are three possible hurdles for servicers. Firstly, loan modification is labor intensive and time consuming, essentially equivalent to re-underwriting the mortgages. Secondly, the compensation structure of services does not cover these extra modification costs. In particular, the main source of income for servicers is their monthly servicing fee, which is a fixed percentage of the unpaid principal balance of the loans in the pool. When dealing with defaults, although successfully keeping a loan alive might help maintain the servicer's future income, the servicer can recover foreclosure cost but not modification cost. In the recent foreclosure crisis, record default rates caused servicers to favor cost-cutting through automated foreclosure processes rather than risking incurring modification costs with a low likelihood of success. Finally, the conflict of interest between servicers and borrowers as well as investors further worsens the modification disincentive problem. In fact, anecdotal evidence shows that many servicers have engaged in abusive practices to increase their income at the expense of borrowers.

[8] EPD clauses demand issuers to repurchase mortgages that become delinquent within 90 days after securitization. Essentially, whether a borrower first misses a payment in the third or fourth month after issuance is close to random but the former group will ultimately be repurchased and end up on a bank's portfolio. Thus, by restricting their analysis to default loans surrounding the 3-month cutoff, Piskorski et al. (2010) argue that they have a plausible instrument for securitization.

strictly enforce EDP clauses. Adelino et al. (2014) seek to improve their method through a two-stage approach to achieve identification and find no differences in the renegotiation rates for securitized loans and loans held on banks' balance sheets, thereby refuting the prevalent conclusion in other research that securitization creates frictions to loan modification. A similar conclusion is also found in Foote et al. (2009). Another recent attempt to settle this debate by Kruger (2014) uses a quasi-experiment to estimate that securitization increases the probability of foreclosure by 4.7 percentage point and decrease the probability of modification by 3.6 percentage point.

In an earlier paper, Adelino et al. (2013) develop a theoretical model to explain renegotiation activities. Their theory is built upon the information asymmetry problem mentioned in Wang et al. (2002) that borrowers have private information about their financial conditions and willingness to repay the mortgage. In particular, modification rates will be lower when it is more difficult for lenders and servicers to evaluate borrowers' ability and willingness to repay the mortgage. Adelino et al. (2013) find a negative correlation between modification rates and self-cure rates of delinquent loans, which serve as a proxy for the information problem, for the period from 2005 to 2010. Such a finding is in stark contrast with the securitization explanation often cited by others.

Advocates of the securitization hypothesis usually compare the low modification rate in the recent crisis with the prior period when securitization was uncommon. However, there is no concrete evidence about the popularity of loan workouts until Ghent (2011). Using a sample of mortgages originated in New York during the Great Depression when mortgage securitization was almost non-existent (1920-1939), she shows that less than 2% of outstanding loans received any concessionary modification. In other words, loan modification was as rare in the old days as it is during the latest era. Therefore, the debate on the effect of securitization on post-default outcomes remains unresolved in the current literature, notwithstanding several attempts to reconcile the evidence. Although it is tempting to blame securitization for the lack of modification effort by servicers, many studies have proven that there are other elements to the story.

Despite its importance, research on the factors affecting the success rate of loan modification is sparse. Overall, studies in this area have suggested that the probability of re-default depends on the type and timing of modification, the characteristics of the borrowers, and whether the loan is securitized (Acoca and Wachter, 2012; Adelino et al., 2013; Agarwal, Amromin, et al., 2011; Goodman, Ashworth, Landy, and Yang, 2011a, 2011b; Haughwout, Okah, and Tracy, 2010; Quercia and Ding, 2009; Voicu, Jacob, Rengert, and Fang, 2012). Das and Meadows (2013) develop a model conjecturing that, among the various ways to lower monthly payment for borrowers, reducing the principal amount (which effectively reduces LTV ratio) is the optimal type of loan modification. By writing down the LTV ratio, lenders lower both the payment burden and the incentive for strategic default for borrowers.[9] Empirical tests by Quercia and Ding (2009) as well as Haughwout et al. (2010) confirm the above theoretical predictions, but they also acknowledge that their small samples have limited statistical power because principal forbearance is very rare. On the other hand, Agarwal, Amromin, et al. (2011) find that greater reductions in interest rates are associated with lower re-default rates. There are conflicting findings regarding the effect of extending loan duration on re-default probability, with a positive correlation found in Voicu et al. (2012) and a negative link found in Agarwal, Amromin, et al. (2011). Finally, Quercia and Ding (2009) assert that earlier intervention helps reduce the risk of subsequent defaults.

[9] There are two main reasons driving default decisions: negative equity (low willingness to pay) and negative income shock (low ability to pay). If the value of the house falls below the outstanding mortgage balance (negative equity), it is optimal for borrowers to default (strategic default). Reducing LTV ratio will help alleviate the negative equity problem, thereby lowering the incentive to default for borrowers.

Not surprisingly, the success rate of loan modification also depends to a great extent on the characteristics of loans and borrowers. For example, the post-modification performance of high FICO score loans, full document loans, smaller balance loans, loans with positive equity, refinance loans, prime loans, first lien loans, and fixed rate loans is superior to their counterparts (Adelino et al., 2013; Agarwal, Amromin, et al., 2011; Goodman et al., 2011a; Haughwout et al., 2010; Quercia and Ding, 2009). Agarwal et al. (2011) also show that modifications of bank-held loans are more efficient as they have 9% lower re-default rates than mortgages underlying MBS. Interestingly, Zhang (2013) shows that the differences self-cure and re-default rates between the two types of loans converge in the long run.

In an independent, but related paper, Conklin, Diop, and D'Lima (2016) also study the effect of servicer-originator affiliation on loan modification. Our papers differ and complement each other in several aspects. In particular, I focus exclusively and extensively on the performance of loan modification, while Conklin et al. (2016) provide a more comprehensive study on the likelihood of loan modification. Although they observe a lower re-default rate among loans modified by servicers who are related to originators, Conklin et al. (2016) do not address the endogeneity problem concerning the choice of originators to retain safer loans in their servicing portfolios. In addition, Conklin et al. (2016) do not examine the driving forces behind the observed superior performance of affiliated servicers. Lastly, our empirical models, definitions of re-default as well as definitions of originator-servicer affiliation differ. Hence, our papers are natural complements that contribute to the understanding of the performance of MBS servicers in loan modification.

## 1.3 Data

The data used in this study come from Blackbox Logic (BBx), a large database of non-agency mortgages underlying more than 90% of the residential mortgage-backed securities market in the U.S. BBx aggregates data from mortgage servicing companies and standardizes them to ensure consistency across different data providers (see www.bbxlogic.com for more information). It provides detailed information on borrower, property and loan characteristics at origination, as well as the monthly performance of each loan from 2000 to 2013. In this study, I focus on subprime loans whose information about originators and servicers are available in the database. After filtering, the sample has 362,427 loans, and their summary statistics are reported in Panel A of **Table 1.1**.[10] The average mortgage in my sample has similar characteristics as one would expect of a subprime loan in the market, with a FICO score of about 586, principal of $189,944, LTV ratio of 80%, loan term of 359 months, and annual interest rate of 8.66%. The majority of them are ARM loans (77%), and 37% have low or no documentation.

As described earlier, the variable OS is an indicator for a loan originated and serviced by the same financial institution, which can be identified by matching the names of its originator and servicer in the BBx database. Approximately 8% of the sample are OS loans. I identify OS using an exact originator and servicer name match to ensure that they are the same entity. A broader approach to defining the OS variable would consider whether the two institutions are affiliated with each other, such as in a subsidiary-parent company relationship. However, identifying affiliation among servicers and originators is particularly tricky during and after the 2008 financial crisis when the market went through a large wave of bankruptcies, mergers and acquisitions. In

[10] The following loans are excluded from the final sample: original balance less than $40,000, LTV ratio less than 25% or more than 100%, and loan term less than 5 years.

light of such complications, I employ the stricter definition, requiring the two entities to have the same name in order to qualify as an OS. As a result, the estimates of the OS impact are biased away from finding an effect because the non-OS group may include mortgages whose servicers are related to the loan originators. Thus, my estimated coefficient for the OS effect is a conservative approximation of the true effect.

Although the sample covers mortgages originated from 2000, the majority of them concentrate within the short window from 2005 to early 2007, with a peak of 71,205 loans originated in the second quarter of 2006. The top ten originators account for more than 56% of the sample, while the remaining market were shared by 2,084 smaller institutes (Table A1 in the appendix lists the top ten originators and servicers in the sample). New Century is the largest originator in terms of the number of loans issued, followed by Option One and Fremont. In comparison, the sample of servicers is much smaller with only 80 entities, but the market is also dominated by only a few big players. The largest servicer, Ocwen, is responsible for 28% of the sample (see Appendix A).

The focus of this paper, however, is the performance of mortgages that received modification during the study period. The sample of modified loans has 144,148 observations and their summary statistics are shown in Panel B of **Table 1.1.** With more than half of the sample originated and modified during the mortgage crisis (2005-2007), it is not surprising that 60% of the modified loans re-defaulted within 24 months, where default is defined as at least 90 days delinquency. The rates of re-default within 12 months and 6 months are lower at 44% and 23%, respectively, but are still considerably high. Only 3% of this sample are OS loans. There are no discernible differences between the full sample (Panel A) and this sample of modified mortgages in terms of average FICO scores, principal amounts, LTV ratios, loan terms, interest rates, type of interest rate, and level of documentation. At the time of modification, a loan in my sample had slightly less than 320 months remaining in its term. Given that the average loan term in the full sample is 359 months, this implies that the average borrower had made payments for close to 3.5 years when they received loan modification. I also estimate the amount of negative equity for each mortgage at the time of modification using the Federal Housing Finance Agency (FHFA) housing price index.[11] On average, the outstanding balance was equivalent to about 99.35% of the collateral value when the borrower received assistance. Nonetheless, there is a wide range of negative equity among borrowers in the sample, as indicated by the standard deviation of 29.89%.

Four common types of modification are listed in the last four rows of Panel B. Capitalizing arrears is most widely used (83% of the sample), followed by interest rate reduction (79%). In capitalization modification, the delinquent amount is added to the outstanding balance and the borrower is brought to current. This is the most popular method in general, and is often used in combination with other modification types. In comparison, lenders are much more reluctant to reduce principal (41% of the sample), as it requires them to immediately recognize a loss. Only 1% of the loans received term adjustment, because this type of modification is often restricted in order to avoid mismatch in the maturity dates of the underlying mortgages and the trusts that hold them.

## 1.4 Empirical methodology

To test the hypothesis that OS are more effective than non-OS in performing loan modification, I select loans that were modified at least once during the study period and track their

[11] The value of the property at modification is estimated by multiplying its value at origination by the change in the FHFA housing index for the corresponding metropolitan area. Negative equity is then calculated as the LTV ratio at modification minus one.

performance up to two years after modification. The baseline logistic model has the following form:

$$Pr\{Redefault_i\} = \Phi(\alpha + \beta OS_i + \gamma Z_i + \theta S_m + \delta_{sv} + \delta_s + \delta_t + \varepsilon_i). \quad (1)$$

The dependent variable, *Redefault*, takes the value of 1 if loan *i* defaulted within 6 months after modification. For robustness, I also use 12 months and 24 months as alternative cutoffs. This post-modification default rate is used as a measure of the servicers' success in modifying troubled mortgages. The main explanatory variable of interest is *OS*, which is the indicator for a loan modified by an OS. Its coefficient, $\beta$, is expected to be negative if the proposed hypothesis holds. The set of control variables $Z_i$ include loan characteristics measured at the time of origination as well as modification. The variables measured at loan origination include FICO score, principal amount, LTV ratio, loan term, interest rate, interest type (ARM or fixed rate), and documentation type (full or low/no documentation). The model also controls for remaining term and an estimate of the borrower's equity position at the time of modification. In addition, vector $Z_i$ also includes four dummy variables for four types of loan modification: reducing principal, adjusting term, adjusting interest rate, and capitalizing arrears.[12] The set of control variables $S_m$ include two metropolitan-level factors that are important in affecting borrowers' default risk. To account for movement in the housing market, I include the change in the FHFA house price index (HPI) of the metropolitan area (MSA) where the property is located. The change is calculated over the period from the quarter of modification until the time of default. In a similar manner, the change in the MSA-level unemployment rate is also included in vector $S_m$ as a measure of the movement in broad

[12] Note that a loan can receive more than one type of modification.

economic conditions. Finally, the model also controls for servicer fixed effects $\delta_{sv}$, state fixed effects $\delta_s$, and modification month-year fixed effects $\delta_t$.

### 1.4.1 Quality difference between OS and non-OS loans

An important concern in estimating the above model is the selection bias problem associated with the originators' ability to cherry-pick loans to retain in their servicing portfolios. Put it another way, if originators can take advantage of their private information to select low risk loans when bidding for servicing rights, we will observe a lower re-default rate among OS loans that simply results from their lower ex-ante risk. Thus, prior to examining their post-modification performance, I investigate if there exist any systematic differences in the characteristics of OS versus non-OS loans using the following logistic model:

$$\Pr\{OS_i\} = \Phi(\alpha + \beta X_i + \delta_o + \delta_s + \delta_t + \varepsilon_i), \tag{2}$$

where $OS_i$=1 if loan $i$ is an OS loan, and $X$ is a vector of loan characteristics at the time of origination, including FICO score, principal amount, loan term, LTV ratio, interest rate, interest type (fixed rate or ARM), and documentation type (full or low/no documentation). The model includes a set of originator fixed effects $\delta_o$ to control for their heterogeneity, as well as a set of state fixed effects $\delta_s$ and origination time fixed effects $\delta_t$. It is important to note that only mortgage originators who also own a servicing business can choose to service the loans they originate. Originators without servicing businesses, most likely small organizations, do not have such an option and have to sell all servicing rights to an outside institution. By controlling for originator fixed effects, the above regression essentially includes only loans originated by lenders who have servicing businesses, ensuring that the results truly reflect their selection decisions. The first column in **Table 2.1** reports the estimation results of equation (2), which show that OS loans tend to have higher FICO scores, smaller loan amounts, and lower LTV ratios than their counterparts.

Although they also tend to have higher interest rates, as indicated by the positive coefficient on the interest rate variable, the evidence in **Table 1.2** is generally in line with the popular conjecture that originators strategically retain their exposure to low risk loans.

In addition to observable loan and borrower characteristics, there remains a possibility that OS also cherry-pick loans based on private information unobservable to outsiders. Following common practice in the literature, I use a mortgage's default probability measured over 24 months after origination (early defaults) as an indication of its ex-ante quality, and examine whether OS loans indeed have lower early default risk. The intuition goes that good loans are likely to survive at least their first two years, while early defaults are often a signal of risky borrowers. The regression model is specified as follows:

$$\Pr\{Early\_default_i\} = \Phi(\alpha + \beta OS_i + \gamma X_i + \theta S_m + \delta_o + \delta_s + \delta_t + \varepsilon_i), \tag{3}$$

where $X_i$ and $S_m$ are loan- and MSA-specific control variables as described earlier, and $\delta_o, \delta_s$ and $\delta_t$ denote three sets of fixed effects for originator, state and time, respectively. A significantly negative coefficient on OS will provide evidence supporting the postulation that OS loans on average have lower default risks, even after controlling for observable characteristics. This is indeed the case, as shown in the second column of **Table 1.2**. For robustness, I also estimate equation (2) and (3) simultaneously to account for potential endogeneity problem and obtain qualitatively similar results. This finding is congruent with Demiroglu and James (2012) who find that loans serviced by their originators have lower ex-post default rate. Together with the previous result, this test confirms that, on average, originators keep lower risk loans in their servicing portfolios based on both observable and unobservable factors. It is thus important to address this selection bias problem before estimating the OS effect on re-default rate in equation (1).

### 1.4.2 Identification strategy

The difference in observable loan characteristics between the two groups of mortgages can be reduced through a matching estimator, such as propensity score matching. Alternatively, the bias can be controlled for by adding loan characteristics as control variables in regression models. In this paper, I adopt the latter method, due mainly to an econometric issue with the propensity score matching estimator. As described in equation (1) above, the regression model includes a set of loan-specific variables to explicitly control for the observable differences in the two groups of loans.

Clearly, the endogeneity issue associated with unobservable quality is much more problematic. While there are no perfect solutions to correct for this bias, the quality difference between OS and non-OS mortgages can be alleviated if the regression sample includes only the top quality non-OS loans and the lowest quality OS loans. In order to measure their unobserved quality, I use the time a loan remains in the portfolio of its originator until being securitized. In particular, I select only non-OS loans which were securitized more than 24 months after their origination (hereafter, late securitized mortgages). The rationale is that loans which had performed for at least two years before they were securitized are likely high quality loans. On the other hand, we can reasonably raise questions about the quality of mortgages that were securitized within a few months after their origination, for they are likely risky, conduit-type loans that were originated solely for the purpose of securitization. Thus, I restrict the sample to include only OS loans securitized within 6 months after origination (hereafter, early securitized mortgages).

To the best of my knowledge, this paper is the first to use loan age at securitization as a proxy for loan quality. The lack of prior empirical evidence calls for the following test to justify whether there is indeed a positive link between time-to-securitization and loan quality:

$$\Pr\{Default24_i\} = \Phi(\alpha + \beta Age\_at\_securitization_i + \gamma X_i + \theta S_m + \delta_s + \delta_t + \delta_o + \varepsilon_i). \quad (4)$$

The dependent variable in this model is loan $i$'s probability of default within 24 months following its securitization date, which acts as an ex-post measure of the loan's ex-ante, unobservable quality. The independent variables of interest is *Age_at_securitization*, calculated as the number of months between loan *i*'s origination and securitization dates. The vector of control variables $X_i$ includes loan characteristics as described earlier, and LTV ratio at securitization to account for the possibility that loan age can also be a proxy for the amount of equity accumulation in the property up to that point. $S_m$ includes the change in HPI and unemployment rate in the property's MSA. The usual set of state, time, and originator fixed effects are also included.

In the first column of **Table 1.3**, the coefficient on the loan age at securitization variable is negatively significant as predicted by my proposition. When a set of dummy variables is used in the second column in place of the continuous age variable, the statistical significance of the coefficients becomes weaker, although their negative signs still indicate lower default probability for older mortgages, with early securitizers (mortgages securitized within 6 months following their origination) serving as the base case for comparison. Notably, the dummy for loan age between 25 and 36 months is both statistically and economically significant.

Although the above evidence suggests that late securitizers are safer than early securitizers in general, in order for my identification strategy to work, I need to confirm that the selection bias issue is indeed non-existent or insignificant between late non-OS and early OS loans. Thus, I identify a sample of OS loans less than 6 months old and non-OS loans more than 24 months old at the time of securitization (hereafter, the matched sample), and estimate the following regression:

$$\Pr\{Default24_i\} = \Phi(\alpha + \beta OS_i + \gamma X_i + \theta S_m + \delta_s + \delta_t + \delta_o + \varepsilon_i). \quad (5)$$

The estimated coefficients of OS are presented in **Table 1.4**. In the first column, I estimate the above model using the full sample for comparison. Contrast to the negative and significant coefficient obtained using the full sample, the coefficient on OS is positive and insignificant in column (2), indicating that OS loans no longer exhibit a lower ex-post default rate compared to non-OS loans in the matched sample. In summary, the results in **Table 1.3** and **Table** 1.**4** provide evidence to justify the use of loan age at securitization as a proxy for unobserved loan quality, and also support the proposition that the quality difference between the two types of loans can be alleviated using this proxy.

Following the selection criteria described above, the final sample used to estimate equation (1) includes 5,156 modified mortgages. Panel C of **Table 1.1** reports their summary statistics. The matched sample has comparable descriptive statistics to the full sample of modified loans shown in Panel B, with the exception of a few variables. It has a more balanced composition of OS and non-OS loans, with each type comprising about half of the sample. Moreover, the average FICO score, average principal amount, and the proportion of ARM loans are lower than those of the full sample. Regarding negative equity position, an average borrower in this sub-sample had an outstanding loan balance equivalent to about 82% of his property value at the time he received modification, as compared to 99.35% in the full sample of modified loans.

## 1.5 Empirical results

### 1.5.1 Originator-servicers affiliation and loan modification performance

We are now ready to estimate the effectiveness of OS in modifying defaulted mortgages as specified in equation (1). The results are presented in the first three columns of **Table 1.5**; the coefficients of control variables are suppressed for brevity. Our variable of interest *OS* has the expected negative sign. It is both statistically and economically significant across all three horizons used to measure re-default, but its magnitude decreases as the horizon increases to 12 and 24 months. The magnitude of the coefficients suggest that, all else being equal, the probability of re-default of OS loans are 8.7, 6.9 and 7.5 percentage points lower than that of non-OS loans. These figures strongly support the presumption that OS are much more effective in performing modification for the loans they originate compared to their counterparts. For comparison, I report the estimation using the full sample of modified loans in column (4) of **Table 1.5**. Albeit having a negative sign, the coefficient is not statistically significant, contrary to prior expectations. However, it is likely due to the fact that OS loans only comprise less than 3% of the full sample (Panel B of **Table 1.1**). Finally, for robustness check, column (5) of **Table 1.5** shows the estimation result using propensity score matching to control for observable loan characteristics. Even though the standard error reported is incorrectly estimated, the sign and magnitude of the OS coefficient are still consistent with the baseline results obtained in the first three columns of **Table 1.5**.

Even though the tests in Section 4.2 suggest that using 6-month and 24-month as the thresholds to define early and late securitized loans makes the most empirical sense, one may still question how sensitive the above results are to the chosen sample. For robustness, I estimate the baseline model again using different samples created by varying the selection criteria as shown in **Table 1.6**. I focus on the model using the re-default rate within 6 months as the dependent variable, and only the coefficients on *OS* are reported for brevity. In all but the last sample where it becomes statistically weaker at 10% (last row in the last column of **Table 1.6**), the coefficients are qualitatively similar to the earlier estimates. **Table 1.6** thus confirms that the difference in the re-default risks between the two types of loans is persistent across different samples.

### 1.5.2 Explaining the lower re-default rates of OS loans

This section explores several potential explanations for the superior post-modification performance of OS loans reported in **Table 1.5** and **Table 1.6**. Note that using the identification strategy described in the previous section, we can reasonably rule out ex-ante quality as a likely candidate to explain the observed result. The various hypotheses considered in this section belong to two general groups. Those under the first group postulate that OS have stronger incentives than non-OS to allow more substantial modification in hope of avoiding foreclosure, while the second group proposes that OS have certain advantages over non-OS, which allow them to better evaluate and select loans for modification. Put differently, the former set of hypotheses posits that OS exert more effort (the "effort" explanations), while according to the latter OS have better capability in performing modification (the "ability" explanations). The tests in all the following sections use the re-default rate within 6 months after modification as the dependent variable.

#### 1.5.2.1 Effort hypotheses

*a. Foreclosure cost*

The success of loan modification depends to a great extent on the degree of payment reduction given by lenders. Intuitively, the more concession a borrower receives, the more likely he can continue to make payment, implying that the low re-default rate observed among modified OS loans can simply be a mechanical result of OS giving more substantial payment reduction. This explanation is highly conceivable if OS concentrate their lending in states where it is more costly to foreclose on properties, hence stronger incentives to save defaulted loans. Foreclosure laws differ substantially across U.S. states, which have been both theoretically predicted and empirically tested to affect default behaviors as well as lenders' loss severity (Ambrose, Buttimer, and Capone, 1997; Crawford and Rosenblatt, 1995; Ghent and Kudlyak, 2011). A summary provided in Ghent and Kudlyak (2011) shows that many states forbid deficiency judgments (non-recourse states), such as California or Minnesota, or require a lengthy judicial foreclosure process spanning more than 360 days, such as New York and Michigan, as opposed to just 46 days in Maryland.[13] It follows that in borrower-friendly jurisdictions, lenders or servicers are more inclined to seek alternative loan workouts rather than foreclosure. However, I reject the conjecture that OS have stronger incentives to avoid foreclosure for two reasons. Firstly, the regressions in **Table 1.5** include state fixed effects, which effectively allows us compare loans within the same state. Secondly, to visually check whether OS loans are indeed disproportionally located in states with high foreclosure costs, I plot the number of OS and non-OS loans in four types of states in **Figure 1.1**. Following Ghent and Kudlyak (2011), I categorize all states into four types based on whether they have a judicial foreclosure process and allow deficiency judgements. There is no discernable concentration of OS loans in any type of states compared to non-OS loans.

*b. Reputational concern*

Next, reputation can be another potential source of incentives for OS to exert more effort in modifying loans. In their role as servicers, OS may also be wary about their own reputation as originators when handling default cases. For example, if too many foreclosures can damage its name in the mortgage origination business, a servicer may be willing to extend more substantial concessions in order to help borrowers continue to make payment. Using an institution's market share as a proxy for its reputation, I examine if the OS effect changes with reputational costs in **Table 1.7**. *Top 10 originators* and *Top 10 originator-servicers* are two indicators for ten

[13] A deficiency judgment allows a lender to pursue the borrower's personal property if the foreclosure sale is not sufficient to cover the mortgage balance due.

institutions with the highest market share as originators and originator-servicers, respectively.[14] Arguably, if the low re-default rate of OS loans is driven by reputational concerns, we should observe even stronger negative coefficients for the interaction terms in **Table 1.7**, but there is no evidence to support this conjecture. Both interaction terms have positive coefficients, contrary to the prediction by this reputation hypothesis. Note that since interaction terms are difficult to interpret in a logistic model, in this paper I estimate and present results for all regressions with interaction terms using the Ordinary Least Square estimator. Though not reported, for robustness checks I also estimate them using logistic regressions and the results are qualitatively similar.

*c. Type and extent of modification*

More generally, regardless of their motivation, the primary question asked in this section is whether an affiliation with the loan originators causes the servicers to act differently. Thus, I examine the type and extent of modification given by OS and non-OS to look for direct evidence on their willingness to revive the troubled loans. The four types of modification have very different implications for borrowers and mortgage owners. Arrear capitalization is the most commonly used but least helpful for borrowers because it does not reduce their payment burden. Among the remaining three options, several theoretical and empirical papers have shown that reducing principal amount is the optimal type of modification, owing to its dual relief effect on payment burden and negative equity for the borrowers (Das and Meadow, 2013; Quercia and Ding, 2009; Goodman et al., 2011a, 2011b; Haughwout et al., 2010). Ambrose and Buttimer (2012) even propose that mortgage contracts should allow for automatic principal adjustment in order to minimize default risk. However, this is the most costly option for the mortgage owners because they have to immediately recognize the loss. On the other hand, Agarwal, Amromin, et al., (2011)

[14] The market share of each institution is calculated using the full sample of mortgages described in Panel A of Table 1.

find that greater reductions in interest rates are associated with lower re-default rates. There are conflicting findings regarding the effect of extending loan duration on re-default probability, with a positive correlation found in Voicu et al. (2012) and a negative link found in Argawal et al. (2011).

The purpose of the test presented in **Table 1.8**, however, is not to resolve the debate on which modification method is more effective, but rather to study how effectively OS employ them. The coefficients of interests are the interaction terms between the *OS* dummy and the three type of modification indicators. Note that only 5% of the sample received term adjustment (Panel C of **Table 1.1**), all of which are non-OS loans, hence the missing interaction term between *OS* and *Adjust term*. Interestingly, the only type of modification that OS appear to have used more effectively is arrear capitalization, but it is the only method that offers no help for borrowers in terms of easing their financial burden.

Finally, an examination of the extent of payment reduction given by OS and non-OS can offer direct evidence on the "effort" conjecture proposed in this section. Since the BBx database only provide information on the type but not the amount of modification, I use the following approach to estimate the payment change for each loan after it was modified. Denote the month a borrower received assistance as time *t*. Using the data on the outstanding loan amount, interest rate and remaining term at the end of time *t-1,* I recursively calculate the scheduled payments for time *t*, *t+1* and *t+2*. These are the payments that the borrower would have had to pay for these three months, had they not received any modification at time *t*. The next step involves subtracting the calculated payment for period *t+2* from the actual payment for that month reported by BBx.[15]

[15] I do not use the payments at time *t+1* to allow for any possible delays in implementation by the servicers or data reporting by BBx.

This difference is the payment change resulting from the modification. As shown in the first row of **Table 1.9**, the average borrower in my sample received a reduction of about $214.27, which is equivalent to 15.63% of the payment they were originally required to pay. The next two rows report the statistics for OS and non-OS loans separately. In terms of dollar amount, the t-test on the difference in the means of the two groups ($208.4 versus $221.01) is not statistically significant. However, when the change is expressed as a percentage of the borrower's originally required payment, non-OS loans seem to receive a higher concession of 16.98% compared to only 14.46% for OS loans, and the difference is statistically significant. Interestingly, despite having smaller payment reduction, OS loans are found to perform better post-modification. Alongside all the previous evidence presented in this section, this finding essentially implies that the lower re-default rate of OS loans is not attributable to the type and extent of modification the borrowers received. A natural question then follows, is it the result of OS being better at evaluating and selecting borrowers to offer help?

#### 1.5.2.2 Ability hypotheses

*a. Organizational advantages*

We now turn to the second strand of hypotheses regarding servicers' ability to perform effective modification. First of all, OS might possess unique characteristics that make them inherently better servicers than their counterparts. For example, since all OS must have both loan originating and servicing businesses, they are often big and more diversified institutions, which might imply better servicing capability. In this sense, the OS variable is simply a proxy for size, and all loans serviced by these institutions will have lower re-default probability compared to those serviced by smaller institutions. However, since the baseline model in equation (1) already controls for servicer fixed effects, it effectively compares the performance of loans modified by the same servicer. We can therefore conclude from the findings in **Table 1.5** that a servicer is indeed more effective in modifying the loans it originates than those it does not.

*b. Soft information*

The above observation leads to the next and final hypothesis that servicers involved in the origination process gain valuable soft information to assist with their subsequent modification decisions. Research on the role of "soft" information has gain increased prominence since the 1990s as the rapid advancement of information technology drastically reduces the need for face-to-face interactions. While there is yet any formal definition of soft information in the literature, it is often associated with being qualitative and personal in nature, as opposed to hard information that can be objectively measured and quantified (Berger and Udell, 2002; Petersen, 2004). Soft information can only be obtained through close, and often repeated, interactions between two or more parties. The consensus in the literature is that such information is beneficial and important for both lenders and borrowers. As a supplement to hard information, soft information provides lenders with a more precise assessment of borrowers' risk characteristics, therefore improving their credit decision making (Diamond, 1984; Rajan, 1992).[16] Additionally, through building a personal relationship with lenders, not only can borrowers increase their access to credit but also negotiate more favorable terms, especially borrowers whose hard information is difficult to verify (eg., Agarwal and Hauswald, 2010; Berger and Udell, 1995; Petersen and Rajan, 1994). At the retail lending level, Agarwal, Ambrose, Chomsisengphet, and Liu (2011) use data on home equity loans to show that financial institutions do rely on soft information to assess borrowers' risk and

[16] However, newer evidence from Petersen and Rajan (2002) indicates that banks were relying increasingly on hard information and impersonal interactions in their lending to small firms from 1973 to 1993. This is attributed to increased labor productivity brought about by advancement in communication technology.

determine the appropriate annual percentage rate requirement for them. Conklin (2015) also finds that face-to-face interaction between a mortgage broker and borrower in the loan origination process can help educate borrowers and thus reduce problems associated with financial illiteracy.

Success in loan modification depends largely on the servicers' ability to determine the likelihood that borrowers can continue making payments after their contracts are modified. Essential to this decision is the evaluation of the borrower's characteristics and financial conditions, which requires collecting and assessing information, especially soft information that is often not reflected in their loan documents. Since this is a costly and time-consuming process, the originator of a loan who has already gone through the initial underwriting process will have an advantage over other institutions performing this task. Such information friction can present a major challenge to loan modification for securitized mortgages when servicers are not the loan originators.

Since soft information is unobservable to outsiders, testing this hypothesis directly is challenging. The following indirect tests are based on its several testable implications. Firstly, an observation that OS are more successful in modifying loans that are informationally opaque, such as low-doc loans, will be consistent with the information hypothesis. However, the estimation of model (1) in **Table 1.10** does not support this proposition, as shown by the interaction term between *OS* and *Lowdoc*. Secondly, given that they have lower information cost and better information quality to aid with their decision making, OS are expected to be more effective in modifying loans whose success probability is more difficult to assess. In particular, it is much riskier to modify mortgages with negative equity because these borrowers have high propensity to give up their properties. Sorting borrowers in my sample by negative equity position, I find that about 24% of my sample owed more than their house value when they received modification, with the top 10% of borrowers having outstanding loan balance equivalent to more than 117% of their house value. The dummy indicator for borrowers with negative equity, *Negative equity dummy*, is interacted with the *OS* indicator in the second column in **Table 10**. Consistent with the prediction by the information hypothesis, the interaction term is negative and significant at the 5% level, suggesting that OS are especially better at dealing with borrowers with negative equity.

Finally, this hypothesis also implies that useful information about the borrowers obtained in the origination process, if any, should become less relevant as the modification is done further from the origination date. I test this conjecture using a set of dummy variables indicating the loan age at modification in **Table 1.11**. As predicted, the coefficients of the interaction terms between OS and these dummy variables are positive and monotonically increasing as loan age increases. Interestingly, if the loan is modified more than 48 months after its origination date, there is almost no difference in the performance of OS versus non-OS. In other words, the positive effect of involvement in loan origination dissipates over time and does not extend beyond 4 years.

To summarize the findings in this section, I find that OS do not possess institutional advantages that lead to their higher success rate in loan modification. Instead, it is likely that they benefit from the original loan underwriting process and are able to use these informational advantages effectively in making later modification decisions.

## 1.6 Conclusion

The detachment between originators and the loans they underwrite in mortgage securitization has attracted much attention from researchers, practitioners and policy makers alike. The focus of virtually all discussion thus far is on the lenders' disincentives to carefully screen borrowers in the initial loan origination process, and the asymmetric information problem when these loans are sold to investors. This paper adds to the existing literature by examining whether and how originators also have important advantages over other parties in managing their loans in the long run. More precisely, I compare the success rate in loan modification among MBS servicers when they also act as originators versus when they do not, after controlling for any pre-existing difference in loan quality.

Using a sample of mortgages underlying residential MBS in the U.S., I find that servicers are much more effective in modifying the mortgages originated by themselves. The probability of re-default within 6 months, 12 months and 24 months after receiving modification for OS loans is 8.7, 6.9, and 7.5 percentage points lower than non-OS loans, respectively. More importantly, further evidence shows that there are no discernible differences in the type and extent of modification given by OS in comparison to their counterparts. Indeed, it is their ability to assess and select the right loans to modify that drives the result, which likely comes from the soft information collected during the initial underwriting process. This is evidenced in the finding that OS are especially effective in working with mortgages that are often more challenging to evaluate, in particular underwater loans. In addition, the effect tapers off as the modification occurs further from the origination dates and disappears after about four years, consistent with old information being less valuable. However, these observations are limited in that they only offer indirect support for the soft information story. Further research is needed to shed light on this hypothesis. Nevertheless, the findings in this paper raise important questions for practitioners and policy makers concerning about the effectiveness of MBS servicers, as well as servicers of securitized loans in general, in performing important tasks such as loan modification. The cost associated with mortgage securitization is obviously not limited to moral hazard and adverse selection problems at the beginning of the process, but extends throughout the life of the underlying mortgages.

### 1.7 Bibliography


Acoca, A., and Wachter, S. (2012). The Performance of New Private-Label Mortgage Loan Modifications. *Working Paper*.

Adelino, M., Gerardi, K., and Willen, P. (2014). Identifying the Effect of Securitization on Foreclosure and Modification Rates Using Early Payment Defaults. *Journal of Real Estate Finance and Economics*, *49*, 352–378.

Adelino, M., Gerardi, K., and Willen, P. S. (2013). Why don't Lenders renegotiate more home mortgages? Redefaults, self-cures and securitization. *Journal of Monetary Economics*, *60*(7), 835–853.

Agarwal, S., Ambrose, B. W., Chomsisengphet, S., and Liu, C. (2011). The role of soft information in a dynamic contract setting: Evidence from the home equity credit market. *Journal of Money, Credit and Banking*, *43*(4), 633–655.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., and Evanoff, D. D. (2011). The role of securitization in mortgage renegotiation. *Journal of Financial Economics*, *102*(3), 559–578.

Agarwal, S., Chang, Y., and Yavas, A. (2012). Adverse selection in mortgage securitization $. *Journal of Financial Economics*, *105*(3), 640–660. http://doi.org/10.1016/j.jfineco.2012.05.004

Agarwal, S., and Hauswald, R. (2010). Distance and private information in lending. *Review of Financial Studies*, *23*(7), 2758–2788.

Ambrose, B. W., and Buttimer, R. J. (2012). The Adjustable Balance Mortgage: Reducing the Value of the Put. *Real Estate Economics*, *40*(3), 536–565.

Ambrose, B. W., Buttimer, R. J., and Capone, C. a. (1997). Pricing Mortgage Default and Foreclosure Delay. *Journal of Money, Credit, and Banking*, *29*(3), 314–325.

Ambrose, B. W., and Capone, C. a. (1996). Cost-benefit analysis of single-family foreclosure alternatives. *The Journal of Real Estate Finance and Economics*, *13*(2), 105–120.

Ambrose, B. W., Lacour-Little, M., and Sanders, A. B. (2005). Does Regulatory Capital Arbitrage , Reputation , or Asymmetric Information Drive Securitization ? *Journal of Financial Services Research*, *28*, 113–133.

Berger, A. N., and Udell, G. F. (1995). Relationship Lending and Lines of Credit in Small Firm Finance *, *68*(3), 351–381.

Berger, A. N., and Udell, G. F. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *Economic Journal*, *112*(477), 32–53.

Conklin, J. N. (2015). Financial Literacy , Broker Borrower Interaction , and Mortgage Default, *Forthcomin*.

Conklin, J. N., Diop, M., and D'Lima, W. (2016). Network Affiliation and Mortgage Modification in the Non-Agency MBS Market. *Working Paper*.

Crawford, W. G., and Rosenblatt, E. (1995). Efficient Mortgage Default Option Exercise: Evidence from Loss Severity. *Journal of Real Estate Research*, *10*(5), 543–555.

Das, S. R., and Meadows, R. (2013). Strategic loan modification: An options-based response to strategic default. *Journal of Banking and Finance*, *37*(2), 636–647.

Dell'Ariccia, G., Igan, D., and Laeven, L. (2008). "Credit Booms and Lending Standards : Evidance from the Subprime Mortgage Market." *Journal of Money, Credit and Banking, 44*(2), 367–384.

Demiroglu, C., and James, C. (2012). How important is having skin in the game? Originator-sponsor affiliation and losses on mortgage-backed securities. *Review of Financial Studies, 25*(11), 3217–3258.

Diamond, D. W. (1984). Financial Intermedation and Delegated Monitoring. *Review of Economic Studies*, *51*(3), 393–414.

Eggert, K. (2007). Comment on Michael A . Stegman et al .' s " Preventive Servicing Is Good for Business and Affordable Homeownership Policy ": What Prevents Loan Modifications ? *Housing Policy Debate*, *18*(2), 279–298.

Elul, R. (2016). Securitization and Mortgage Default. *Journal of Financial Services Research, 49*, 281–309. http://doi.org/10.1007/s10693-015-0220-3

Foote, C., Gerardi, K., Goette, L., Willen, P., Foote, C., Gerardi, K., … Willen, P. (2009). FEDERAL RESERVE BANK o f ATLANTA Reducing Foreclosures : No Easy Answers Reducing Foreclosures : No Easy Answers. *Working Paper*.

Ghent, A. C. (2011). Securitization and mortgage renegotiation: Evidence from the great depression. *Review of Financial Studies*, *24*(6), 1814–1847.

Ghent, A. C., and Kudlyak, M. (2011). Recourse and residential mortgage default: Evidence from US states. *Review of Financial Studies*, *24*(9), 3139–3186.

Goodman, L., Ashworth, R., Landy, B., and Yang, L. (2011a). Modification Success— What Have We Learned? *The Journal of Fixed Income*, *21*(2), 57–67.

Goodman, L., Ashworth, R., Landy, B., and Yang, L. (2011b). The Case for Principal Reduction. *Journal of Structured Finance*, *Fall*, 29–41.

Haughwout, A., Okah, E., and Tracy, J. (2010). Federal Reserve Bank of New York Staff Reports Second Chances : Subprime Mortgage Modification and Re-Default. *Working Paper, Federal Reserve Bank of New York*.

Keys, B. J., Mukherjee, T., Seru, A., and Vig, V. (2010). Did Securitization Lead to Lax Screening? Evidence from Subprime Loans. *Quarterly Journal of Economics*, *125*(1), 307–362.

Kruger, S. (2014). The Effect of Mortgage Securitization on Foreclosure and Modification. *Working Paper*. Retrieved from http://scholar.harvard.edu/files/skruger/files/securitization_and_foreclosure.pdf

Maturana, G. (2014). When are Modifications of Securitized Loans Beneficial to Investors? *Working Paper*. http://doi.org/10.2139/ssrn.2524727

Mayer, C., Morrison, E., Piskorski, T., and Gupta, A. (2014). Mortgage Modification and Strategic Default : Evidence from a Legal Settlement with Countrywide. *American Economic Review*, *104*(9), 2803–2857.

Mian, A., and Sufi, A. (2009). The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *Quarterly Journal of Economics*, *124*(4), 1449–1497.

Petersen, M. A. (2004). Information : Hard and Soft. *Working Paper*.

Petersen, M. A., and Rajan, R. G. (1994). The Benefits of Lending Relationships : Evidence from Small Business Data. *Journal of Finance*, *49*(1), 3–37.

Petersen, M. A., and Rajan, R. G. (2002). Does distance still matter? The information revolution in small business lending. *Journal of Finance*, *57*(6), 2533–2570.

Piskorski, T., Seru, A., and Vig, V. (2010). Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis. *Journal of Financial Economics*, *97*(3), 369–397.

Purnanandam, A. (2011). Originate-to-distribute model and the subprime mortgage crisis. *Review of Financial Studies*, *24*(6), 1881–1915.

Quercia, R. G., and Ding, L. (2009). Loan Modifications and Redefault Risk : An Examination of Short-Term Impacts. *Cityscape: A Journal of Policy Development and Research*, *11*(3), 171–194.

Rajan, R. G. (1992). Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt. *The Journal of Finance*, *47*(4), 1367–1400.

Riddiough, T. J., and Wyatt, S. B. (1994). Strategic default, workout, and commercial mortgage valuation. *The Journal of Real Estate Finance and Economics*, *9*(1), 5–22.

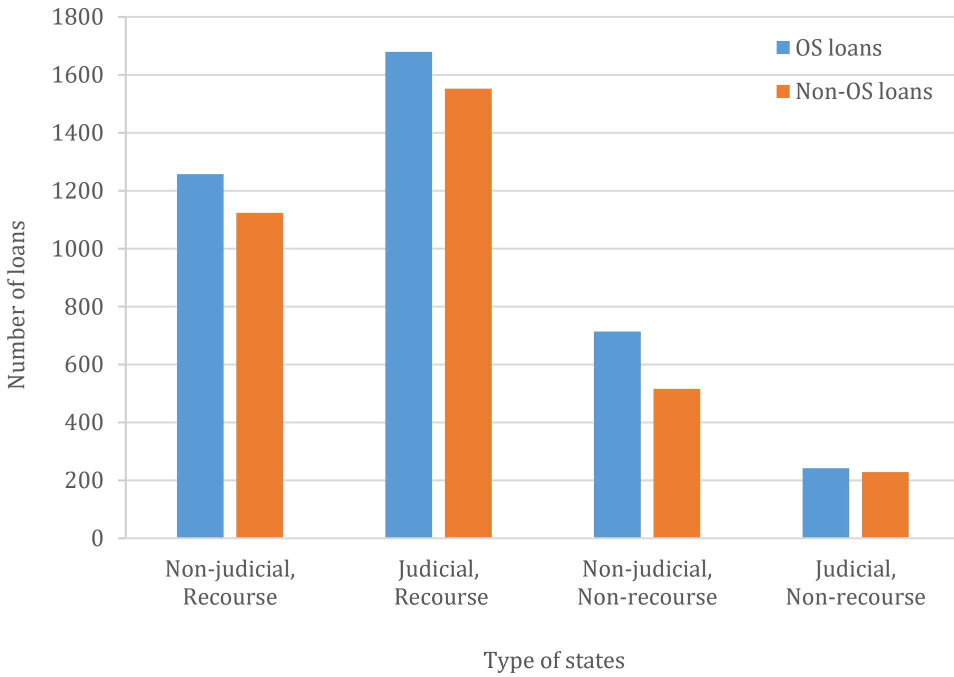
Voicu, I., Jacob, M., Rengert, K., and Fang, I. (2012). Subprime Loan Default Resolutions: Do They Vary Across Mortgage Products and Borrower Demographic Groups? *Journal of Real Estate Finance and Economics*, *45*(4), 939–964.

Wang, K., Young, L., and Zhou, Y. (2002). Nondiscriminating Foreclosure and Voluntary Liquidating Costs. *Review of Financial Studies*, *15*(3), 959–985.

Zhang, Y. (2013). Does loan renegotiation differ by securitization status? A transition probability study. *Journal of Financial Intermediation*, *22*(3), 513–527. http://doi.org/10.1016/j.jfi.2012.11.003

## Figure 1.1. Distribution of loans by State

This figure plots the number of OS and non-OS loans in four type of states as categorized in Ghent and Kudlyak (2011). *Non-judicial, Recourse* states include AL, CO, DC, GA, HI, ID, MI, MO, MS, NE, NV, NH, OK, RI, TN, TX, UT, VA, WV, and WY. *Judicial, Recourse* states include CT, DE, FL, IL, IN, KS, KY, LA, ME, MD, MA, NJ, NM, NY, OH, PA, SC, SD, and VT. *Non-judicial, Non-recourse* states include AK, AZ, CA, MN, MT, NC, OR, and WA. *Judicial, Non-recourse* states include AR, IA, NC, ND and WI.



## Table 1.1. Summary statistics

This table reports the summary statistics of all loans in the sample. *OS* is an indicator for a loan originated and serviced by the same institution; *FICO* is the credit score developed by FICO; *Principal* is the original loan amount in USD; *LTV* is the loan-to-value ratio; *Term* is the loan term in months; *Interest rate* is the annual interest rate charged on the loan; *ARM* is dummy for adjustable rate mortgages; *Lowdoc* is a dummy for loans with low or no documentation; *Re-default* is a dummy for a loan that defaults within 24, 12 or 6 months after it is modified, where default is defined as at least 90 days delinquency; *Investment property* is a dummy for a house bought for investment purposes; *Remaining term* is the number of months remaining at the time of modification; *Negative equity* is calculated as the LTV ratio at modification minus one; *Reduce principal* is an indicator for a loan receiving a reduction in outstanding principal; *Adjust rate* is an indicator for a loan receiving an adjustment in interest rate, *Adjust term* is an indicator for a loan receiving an adjustment in loan term; *Capitalize arrears* is an indicator for a loan receiving arrear capitalization.

| **Variable** | **Mean** | **Std. Dev.** | **Min** | **Max** |
|---|---|---|---|---|
| ***Panel A: Full sample (n=362,427)*** | | | | |
| OS | 0.08 | 0.28 | 0 | 1 |
| FICO | 586.33 | 44.55 | 500 | 730 |
| Principal (USD) | 189,944 | 124,373 | 40,000 | 647,500 |
| LTV (%) | 80.49 | 11.92 | 25 | 100 |
| Term (months) | 359.27 | 34.01 | 180 | 480 |
| Interest rate (%) | 8.66 | 2.53 | 2 | 16.37 |
| ARM | 0.77 | 0.42 | 0 | 1 |
| Lowdoc | 0.37 | 0.48 | 0 | 1 |
| ***Panel B: Sample of modified loans (n=144,148)*** | | | | |
| Re-default (24 months) | 0.60 | 0.49 | 0 | 1 |
| Re-default (12 months) | 0.44 | 0.50 | 0 | 1 |
| Re-default (6 months) | 0.23 | 0.42 | 0 | 1 |
| OS | 0.03 | 0.16 | 0 | 1 |
| *Loan characteristics at origination:* | | | | |
| FICO | 573.99 | 36.39 | 500 | 730 |
| Principal (USD) | 204,019 | 130,932 | 40,000 | 647,500 |
| LTV (%) | 80.66 | 11.29 | 25 | 100 |
| Term (months) | 368.40 | 37.72 | 180 | 480 |
| Interest rate (%) | 8.89 | 4.15 | 2 | 16.35 |
| ARM | 0.78 | 0.42 | 0 | 1 |
| Lowdoc | 0.36 | 0.48 | 0 | 1 |
| *Loan characteristics at modification:* | | | | |
| Remaining term (months) | 319.59 | 41.89 | 132 | 480 |
| Negative equity (%) | -0.65 | 29.89 | -90.93 | 89.45 |
| *Type of modification:* | | | | |
| Reduce principal | 0.41 | 0.49 | 0 | 1 |
| Adjust rate | 0.79 | 0.41 | 0 | 1 |
| Adjust term | 0.01 | 0.11 | 0 | 1 |
| Capitalize arrears | 0.83 | 0.38 | 0 | 1 |
| ***Panel C: Sample of modified loans – Early securitized OS loans and late securitized non-OS loans (n=5,156)*** | | | | |
| Re-default (24 months) | 0.58 | 0.49 | 0 | 1 |
| Re-default (12 months) | 0.43 | 0.49 | 0 | 1 |
| Re-default (6 months) | 0.24 | 0.43 | 0 | 1 |
| OS | 0.54 | 0.50 | 0 | 1 |

| | | | | |
|---|---|---|---|---|
| *Loan characteristics at origination:* | | | | |
| FICO | 558.59 | 40.81 | 500 | 730 |
| Principal (USD) | 158,127 | 111,972 | 40,000 | 647,500 |
| LTV (%) | 81.61 | 12.61 | 25 | 100 |
| Term (months) | 372.23 | 57.81 | 180 | 480 |
| Interest rate (%) | 7.92 | 3.58 | 2 | 16.35 |
| ARM | 0.58 | 0.49 | 0 | 1 |
| Lowdoc | 0.32 | 0.46 | 0 | 1 |
| *Loan characteristics at modification:* | | | | |
| Remaining term (months) | 300.71 | 65.66 | 132 | 480 |
| Negative equity (%) | -17.95 | 29.94 | -90.93 | 89.45 |
| *Type of modification:* | | | | |
| Reduce principal | 0.35 | 0.48 | 0 | 1 |
| Adjust rate | 0.66 | 0.47 | 0 | 1 |
| Adjust term | 0.05 | 0.22 | 0 | 1 |
| Capitalize arrears | 0.86 | 0.35 | 0 | 1 |

## Table 1.2. Characteristics of OS loans

This table reports the estimations of equation (2) and (3). *OS* is an indicator for a loan originated and serviced by the same institution; *Early default* is a dummy for a loan that defaults within 24 months after it is originated, where default is defined as at least 90 days delinquency; *FICO* is the credit score developed by FICO; *Principal* is the original loan amount (in log); *LTV* is the loan-to-value ratio; *Term* is the loan term in months; *Interest rate* is the annual interest rate charged on the loan; *ARM* is dummy for adjustable rate mortgages; *Lowdoc* is a dummy for loans with low or no documentation; *Change in HPI* is the change in the FHFA house price index of the MSA where the property is located, calculated from origination to the time the loan defaults; *Change in unemployment rate* is the change in the unemployment rate of the MSA where the property is located, calculated from origination to the time the loan defaults.

| VARIABLES | OS | Early default |
|---|---|---|
| | (1) | (2) |
| OS | | -0.392*** |
| | | (0.081) |
| FICO | 0.002*** | 0.001 |
| | (0.000) | (0.000) |
| Principal (log) | -0.118*** | 0.302*** |
| | (0.036) | (0.056) |
| LTV | -0.013*** | 0.022*** |
| | (0.001) | (0.004) |
| Term | -0.001 | 0.000 |
| | (0.000) | (0.001) |
| Interest rate | 0.037*** | 0.072*** |
| | (0.008) | (0.004) |
| ARM | -0.069* | 0.386*** |
| | (0.039) | (0.052) |
| Lowdoc | -0.014 | 0.322*** |
| | (0.036) | (0.019) |
| Change in HPI (MSA) | | -0.029* |
| | | (0.016) |
| Change in unemployment rate (MSA) | | -0.035*** |
| | | (0.003) |
| Constant | 5.447*** | -9.966*** |
| | (0.603) | (1.607) |
| Originator FE | Yes | Yes |
| State FE | Yes | Yes |
| Time FE | Yes | Yes |
| Observations | 209,960 | 359,227 |
| Pseudo R | 0.462 | 0.183 |

## Table 1.3. Testing the time-to-securitization assumption

This table reports the estimation of equation (4). *Age at securitization* is calculated as the number of months between loan origination and securitization dates. Other control variables include *FICO, Principal (log), LTV, Term, Interest rate, ARM, Lowdoc, LTV at securitization*, *Change in HPI (MSA),* and *Change in unemployment rate (MSA).*

| | Default within 24 months after securitization | |
|---|---|---|
| VARIABLES | (1) | (2) |
| Age at securitization | -0.035** | |
| | (0.016) | |
| Dummies for age at securitization: | | |
| 6-24 months dummy | | -0.039 |
| | | (0.047) |
| 25-36 months dummy | | -0.578** |
| | | (0.295) |
| 37-48 months dummy | | -0.361 |
| | | (0.497) |
| 48-60 months dummy | | -0.504 |
| | | (0.678) |
| > 60 months dummy | | -1.195 |
| | | (0.997) |
| Other controls | Yes | Yes |
| State FE | Yes | Yes |
| Time FE | Yes | Yes |
| Originator FE | Yes | Yes |
| Observations | 99,420 | 99,420 |
| Pseudo R | 0.563 | 0.572 |

## Table 1.4. Testing the endogeneity issue

This table reports the estimation of equation (5). *OS* is an indicator for a loan originated and serviced by the same institution. Full sample include all loans that were modified. Match sampled includes modified OS loans which were securitized less than 6 months after securitization, and non-OS loans which were securitized more than 24 months after securitization. Other control variables include *FICO, Principal (log), LTV, Term, Interest rate, ARM, Lowdoc, LTV at securitization*, *Change in HPI (MSA)*, and *Change in unemployment rate (MSA).*

| | Default within 24 months after securitization | |
|---|---|---|
| | Full sample | Matched sample |
| VARIABLES | (1) | (2) |
| OS | -1.132*** | 0.735 |
| | (0.143) | (1.110) |
| Other controls | Yes | Yes |
| State FE | Yes | Yes |
| Time FE | Yes | Yes |
| Originator FE | Yes | Yes |
| Observations | 98,821 | 10,317 |
| Pseudo R | 0.586 | 0.611 |

## Table 1.5. Re-default probability of modified loans

This table reports the estimation of equation (1). *Re-default* is a dummy for a loan that defaults within 24, 12 or 6 months after it is modified, where default is defined as at least 90 days delinquency; *OS* is an indicator for a loan originated and serviced by the same institution. Loan characteristics include *FICO*, *Principal (log)*, *LTV*, *Term*, *Interest rate*, *ARM*, and *Lowdoc.* Other control variables include *Remaining term*, *Negative equity*, *Capitalize arrears*, *Reduce principal*, *Adjust term*, *Adjust rate*, *Change in HPI (MSA)*, and *Change in unemployment rate (MSA).*

| | Re-default within | | | | |
|---|---|---|---|---|---|
| | 6 months | 12 months | 24 months | 6 months | 6 months |
| | Matched sample | Matched sample | Matched sample | Full sample | Propensity score matching |
| VARIABLES | (1) | (2) | (3) | (4) | (5) |
| OS | -0.580*** | -0.342** | -0.370* | -0.090 | -0.726** |
| | (0.216) | (0.179) | (0.197) | (0.078) | (0.302) |
| Loan characteristics | Yes | Yes | Yes | Yes | No |
| Other controls | Yes | Yes | Yes | Yes | Yes |
| State FE | Yes | Yes | Yes | Yes | Yes |
| Time FE | Yes | Yes | Yes | Yes | Yes |
| Servicer FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 4,398 | 4,422 | 4,423 | 144,045 | 4,366 |
| Pseudo R | 0.158 | 0.134 | 0.139 | 0.130 | 0.139 |

## Table 1.6. Testing the baseline model with different samples of early and late securitized loans

This table reports the coefficients on the *OS* variable obtained from estimating equation (1) for various loan samples. The dependent variable is a loan's re-default rate within 6 months following its modification date. Standards errors are in brackets.

| Time between origination and securitization | | Early securitized OS loans | | |
|---|---|---|---|---|
| | | <= 6 months | <=12months | <=24 months |
| Late securitized non-OS loans | > 12 months | -0.460*** | -0.492*** | - |
| | | (0.153) | (0.159) | - |
| | > 24 months | -0.580*** | -0.598*** | -0.295* |
| | | (0.216) | (0.157) | (0.181) |

## Table 1.7. Re-default probability controlling for originator's and servicer's reputation

This table reports the estimation of equation (1) controlling for originator's and servicer's reputation. *Re-default* is a dummy for a loan that defaults within 6 months after it is modified, where default is defined as at least 90 days delinquency; *OS* is an indicator for a loan originated and serviced by the same institution; *Top 10 originators* is an indicator for the ten originators with the highest market share in the sample; *Top 10 originator-servicers* is an indicator for the ten institutions with the highest market share as both originators and servicers. Other control variables are listed in equation (1).

| | Re-default within 6 months | |
|---|---|---|
| VARIABLES | (1) | (2) |
| OS | -0.097** | -0.097* |
| | (0.048) | (0.056) |
| Top 10 originators | 0.020 | |
| | (0.031) | |
| OS*Top 10 originators | 0.026 | |
| | (0.069) | |
| Top 10 originator-servicers | | 0.023 |
| | | (0.049) |
| OS*Top 10 originator-servicers | | 0.023 |
| | | (0.095) |
| Other controls | Yes | Yes |
| Time FE | Yes | Yes |
| Servicer FE | Yes | Yes |
| State FE | Yes | Yes |
| Observations | 4,445 | 4,445 |
| Pseudo R | 0.179 | 0.179 |

## Table 1.8. Modification type and re-default probability

This table reports the estimation of equation (1) with modification type interaction terms. *Re-default* is a dummy for a loan that defaults within 6 months after it is modified, where default is defined as at least 90 days delinquency; *OS* is an indicator for a loan originated and serviced by the same institution; *Reduce principal* is an indicator for a loan receiving a reduction in outstanding principal; *Adjust rate* is an indicator for a loan receiving an adjustment in interest rate, *Adjust term* is an indicator for a loan receiving an adjustment in loan term; *Capitalize arrears* is an indicator for a loan receiving arrear capitalization. Other control variables include all the control variables listed in equation (1).

| VARIABLES | Re-default within 6 months |
|---|---|
| OS | 0.082 |
| | (0.052) |
| OS*Reduce principal | 0.017 |
| | (0.027) |
| OS*Adjust rate | -0.051 |
| | (0.036) |
| OS*Adjust term | - |
| | - |
| OS*Capitalize arrears | -0.147*** |
| | (0.038) |
| Other controls | Yes |
| Time FE | Yes |
| Servicer FE | Yes |
| State FE | Yes |
| Observations | 4,445 |
| Pseudo R | 0.182 |

## Table 1.9. Extent of modification

This table reports the summary statistics on the amount of payment change as a result of loan modification.

| | Payment change ($ amount) | | Payment change (% of original payment) | |
|---|---|---|---|---|
| | Mean | St. Dev. | Mean | St. Dev. |
| All loans | -214.27 | 370.54 | -15.63% | 22.57% |
| OS loans | -208.40 | 341.63 | -14.46% | 19.62% |
| Non-OS loans | -221.01 | 401.17 | -16.98% | 25.48% |
| t-stat | -1.23 | | -4.05*** | |

## Table 1.10. Informational advantages and re-default probability

This table reports the estimation of equation (1) with additional interaction terms. *Re-default* is a dummy for a loan that defaults within 6 months after it is modified, where default is defined as at least 90 days delinquency; *OS* is an indicator for a loan originated and serviced by the same institution; *Lowdoc* is a dummy for loans with low or no documentation; *Negative equity indicator* is a dummy for borrower with outstanding balance higher than his house value at the time of modification. Other control variables include all the control variables listed in equation (1).

| | Re-default within 6 months | |
|---|---|---|
| VARIABLES | (1) | (2) |
| OS | -0.063** | -0.056* |
| | (0.028) | (0.030) |
| OS*Lowdoc | 0.040 | 0.039 |
| | (0.039) | (0.040) |
| OS*Negative equity dummy | | -0.069** |
| | | (0.033) |
| Other controls | Yes | Yes |
| Time FE | Yes | Yes |
| Servicer FE | Yes | Yes |
| State FE | Yes | Yes |
| Observations | 4,611 | 4,611 |
| Pseudo R | 0.176 | 0.177 |

## Table 1.11. Age at modification and re-default probability

This table reports the estimation of equation (1) controlling for loan age at modification. *Re-default* is a dummy for a loan that defaults within 6 months after it is modified, where default is defined as at least 90 days delinquency; *OS* is an indicator for a loan originated and serviced by the same institution; *Loan age at modification* is a set of dummy variables for loan age at securitization, calculated as the number of months between loan origination and modification dates. Other control variables include all control variables listed in equation (1).

| VARIABLES | Re-default within 6 months |
|---|---|
| OS | -0.446*** |
| | (0.106) |
| Loan age at modification: | |
| 25 – 36 months dummy | -0.382*** |
| | (0.126) |
| 37 – 48 months dummy | -0.418*** |
| | (0.119) |
| 49 – 60 months dummy | -0.505*** |
| | (0.105) |
| 60 – 72 months dummy | -0.447*** |
| | (0.112) |
| >72 months | -0.407*** |
| | (0.107) |
| OS * Loan age at modification: | |
| 25 – 36 months dummy | 0.329** |
| | (0.129) |
| 37 – 48 months dummy | 0.391*** |
| | (0.124) |
| 49 – 60 months dummy | 0.445*** |
| | (0.107) |
| 60 – 72 months dummy | 0.480*** |
| | (0.120) |
| >72 months | 0.406*** |
| | (0.112) |
| Other controls | Yes |
| Time FE | Yes |
| Servicer FE | Yes |
| State FE | Yes |
| Observations | 4,611 |
| Pseudo R | 0.179 |

# Chapter 2

# Does peer sentiment affect firm investment? Evidence from the home building industry

## 2.1 Introduction

It is well established that peer effects play an essential role in shaping the behaviors of not only individuals but also institutions. Many studies have shown that the characteristics and behaviors of peer firms affect a wide variety of corporate decisions, including capital structure (Leary and Roberts, 2014; MacKay and Phillips, 2005; Welch, 2004), financial reporting (Beatty, Liao, and Yu, 2013), corporate governance (Foroughi, Marcus, Nguyen, and Tehranian, 2016; John and Kadyrzhanova, 2008) and corporate investment (Foucault and Fresard, 2014; Ozoguz and Rebello, 2013). For example, in an extensive survey on corporate governance by Graham and Harvey (2001), a large number of CFO respondents acknowledge the importance of considering peer firm financing decisions in forming their own capital structure policies. Consistent with this finding, industry average leverage ratios are found to be an important determinant of firms' capital structures in Welch (2004) and MacKay and Phillips (2005). Furthermore, a recent paper by Leary and Roberts (2014) documents that debt and equity issuance decisions of firms are associated with equity return shocks of peer firms. In addition to financing policies, stock prices of peer firms also influence corporate investment, as evidenced in Foucault and Fresard (2014) and Ozoguz and Rebello (2013).

Unlike firm characteristics (such as stock price returns) or behaviors (such as debt and equity issuance), which can be directly observed, firm sentiment has received much less attention from researchers in this field, due likely to the lack of an appropriate measure for firm sentiment. The effect of peer firm attitude on firm behaviors therefore remains unexplored to date. To the best of my knowledge, this paper is the first attempt to fill this gap in the literature. In particular, I focus on firms in the homebuilding industry and examine whether and how their building activities are influenced by peer sentiment. An important contribution of this paper lies in the use of the NAHB/Wells Fargo Housing Market Index (HMI) published by the National Association of Homebuilders (NAHB) to measure homebuilder belief. This index is derived from a monthly survey conducted by NAHB since 1985, in which approximately 400 members are asked to rate their views about the current and future conditions of the single-family housing market. The Housing Market Index is then constructed based on their responses, ranging from 0 (all respondents are pessimistic) to 100 (all respondents are optimistic). This index is widely followed by various participants in the homebuilding industry and often interpreted as homebuilder confidence. The primary research question in this paper asks whether there exists a significant relationship between this index and building activities.

This paper contributes a new insight into the effect of peer firms on not only firm behaviors but also the housing market in general. Although herding in the housing market is relatively well researched, the majority of the current literature concentrates on the demand side, that is, the herding behaviors of homebuyers and speculators. DeCoster and Strange (2012), Wang and Zhou (2000) and Grenadier (1996) are the few theoretical work focusing on the tendency of homebuiders to overbuild. This research compliments these earlier papers by offering the first empirical evidence on herding behaviors from the supply side of the housing market at the firm level.

Essential to the goal of this paper is determining if peer firm sentiment affects building activities beyond the effect of market fundamentals. In order to isolate the sentiment component in HMI, I first regress the index against a set of market fundamentals and then use the residuals as the measure of homebuilder sentiment. Using data from 22 public homebuilders in the US over 2003Q1-2016Q3, I find that a one-standard-deviation increase in the orthogonalized HMI will induce homebuilders to increase their lot inventory by 8.4%-12.6%. It also has a similar effect on the number of units built. This result remains robust even when homebuyer sentiment is controlled for. However, the positive relationship between building activities and peer sentiment is only strong when it is clear that the majority of the builders share similar beliefs. When the survey respondents seem to be divided in their opinions about the direction of the housing market, developers tend to reduce their building activities.

Research on peer effects often points to learning and reputational concerns as two potential explanations. The first sort of herding is associated with information cascades, where followers mimic leaders because the former infer useful information from the latter' actions (Banerjee, 1992). Alternatively, firms can be under pressure to herd in order to avoid falling behind their competitors and being perceived as losers (Scharfstein and Stein, 1990). It follows from these theories that herding behaviors are more pronounced among followers than leaders. Contrary to this hypothesis, I find that the biggest homebuilders in my sample are just as prone to peer sentiment as smaller firms. However, evidence shows that firms held by short term investors are more likely to follow their peers than those held by institutional shareholders. Finally, I test whether homebuilders who follow their peers will be rewarded in the short run with higher stock performance. Interestingly, firms that overbuild compared to their peers have lower stock returns while underbuilding is rewarded with higher stock prices, but this effect decreases as the magnitude of underbuilding increases. This result is consistent with our intuition that, since building is an irreversible investment, overbuilding is more detrimental than underbuilding.

## 2.2 Literature review

This paper is related to the growing strand of corporate finance literature focusing on peer effects on corporate decisions. Capital structure is one of the most well researched corporate policies. Industry average leverage ratios are found to be an important determinant of firms' capital structures in Welch (2004) and MacKay and Phillips (2005). Motivated by the prior findings that changes in stock prices contain useful information about financing decisions, Leary and Roberts (2014) use return shocks of peer firms as an instrument for their leverage ratios. They show that a firm's debt and equity issuance decisions are negatively and positively related to its peers' return shocks. The magnitude of this peer effect is found to be the largest compared to other determinants of capital structure, with a one-standard deviation increase in peer firms' leverage ratios leading to a 10% increase in firm *i*'s leverage ratio.

Investment decision is another corporate issue where peer effects are found to have a strong presence. Also using stock performance of peer firms as a measure of their influence, Foucault and Fresard (2014) document that a one standard deviation increase in peer stock prices induces a 5.9% increase in corporate investment. This positive relationship between investment and peer stock valuation is again confirmed in Ozoguz and Rebello (2013). A more recent paper by Fracassi (2016) takes a different approach in measuring peer influence by linking social ties among firm managers with the degree of similarity in their investment. Besides capital structure and investment, the behaviors and characteristics of peer firms have also been found to impact a wide variety of other corporate decisions, examples including corporate governance (Bouwman, 2011;

Foroughi et al., 2016; John and Kadyrzhanova, 2008), product differentiation (Foucault and Fresard, 2016), stock splitting decisions (Kaustia and Rantala, 2015), IPO decisions (Nelson, 2002), financial reporting practices (Beatty et al., 2013).

The literature offers two key theoretical justifications for the observed herding behaviors of firms, learning and reputational concerns, both of which are rational explanations. The learning motivation is often associated with the work of Banerjee (1992), who proposes that managers may rationally ignore their own information and rely more on the decisions of peers when their own signal appears noisy. Information cascade is especially likely when there are leaders in the industry who are perceived to possess better information (Bikhchandani, Hirshleifer, and Welch, 1998). The second theory suggests reputational concerns as a potential motive for herding. In the seminal paper by Scharfstein and Stein (1990), lower quality managers imitate actions by higher quality managers in order to mask their true type. In this setting, it is more beneficial for agents to follow the crowd than make efficient investment decisions because they can share the blame in the event of a bad outcome. Empirical research has found supporting evidence for both theories; for example, herding is more prevalent among smaller, younger, and less profitable firms (Leary and Roberts, 2014), or firms whose own stock price is less informative or managers appear less informed (Foucault and Fresard, 2014). However, Kaustia and Rantala (2015) find no benefits in mimicking stock splitting peers, suggesting that firm managers may simply succumb to imitation pressure.

Building on learning and reputational herding models, DeCoster and Strange (2012) demonstrate that both forces can lead to rational overbuilding in the residential market. On one hand, overbuilding can occur when later developers choose to follow early movers who receive inaccurate signals. On the other hand, herding can also arise because builders want to prevent banks from making inferences about their true quality. In either case, it is optimal for developers to herd even when they believe that the "wisdom of the crowd" is in the wrong. Two earlier papers adopt a game-theoretic frameworks to explain overbuilding in the real estate market. Grenadier (1996) treats development decisions as optimal strategies in a game of option exercise. Faced with declining demand level and building values, developers may simultaneously rush to build in fear of preemption by competitors, causing a concentration of building activities in a market downturn. Employing the game-theoretic approach, Wang and Zhou (2000) present a different framework based on two stages. In their model, developers decide on quantity first and price second. Regardless of their approaches, all three papers arrive at the same conclusion that overbuilding is the outcome of rational decision making on the part of builders.

## 2.3 Data and Methodology

### 2.3.1 Measure of peer sentiment

To measure homebuilder sentiment, I use the NAHB/Wells Fargo Housing Market Index developed by the National Association of Homebuilders[17]. The index is computed from responses to a monthly survey of NAHB members about their view of the single-family housing market. The survey asks approximately 400 homebuilders to rate market conditions for (1) current sales of new homes, (2) expected sales of new homes in the next six months, as well as (3) traffic of prospective buyers of new homes. In particular, builders rate current sales and sales expectations as "good," "fair" or "poor", and traffic of prospective buyers as "high to very high," "average" or "low to very low". For each of the three questions, the percentages of responses in the Good/High and Poor/Low categories are computed and seasonally adjusted. An index is computed for each series according to the formula (Good/High – Poor/Low +100)/2. The overall HMI is a weighted average

[17] Data can be obtained from: http://www.nahb.org/en/research/housing-economics/housing-indexes/housing-market-index.aspx

of these three component indices and their weights are based on their correlations with single-family housing starts. All indices range between 0 and 100, with an index number above 50 indicating that more builders view market conditions as good than poor. An index number of 100 means that all respondents answer Good/High and vice versa.

The index is covered regularly on major news media, such as The Wall Street Journal, Fox Business News, and CNBC, to name a few, as an indication of homebuilders' sentiment. Market analysts also commonly rely on HMI as an indicator of the direction towards which the housing market is heading in the near term. Most important for this paper, this index reveals the attitude of a large group of homebuilders in the country, hence serving as a good candidate for measuring peer firm belief. There are, however, two potential concerns that need to be addressed prior to using HMI for this purpose.

Firstly, as with any other survey data, one may question whether index users can trust that respondents answer the survey questions honestly. While there are no perfect tests to confirm this, the correlation between HMI and aggregate single-family housing starts and building permits from 1985 to 2016 is more than 0.77. In addition, Goodman (1994) and Marcato and Nanda (2016) show that the index is useful in predicting future housing starts. This suggests that homebuilders' rating of the market conditions in the survey is rather closely aligned with their actual behaviors.

The strong relationship between HMI and building activities suggests another concern with using the index as a measure of peer sentiment. Not surprisingly, the index also highly correlates with other macroeconomic factors, such as GDP and employment, because builder perceptions of the housing market often reflect a combination of both economic fundamentals and subjective judgement. In other words, the index can act as a proxy for macroeconomic conditions. It is thus important to show that peer belief has an incremental effect on building activities beyond the effect of fundamentals. To isolate the component of HMI that is independent of overall economic conditions, I regress the index against a set of market control variables: GPD growth, income growth, change in unemployment rate, change in mortgage rate, and one-quarter lagged change in the Case-Shiller house price index. All variables are measured at quarterly intervals from 1985Q1 to 2016Q3. This orthogonalization method has been used in several prior studies of sentiment in the stock market (see, for example, Baker and Wurgler, 2007; Kumar and Lee, 2006) as well as the housing market (Ling, Ooi, and Le, 2015). The regression results are shown in Table 1. Only GDP and house price growth appear statistically significant with positive signs as predicted. The residuals collected from this regression are then used to test if building activities are driven by the component of peer belief that is independent of fundamentals.

### 2.3.2 Base model and data

Data on homebuilding activities and other financial measures of homebuilders are obtained from SNL Financials, a data service provided by S&P Global Market Intelligence. The database includes 22 public homebuilders, which covers all public companies in the homebuilding industry in the U.S., except Green Brick Partners and producers of manufactured homes. The list of the sample homebuilders is shown in the Appendix. Traditionally, the homebuilding industry in the U.S. consists of three segments: large public homebuilders, large private home builders, and small private home builders. It is important to note that public homebuilders may be fundamentally different from the other two segments in their building capabilities. For example, Ambrose and Peek (2008) find that large public homebuilders have better access to capital, allowing them to grow their market shares at the expense of their private counterparts. Hence, the results in this research are limited to the public segment of the homebuilding industry; future research is need if one wants to understand private and small builders. Nevertheless, understanding the behaviors of these 22 public builders is especially important for the housing market because of their large market share in the total national supply. These companies are regularly present in the top 100 builders by sales and sales revenue reported by Builder Magazine. In 2015, 21 out of 22 are present in the top 100 builders and accounted for 166,258 sales, equivalent to 69.2% of the total sales by the top 100 builders and 33% of the total sales in the nation.

To study the effect of peer sentiment on homebuilding activities, I estimate the following regression model:

$$Building\ activities_{i,t} = \alpha + \beta HMI_{t-1} + \gamma Firm\ controls_{i,t-1} + \delta Market\ controls_{t-1} + \theta_i + \varepsilon_{i,t}, \tag{1}$$

where the dependent variable include (1) the number of lots owned and (2) number of units built by builder *i* in quarter *t*. All the explanatory variables are lagged by one quarter, but for robustness I also estimate the regression with two lags. The main variable of interest is HMI, which is the orthogonalized housing market index as described in the previous section. Since the original HMI is released monthly, I convert the series into quarterly data using the index numbers in the months of March (Q1), June (Q2), September (Q3) and December (Q4).

Firm controls include Tobin's Q and the ratio of cash flow to assets (Cashflow/Asset). These variables have been commonly found to affect investment decisions of firms, starting with the seminal work by Hayashi (1982) and Fazzari, Hubbard, and Petersen (1988) on Q theory and cash flow sensitivity of investment, respectively. Moreover, I control for the cost of debt of each firm, which is approximated by the ratio of interest paid in each quarter to total debt. Finally, equity issuance in times of overpriced stocks can also act as a cheap form of funding for investment. Following Polk and Sapienza (2009), I include the proceeds from sale of common equity (in log) in each quarter as another firm control variables. Regarding the control variables for market conditions, equation (1) takes into account the impact of GDP growth, income growth, change in unemployment rate, change in mortgage rate, and change in the Case-Shiller house price index. Lastly, the regression is run with a set of firm fixed effects $\theta_i$.

The study period covers 2003Q1-2016Q3 because of data unavailability for firm characteristics prior to 2003. In total there are 737 firm-quarter observations and there descriptive statistics are shown in Table 2.2. The original HMI has a mean of 49.5 over the 1985Q1-2016Q3 period, while the orthogonalized index has a mean of 0 as expected of a regression residuals series. On average, a builder in my sample has 62,478 lots in their inventory, with a low of 229 and a high of 396,000 lots. A useful alternative measure of building activities is number of units built in each quarter. Since the data for this variable are not directly available at the firm level, it is estimated using the following procedure. For each builder *i* in quarter *t*, I first estimate the number of units in their inventory by dividing the dollar value of inventory by the average sale price of their units in that quarter. The number of units built by builder *i* in quarter *t* is then calculated as the sum of the number of units sold and units in inventory. The estimated number of built units in each quarter for each builder ranges from 29 to 33,153, with an average of 6,131 units. The four firm control variables are similar to an average firm in other industries reported by prior studies.

## 2.4 Empirical results

### 2.4.1 Base results

The relationship between the orthogonalized builder sentiment index and building activities are shown in Table 2.3. Building activities are measured by the number of lots owned and units built, both expressed in log, in Panel A and Panel B, respectively. I include up to two lags of HMI. The first two specifications in column (1) and (2) do not include any control variables, while the next two specifications control for firm characteristics, and the last two control for both firm factors and economic conditions. The control variables are lagged by one quarter in column (1), (3) and (5), and two quarters in the remaining columns.

The coefficients on the first lag of HMI is positive and statistically significant at 1% in all specifications. The second lag is insignificant when market conditions are added in the last column, albeit sill positive as expected. In addition, the magnitude of the coefficients on the first lag consistently range between 0.008 and 0.012 throughout all specifications, suggesting that a one-standard-deviation increase in the orthogonalized HMI (10.52) will induce homebuilders to increase their lot inventory by 8.4%-12.6%. Given that the average number of lots owned by builders in my sample is 62,478 (Table 1), this translates to an average increase of 5,258-7,887 lots for each homebuilder, and a total of 115,679-173,519 lots for all 22 homebuilders in the sample. In comparison, the magnitude of the coefficients on the second lag reduces by 10 times once control variables are accounted for. Since the impact of the first lag of HMI appears much more robust and consistent, in all subsequent empirical tests, I will use the model specified in column (5) of Panel A, which includes only one lag of HMI and all other control variables.

Focusing on firm-specific control variables, I find that they are mostly not statistically significant, except for the positive effect of Cashflow/Asset. This result is in line with prior findings that cash flow is an important predictor of investment (see, for example, Fazzari et al., 1988; Lamont, 1997). On the other hand, Tobin's Q appears negative in most specifications and is statistically significant in the last column, contrary to the Q theory of investment first proposed by Hayashi (1982) that firms with higher Q values should invest more. Turning to the set of economic fundamentals, it is not surprising that homebuilders increase their land inventory in accordance with GDP growth, income growth and house price growth. However, the positive sign on mortgage rate change suggests that building activities are higher in times of higher interest rates, which seems counter-intuitive. The estimation results of all variables remain qualitatively the same when the number of units built is used as the dependent variable in Panel B of Table 2.3.

In Table 1.4, I investigate the individual effects of the three components of HMI, namely (a) current sale conditions, (b) sale conditions in the next six months, and (c) traffic of prospective buyers. As explained in the previous section, the survey asks homebuilders to provide their belief about these three aspects of the single-family housing market and their answers are used to compute the HMI. For each series, I orthogonalize them in a similar manner as the HMI (see Table 1 for the orthogonalization estimation) and use the residuals to estimate the results presented in Table 1.4. As shown, all three components are statistically and economically significant, and their effects are also comparable to each other in terms of magnitude.

### 2.4.2 Robustness tests

In the first column of Table 2.5, I test whether the effect of HMI remains robust when homebuyer sentiment is controlled for. To measure homebuyer sentiment, I use responses from the Survey of Consumers, which is conducted monthly by the Survey Research Center at the University of Michigan (see http://www.sca.isr.umich.edu/ for details). Approximately 500 households from all states in the U.S. are chosen for each survey. Among the 50 questions covered in the survey, I focus on the following question that addresses respondents' attitude about home-buying conditions: "Generally speaking, do you think now is a good time or a bad time to buy a house?" Responses include "good," "bad," and "uncertain". Furthermore, a follow-up question asks respondents to provide reasons for their answers. The reasons for the "good" response are classified into six groups: “prices will increase,” “prices low,” “interest rates low,” “rising interest rates,” “good investment,” and “time’s good.” I use the percentage of respondents who think it is a “good” time to buy “because price will increase” as my proxy for homebuyer sentiment. Previous research by Ling, Ooi, and Le (2015) and Piazzesi and Schneider (2009) has used the responses to these survey questions as indication of optimism among homebuyers.

Similar to the orthogonalization exercise described earlier, I regress this buyer sentiment series on a set of economic variables and use the residuals as the component of sentiment unexplained by fundamentals. Prior to orthogonalization, the correlation between the original HMI and the homebuyer sentiment index is 0.46, but it reduces to only 0.2 when using the two residual series. Though its magnitude decreases by half, the coefficient on HMI remains significant when buyer sentiment is controlled for in the first column of Table 2.5. Not surprisingly, homebuyer sentiment is also an important driver of housing supply. Comparing their coefficients, I find that a one-standard-deviation increase in builder sentiment (10.52) is associated with a 4.2% increase, while that of buyer sentiment (2.7) leads to a 12.4% increase in the number of lots owned by each homebuilder.

In the second column of Table 2.5, I examine the non-linearity of the peer firm sentiment effect by adding the square term of HMI into the model, which has a negative and significant coefficient as predicted. Finally, the last column of Table 2.5 provides another insight into the asymmetrical effect of sentiment. As previously explained, the original HMI ranges from 0 and 100, with an index number of 50 indicating that there are equal numbers of respondents feeling positive and negative about the housing market. In other words, index numbers close to 50 denote periods of high disagreement, while numbers closer to 0 or 100 are indicative of high consensus among builders. One would expect that the sentiment effect is weaker in the former case and stronger in the latter case. The last model in Table 2.5 includes a dummy for quarters when the original HMI numbers range from 40 to 60, as well as an interaction term between this dummy and the orthogonalized HMI.[18] Interestingly, the sum of the two coefficients on HMI and the interaction term is slightly negative, suggesting that builder tend to reduce their building activities when the beliefs of their peers appear unclear.

### 2.4.3 Which firms herd?

The literature suggests learning and reputational concerns as two potential motivations for these herding behaviors (Banerjee, 1992; Scharfstein and Stein, 1990). Firms with inferior private information can observe the actions of early movers and learn from them. Alternatively, firms can be under pressure to herd in order to avoid falling behind their competitors and being perceived as losers. When a large number of their peers feel optimistic about the market, homebuilders may ignore their own private signals and follow the crowd for fear of losing their window of opportunities if the optimistic forecasts actually materialize in the future. Even if the market takes an unexpected downturn, herding firms can still avoid damages to their reputation because they can “share the blame” with peers. Similarly, builders have strong incentives to reduce their investment when the majority of their peers indicate negative attitudes about market conditions, because they do not want to risk being singled out in case the pessimistic sentiment becomes reality.

These leaders-followers explanations imply that mimicking behaviors are likely more pronounced among smaller firms who are often informationally constrained. However, the result presented in the first specification in Table 2.6 does not support this conjecture. The indicator *Big*

[18] I also test alternative ranges between 30 and 70 and obtain similar results.

equals one for firm-quarters in the top 25% market capitalization. Notwithstanding its negative sign, the coefficient is not statistically significant as proposed by the learning hypothesis. Instead, this result indicates that big and small firms are equally likely to be driven by peer sentiment.

Furthermore, mimicking behaviors should increase with the risk of being punished for going against the "wisdom of the crowd", typically in the form of lower stock price performance. Polk and Sapienza (2009) theorize that shareholders' investment horizon is important in determining firms' incentive to cater to investors. Specifically, firm managers are more inclined to overinvest (underinvest) as a result of stock overvaluation (undervaluation) if firms are held by a higher proportion of short term investors whose frequent trades can influence stock prices to a great extent. Consistent with their theoretical prediction, the authors find that abnormal investment is more sensitive to mispricing for firms with higher share turnover, which is a proxy for the amount of short-term investors trading their stocks.

Following their reasoning, I test the hypothesis that builders held by fewer institutional shareholders are more likely to herd. Data on institutional ownership of outstanding common shares of my sample builders are obtained from FactSet, a popular database on corporate shares holding. HMI is interacted with the proportion of institutional ownership (in percentage) in the second column of Table 2.6, and a dummy variable indicating firm-quarters with institutional holdings less than 40% in the third column.[19] In both cases, the interaction terms are significant and consistent with my prior expectation that firms held by more short-horizon shareholders are

[19] The fraction of institutional ownership reported by FactSet can be higher than 100% for certain firm-quarter due to various reasons such as double counting or large amounts of short transactions. See their manuals for more details on their reporting standards. To alleviate this problem, in the last specification of Table 2.6, I use a binary indicator for institutional holdings less than 40% in place of the percentage numbers. The results are qualitatively similar when using lower thresholds (10%-40%).

more prone to herding. This result lends further support to the catering theory in the literature (Polk and Sapienza, 2009).

### 2.4.4 Benefits of herding

In this section, I test the hypothesis is that firms who follow their peers will be rewarded in the short run with higher stock performance. The regression model is specified as follows:

$$Return_{i,t} = \alpha + \beta Deviation_{i,t} + \gamma Deviation_{i,t}^2 + \delta Controls_t + \theta_i + \varepsilon_t, \tag{2}$$

where *i* and *t* denote firm and quarter, respectively. The dependent variable is the excess return of firm *i*'s stocks:

$$Return_{i,t} = \left(\frac{Price_{i,t+1}}{Price_{i,t}} - 1\right) * 100 - RF_t, \tag{3}$$

where $RF_t$ is the Treasury bill rate in period *t*, and $Price_{i,t}$ and $Price_{i,t+1}$ are the closing prices of firm *i* at time *t* and *t+1*, respectively.

The explanatory variable of interest is $Deviation_{i,t}$. For each builder *i* in each quarter, the predicted number of lots (in log) can be calculated based on the sentiment model shown in column 5 of Table 2.3. This is the level of land lots that builder *i* is expected to carry in quarter *t* if they follow peer sentiment. The difference between the actual and predicted lot inventory is then calculated, and $Deviation_{i,t}$ is defined as the absolute value of this difference. In other words, it is the absolute value of the residuals from estimating the model in column 5 of Table 2.3. Essentially, this variable is a measure of the degree firm *i* deviates from their expected level of lot inventory at quarter *t*. If firms mimic to cater to investors, we should observe higher errors to be associated with lower stock return. Hence, the coefficient $\beta$ is expected to be negative.

Regarding the control variables, I employ three popular asset pricing models: capital asset pricing model ($R_m$-$R_f$), Fama-French three-factor model ($R_m$-$R_f$, SMB, and HML), and Fama-French five-factor model ($R_m$-$R_f$, SMB, HML, RMW, CMA). Data on these variables, their definition and construction are available from Kenneth French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). These control variables are also calculated over the period from quarter *t* to quarter *t+1* in a similar manner to the *Return* variable. The regression model includes a set of firm fixed effect $\theta_i$.

Contrary to my expectation, both the *Deviation* variable and its square term appear insignificant in the first three models in Table 2.7. However, more interesting insights are revealed when the sign of the residuals is taken into consideration in the last model (column 4). *Positive error* is an indication variable for a positive residual, meaning the actual number of lots owned is higher than that suggested by the sentiment model. The variable *Deviation*, its square term and interaction term all become significant. The coefficient on *Deviation* suggests that firms which underbuild, defined as those with actual lot inventory lower than predicted level, have higher excess returns in the next quarter, but this reward decreases as the degree of underbuilding increases. On the contrary, firms that overbuild are punished with lower stock prices and this penalty increases further with the degree of overbuilding. This result is consistent with our intuition that, since building is an irreversible investment, overbuilding is more detrimental than underbuilding.

## 2.5 Conclusion

This paper examines whether and how building activities are influenced by peer sentiment, using the NAHB/Wells Fargo Housing Market Index as a measure of homebuilder belief. In order to isolate the sentiment component in HMI, I first regress the index against a set of market fundamentals and then use the residuals as the measure of homebuilder sentiment. Using data from 22 public homebuilders in the US over 2003Q1-2016Q3, I find that a one-standard-deviation increase in the orthogonalized HMI will induce homebuilders to increase their lot inventory by 8.4%-12.6%. It also has a similar effect on the number of units built. This result remains robust even when homebuyer sentiment is controlled for. However, the positive relationship between building activities and peer sentiment is only strong when it is clear that the majority of the builders share similar beliefs. When the survey respondents seem to be divided in their opinions about the direction of the housing market, developers tend to reduce their building activities.

Contrary to the learning hypothesis, I find that the biggest homebuilders in my sample are just as prone to peer sentiment as smaller firms. However, evidence shows that firms held by short term investors are more likely to follow their peers than those held by institutional shareholders. Finally, I test whether homebuilders who follow their peers will be rewarded in the short run with higher stock performance. Interestingly, firms that overbuild compared to their peers have lower stock returns while underbuilding is rewarded with higher stock prices, but this effect decreases as the magnitude of underbuilding increases.

## 2.6 References


Ambrose, B., and Peek, J. (2008). Credit Availability and the Structure of the Homebuilding Industry. *Real Estate Economics*, *36*(4), 659–692.

Baker, M., and Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, *21*(2), 129–151.

Banerjee, A. V. (1992). A simple model of herd behavior. *Quarterly Journal of Economics*, *107*(3), 797–817.

Beatty, A., Liao, S., and Yu, J. J. (2013). The spillover effect of fraudulent financial reporting on peer firms' investments. *Journal of Accounting and Economics*, *55*(2–3), 183–205.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *Journal of Economic Perspectives*, *12*(3), 151–170. http://doi.org/10.1257/jep.12.3.151

Bouwman, C. H. S. (2011). Corporate governance propagation through overlapping directors. *Review of Financial Studies*, *24*(7), 2358–2394.

DeCoster, G. P., and Strange, W. C. (2012). Developers, Herding, and Overbuilding. *Journal of Real Estate Finance and Economics*, *44*(1–2), 7–35.

Fazzari, S. M., Hubbard, R. G., and Petersen, B. C. (1988). Financing Constraints and Corporate Investment. *Brookings Papers on Economic Activity*, *1*, 141–195.

Foroughi, P., Marcus, A., Nguyen, V., and Tehranian, H. (2016). Peer Effects in Corporate Governance Practices : Evidence from Universal Demand Laws. *Working Paper*.

Foucault, T., and Fresard, L. (2014). Learning from peers' stock prices and corporate investment. *Journal of Financial Economics*, *111*(3), 554–577.

Foucault, T., and Fresard, L. (2016). Corporate Strategy , Conformism , and the Stock Market. *Working Paper*.

Fracassi, C. (2016). Corporate Finance Policies and Social Networks. *Management Science*, (Forthcoming).

Goodman, J. L. (1994). Using Attitude Data to Forecast Housing Activity. *Journal of Real Estate Research*, *9*(4), 445–453.

Graham, J. R., and Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, *60*(2–3), 187–243.

Grenadier, S. R. (1996). The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets. *Journal of Finance*, *51*(5), 1653–1679.

Hayashi, F. (1982). Tobin's Marginal Q and Average Q: A Neoclassical Interpretation. *Econometrica*, *50*(1), 213–224.

John, K., and Kadyrzhanova, D. (2008). Peer Effects in Corporate Governance. *Working Paper*.

Kaustia, M., and Rantala, V. (2015). Social learning and corporate peer effects. *Journal of*

*Financial Economics*, *117*(3), 653–669.

Kumar, A., and Lee, C. M. C. (2006). Retail Investor Sentiment and Return Comovements. *Journal of Finance*, *61*(5), 2451–2486.

Lamont, O. (1997). Cash Flow and Investment : Evidence from Internal Capital Markets. *Journal of Finance*, *52*(1), 83–109.

Leary, M. T., and Roberts, M. R. (2014). Do Peer Firms Affect Corporate Financial Policy? *Journal of Finance*, *69*(1), 139–178.

Ling, D. C., Ooi, J. T. L., and Le, T. T. T. (2015). Explaining house price dynamics: Isolating the role of nonfundamentals. *Journal of Money, Credit and Banking*, *47*(S1), 87–125.

MacKay, P., and Phillips, G. M. (2005). How does industry affect firm financial structure? *Review of Financial Studies*, *18*(4), 1433–1466.

Marcato, G., and Nanda, A. (2016). Information Content and Forecasting Ability of Sentiment Indicators: Case of Real Estate Market. *Journal of Real Estate Research*, *38*(2), 165–203.

Nelson, L. (2002). Persistence and Reversal in Herd Behavior: Theory and Application to the Decision to Go Public. *Review of Financial Studies*, *15*(1), 65–95.

Ozoguz, A., and Rebello, M. J. (2013). Information, competition, and investment sensitivity to peer stock prices. *Working Paper*.

Piazzesi, M., and Schneider, M. (2009). Momentum Traders in the Housing Market: Survey Evidence and a Search Model. *American Economic Review Papers and Proceedings*, *99*, 406–411.

Polk, C., and Sapienza, P. (2009). The stock market and corporate investment: A test of catering theory. *Review of Financial Studies*, *22*(1), 187–217.

Scharfstein, D. S. ., and Stein, J. C. . (1990). Herd Behavior and Investment. *American Economic Review*, *80*(3), 465–479.

Wang, K., and Zhou, Y. (2000). Overbuilding: A Game-Theoretic Approach. *Real Estate Economics*, *28*(3), 493–522.

Welch, I. (2004). Capital Structure and Stock Returns. *Journal of Political Economy*, *112*(1), 106–132.

**Table 2.1. Orthogonalization results**

This table reports the result of orthogonalizing HMI and its three components on a set of economic fundamental variables. *HMI* is the NAHB/Wells Fargo Housing Market Index. *Current sale conditions*, *Sale conditions in the next 6 months*, and *Buyer traffic* are the three component indices comprising HMI. *GDP growth* is the quarterly change in real GDP at the national level. *Income growth* is the quarterly change in real average household income at the national level. *Change in unemployment* is the quarterly change in national unemployment rate. *Change in mortgage rate* is the quarterly change in the average interest rate for 30-year fixed rate mortgages. *Change in HPI* is the quarterly change in the real Case-Shiller national home price index.

| | HMI | Sale conditions Current | Sale conditions Next 6 months | Buyer traffic |
|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) |
| GDP growth | 2.005*** | 2.078*** | 2.041*** | 1.790*** |
| | (0.528) | (0.585) | (0.555) | (0.428) |
| Income growth | 1.655 | 1.795 | 1.803 | 1.103 |
| | (1.224) | (1.354) | (1.284) | (0.992) |
| Change in unemployment | 2.901 | 2.799 | 4.230 | 2.063 |
| | (4.044) | (4.476) | (4.244) | (3.277) |
| Change in mortgage rate | 1.107 | 2.035 | 0.404 | -0.731 |
| | (2.788) | (3.086) | (2.926) | (2.259) |
| Change in HPI (lagged one quarter) | 6.856*** | 7.798*** | 6.572*** | 4.950*** |
| | (0.620) | (0.686) | (0.650) | (0.502) |
| Constant | 39.32*** | 42.08*** | 46.95*** | 30.05*** |
| | (2.004) | (2.218) | (2.104) | (1.624) |
| Observations | 127 | 127 | 127 | 127 |
| Adjusted R-squared | 0.614 | 0.620 | 0.569 | 0.578 |

**Table 2.2 Summary statistics**

This table reports the summary statistics of the sentiment indices and the homebuilder sample. *HMI* is the NAHB/Wells Fargo Housing Market Index. *Number of lots* in the number of land lots owned by developers. *Number of unit built* is the estimated number of units built by developers. *Tobin's Q* is the ratio of a firm's market capitalization to total assets. *Cashflow/Asset* is the ratio of cash flow to total assets. *Cost of debt* is the ratio of total amount of interest paid to total debt. *Common equity issuance* is the dollar value in millions of common equity issuance.

| Variable | Observations | Mean | Std. Dev | Min | Max |
|---|---|---|---|---|---|
| *Housing Market Index (HMI)* | | | | | |
| Original index | 127 | 49.05 | 17.27 | 9.00 | 78.00 |
| Orthogonalized index (residuals) | 127 | 0.00 | 10.52 | -33.79 | 19.81 |
| *Firm controls* | | | | | |
| Number of lots | 737 | 62,478.21 | 68,523.76 | 229.00 | 396,000.00 |
| Number of units built | 737 | 6,131.21 | 6,185.30 | 19.39 | 33,153.65 |
| Tobin's Q | 737 | 1.12 | 0.36 | 0.55 | 2.96 |
| Cashflow/Asset | 737 | 0.54 | 3.69 | -17.18 | 26.25 |
| Cost of debt (%) | 737 | 1.53 | 1.15 | 0.01 | 24.57 |
| Common equity issuance ($mil) | 737 | 7,185.41 | 22,126.03 | -4,526.00 | 238,886.00 |

**Table 2.3. Effect of peer sentiment on homebuilding activities**

This table reports the estimation of equation (1). *HMI* is the orthogonalized NAHB/Wells Fargo Housing Market Index. *Number of lots* in the number of land lots owned by developers. *Number of unit built* is the estimated number of units built by developers. *Tobin's Q* is the ratio of a firm's market capitalization to total assets. *Cashflow/Asset* is the ratio of cash flow to total assets. *Cost of debt* is the ratio of total amount of interest paid to total debt. *Common equity issuance* is the dollar value in millions of common equity issuance. *GDP growth* is the quarterly change in real GDP at the national level. *Income growth* is the quarterly change in real average household income at the national level. *Change in unemployment* is the quarterly change in national unemployment rate. *Change in mortgage rate* is the quarterly change in the average interest rate for 30-year fixed rate mortgages. *Change in HPI* is the quarterly change in the real Case-Shiller national home price index.

*Panel A. Effect of sentiment on number of lots owned*

| | Number of lots (log) | | | | | |
|---|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
| $HMI_{t-1}$ | 0.012*** | 0.008*** | 0.009*** | 0.008*** | 0.008*** | 0.011*** |
| | (0.001) | (0.001) | (0.002) | (0.001) | (0.001) | (0.001) |
| $HMI_{t-2}$ | | 0.010*** | | 0.007*** | | 0.001 |
| | | (0.001) | | (0.001) | | (0.001) |
| Tobin's $Q_{t-k}$ | | | -0.062 | 0.050 | -0.253* | -0.308*** |
| | | | (0.102) | (0.065) | (0.134) | (0.108) |
| Cashflow/Asset $_{t-k}$ | | | 0.044*** | 0.046*** | 0.020** | 0.016* |
| | | | (0.010) | (0.011) | (0.009) | (0.009) |
| Cost of debt $_{t-k}$ | | | -0.047 | -0.026 | -0.050 | -0.030 |
| | | | (0.041) | (0.029) | (0.038) | (0.024) |
| Common equity issuance (log) $_{t-k}$ | | | 0.031* | 0.027 | 0.032* | 0.025 |
| | | | (0.017) | (0.017) | (0.016) | (0.015) |
| GDP growth $_{t-k}$ | | | | | 0.032*** | 0.016* |
| | | | | | (0.007) | (0.009) |
| Income growth $_{t-k}$ | | | | | 0.092*** | 0.077*** |
| | | | | | (0.016) | (0.014) |
| Chang in unemployment $_{t-k}$ | | | | | 0.104 | 0.071 |
| | | | | | (0.072) | (0.077) |
| Change in mortgage rate $_{t-k}$ | | | | | 0.092** | 0.077** |
| | | | | | (0.037) | (0.037) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Change in HPI $_{t-k}$ | | | | | 0.047*** | 0.092*** |
| | | | | | (0.011) | (0.014) |
| Constant | 10.532*** | 10.562*** | 10.749*** | 10.664*** | 10.809*** | 10.965*** |
| | (0.005) | (0.008) | (0.200) | (0.169) | (0.216) | (0.199) |
| Observations | 787 | 787 | 539 | 533 | 539 | 533 |
| Number of builders | 22 | 22 | 22 | 22 | 22 | 22 |
| Adjusted R-squared | 0.079 | 0.130 | 0.263 | 0.360 | 0.372 | 0.499 |

*Panel B. Effect of sentiment on number of units built*

| | Number of units built (log) | | | | | |
|---|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
| $HMI_{t-1}$ | 0.013*** | 0.009*** | 0.011*** | 0.009*** | 0.009*** | 0.013*** |
| | (0.002) | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) |
| $HMI_{t-2}$ | | 0.011*** | | 0.008*** | | 0.001 |
| | | (0.001) | | (0.002) | | (0.002) |
| Tobin's Q $_{t-k}$ | | | -0.247 | -0.218 | -0.433 | -0.639** |
| | | | (0.194) | (0.182) | (0.260) | (0.289) |
| Cashflow/Asset $_{t-k}$ | | | 0.024*** | 0.029*** | 0.002 | -0.002 |
| | | | (0.008) | (0.009) | (0.008) | (0.008) |
| Cost of debt $_{t-k}$ | | | -0.050 | -0.042 | -0.050 | -0.041 |
| | | | (0.046) | (0.046) | (0.041) | (0.036) |
| Common equity issuance (log) $_{t-k}$ | | | 0.015 | 0.022 | 0.017 | 0.021* |
| | | | (0.015) | (0.015) | (0.013) | (0.012) |
| GDP growth $_{t-k}$ | | | | | 0.032*** | 0.016* |
| | | | | | (0.007) | (0.009) |
| Income growth $_{t-k}$ | | | | | 0.027*** | 0.002 |
| | | | | | (0.009) | (0.011) |
| Chang in unemployment $_{t-k}$ | | | | | 0.085*** | 0.088*** |
| | | | | | (0.021) | (0.017) |
| Change in mortgage rate $_{t-k}$ | | | | | 0.067 | 0.022 |
| | | | | | (0.048) | (0.062) |
| Change in HPI $_{t-k}$ | | | | | 0.197*** | 0.081 |
| | | | | | (0.053) | (0.058) |
| Constant | 8.155*** | 8.186*** | 8.801*** | 8.723*** | 8.861*** | 9.104*** |
| | (0.007) | (0.011) | (0.284) | (0.267) | (0.336) | (0.354) |
| Observations | 782 | 782 | 526 | 519 | 526 | 519 |
| Number of builders | 22 | 22 | 21 | 21 | 21 | 21 |
| Adjusted R-squared | 0.0837 | 0.135 | 0.153 | 0.228 | 0.256 | 0.382 |

## Table 2.4. Components of HMI

This table reports the estimation of equation (1) using the three component indices: *Current sale conditions*, *Sale conditions in the next 6 months*, and *Buyer traffic*. Control variables include: Tobin's Q, Cashflow/Asset, Cost of debt, Common equity issuance, GDP growth, Income growth, Change in unemployment, Change in mortgage rate, Change in HPI.

| | Number of lots (log) | | |
|---|---|---|---|
| VARIABLES | (1) | (2) | (3) |
| Current sale conditions | 0.008*** | | |
| | (0.001) | | |
| Sale conditions in the next 6 months | | 0.010*** | |
| | | (0.001) | |
| Traffic of prospective buyers | | | 0.009*** |
| | | | (0.002) |
| Control variables | Yes | Yes | Yes |
| Observations | 539 | 539 | 539 |
| Number of builders | 22 | 22 | 22 |
| Adjusted R-squared | 0.369 | 0.388 | 0.359 |

## Table 2.5. Robustness tests

*HMI* is the orthogonalized NAHB/Wells Fargo Housing Market Index. *Buyer sentiment* is the percentage of buyers who think it is a good time to buy a house because "prices will increase" in the Survey of Consumer by the University of Michigan. It is orthogonalized against a set of economic fundamental variables. *Indicator(40<=HMI$_{t-1}$<=60)* is a dummy variable for quarters in which the original HMI numbers range from 40 to 60. Control variables include: Tobin's Q, Cashflow/Asset, Cost of debt, Common equity issuance, GDP growth, Income growth, Change in unemployment, Change in mortgage rate, Change in HPI.

| | Number of lots (log) | | |
|---|---|---|---|
| VARIABLES | (1) | (2) | (3) |
| $HMI_{t-1}$ | 0.004*** | 0.0064*** | 0.012*** |
| | (0.001) | (0.0013) | (0.002) |
| Buyer sentimen$_{t-1}$ | 0.046*** | | |
| | (0.007) | | |
| $HMI^2_{t-1}$ | | -0.0002** | |
| | | (0.0001) | |
| Indicator(40<=HMI$_{t-1}$<=60) | | | -0.050 |
| | | | (0.044) |
| HMI$_{t-1}$*Indicator(40<=HMI$_{t-1}$<=60) | | | -0.013*** |
| | | | (0.002) |
| Control variables | Yes | Yes | Yes |
| Observations | 539 | 539 | 539 |
| Number of builders | 22 | 22 | 22 |
| Adjusted R-squared | 0.419 | 0.374 | 0.392 |

## Table 2.6. Which firms herd?

*HMI* is the orthogonalized NAHB/Wells Fargo Housing Market Index. *Big* is a dummy for firm-quarters in the top 25% of the sample in terms of market capitalization. *Institutional ownership* is the percentage of outstanding common shares held by institutional investors. *Indicator (Institutional ownership < 40%)* is a dummy for firm-quarters with institutional holding less than 40%. Control variables include: Tobin's Q, Cashflow/Asset, Cost of debt, Common equity issuance, GDP growth, Income growth, Change in unemployment, Change in mortgage rate, Change in HPI.

| | Number of lots (log) | | |
|---|---|---|---|
| VARIABLES | (1) | (2) | (3) |
| $HMI_{t-1}$ | 0.0088*** | 0.0261*** | 0.0078*** |
| | (0.0014) | (0.0059) | (0.0014) |
| Big | 0.3168*** | | |
| | (0.1045) | | |
| $HMI_{t-1}$*Big | -0.0038 | | |
| | (0.0027) | | |
| Institutional ownership (%) | | 0.0026 | |
| | | (0.0027) | |
| $HMI_{t-1}$*Institutional ownership (%) | | -0.0002*** | |
| | | (0.0001) | |
| Indicator (Institutional ownership < 40%) | | | -0.5437 |
| | | | (0.3654) |
| $HMI_{t-1}$* Indicator (Institutional ownership < 40%) | | | 0.0376*** |
| | | | (0.0112) |
| Control variables | Yes | Yes | Yes |
| Observations | 539 | 539 | 539 |
| Number of builders | 22 | 22 | 22 |
| Adjusted R-squared | 0.415 | 0.389 | 0.405 |

### Table 2.7. Herding behaviors and stock price returns

*Deviation* is the residuals obtained from estimating equation (1). *Positive error* is an indication variable for positive residuals. Control variables in the CAPM model include ($R_m$-$R_f$). Control variables in the Fama-French three-factor model include $R_m$-$R_f$, SMB, and HML. Control variables in the Fama-French five-factor model include $R_m$-$R_f$, SMB, HML, RMW, CMA.

| | Stock return | | | |
|---|---|---|---|---|
| | CAPM | 3-factor | 5-factor | 5-factor |
| VARIABLES | (1) | (2) | (3) | (4) |
| Deviation | 6.711 | 5.157 | 5.474 | 20.977*** |
| | (6.073) | (6.756) | (6.067) | (7.216) |
| Deviation $^2$ | -1.4709 | -1.3177 | -1.5894 | -3.4303*** |
| | (1.4835) | (1.6443) | (1.5559) | (1.2021) |
| Positive error | | | | -4.962 |
| | | | | (4.702) |
| Deviation* Positive error | | | | -22.504*** |
| | | | | (6.464) |
| Control variables | Yes | Yes | Yes | Yes |
| Observations | 531 | 531 | 531 | 531 |
| Number of builders | 22 | 22 | 22 | 22 |
| Adjusted R-squared | 0.0943 | 0.143 | 0.156 | 0.176 |

# Chapter 3

# Comovement between Real Estate and the Financials Sector: Evidence from the New GICS Structure

## 3.1 Introduction

Aug 31, 2016 is marked as one of the most important dates for real estate as it become an independent sector in the Global Industry Classification Standard (GICS). Developed by Standard and Poor's and Morgan Stanley Capital International, GICS is a standardized classification system for equities that is commonly used by investors worldwide in making asset allocation and portfolio diversification decisions. Since the GICS structure was first introduced in 1999, real estate had been classified as an industry group under the Financials sector, together with banks and insurance companies. In a press release in 2014, Standard and Poor's announced that real estate would be elevated to an independent sector in the GICS system (effective Aug 31, 2016) following their finding that "…Real Estate is now viewed as a distinct asset class and is increasingly being incorporated separately into the strategic asset allocation of asset owners."

This move was highly applauded by market participants, especially real estate professionals, as a timely recognition of the growing prominence and distinction of real estate. This paper is the first to empirically study the effect this momentous event on the real estate sector. The main question in this research asks if this separation will reduce comovement between real estate and financials stocks, leading to enhanced diversification benefits. This is analogous to prior findings in the literature on index inclusion, which suggests that including a new stock into an index increases its correlation with other index members and divergence from non-index peers

(Vijh, 1994; Goetzmann and Massa, 2003; Barberis and Shleifer, 2003). In the case of REITs, Ambrose, Lee, and Peek (2007) examine the change in beta surrounding the addition of several REITs into the S&P 400, S&P 500 and S&P 600 indices. They show that the returns on non-index REITs become more highly correlated with the returns on the indices after some of their peers join those indices, and attribute this effect to spillovers of investor sentiment and market frictions across firm categories.

Using Real Estate Investment Trusts (REITs) to represent the new GICS Real Estate sector, I find that their correlation with the Financials sector fell from 0.568-0.775 to 0.338-0.581 after their departure. The reduction in their connection occurred first at the time of announcement (10 Nov 2014), and again at the time the change came into effect (1 Sep 2016). In addition, for robustness test I also examine the link between real estate, financials, and a third sector that did not experience any changes to its structure – the Industrials sector. Prior to the event, the correlation coefficients of REITs and Banks with Industrials were relatively close to each other, but REIT stocks then became less while Banks stocks grew more connected with the Industrials sector. Together, these findings support my hypothesis that real estate exhibits improved diversification benefits as a result of its separation from the Financials sector.

Turning to return volatility, my results show that REIT returns became as much as 60% less volatile after it became an independent sector.  The financials sector also exhibits lower return volatility during this period in but it is not statistically significant. Finally, I find that there was no change in the sensitivity of REIT trading volume to that of the Financials' sector. Contrary to prior expectations by market participants, the daily trading volatility of REITs also did not lower. Taken together, it appears that becoming an independent sector did not affect trading activities in the REIT market, at least in the short post-implementation period covered in this paper.

## 3.2 Literature review

This paper is related to the extensive debate on the extent to which REIT returns correlate with stock market factors relative to the performance of the real estate assets underlying their portfolios. On one hand, the literature seems to reach a consensus that there is a strong long-run relationship between public (REITs) and private real estate returns (Oikarinen, Hoesli, and Serrano, 2011; Yunus, Hansz, and Kennedy, 2012; Boudry, Coulson, Kallberg, and Liu, 2012; Hoesli and Oikarinen, 2012). On the other hand, there is only weak evidence on the contemporaneous association between them, due to the long trading lag in the private real estate market that requires more time for prices to adjust to new information (Geltner and Kluger, 1998; Ling and Naranjo, 2015). With regards to REIT linkages with the stock market, research has produced evidence suggesting that the strength of their relationship changes in different market conditions, such as in periods of high stock market volatility (Chong, Miffre, and Stevenson, 2009; Liow, Ho, Ibrahim, and Chen, 2009). It is also time-varying (Clayton and MacKinnon, 2001; Cotter and Stevenson, 2006). Interestingly, a recent paper by Case, Yang, and Yildirim (2012) establishes three distinct phases with three structural breaks in the REIT-stock dependency pattern. The first one is the pre-1991 period, when the correlation was 60% without trend. The second phase started in 1991, which marked the modern REIT era, and the final phase started in 2001 with the first inclusion of REITs in the S&P stock market indices. The correlations dropped to 30% in the second period but reverted back to their original level of 60% in the final phase.

Also highly relevant to the topic of this research is the literature on the effect of index inclusion. A common conclusion in this area is that index inclusion leads to an increase in comovement of the added firms with their index peers. For example, following index addition, Vijh (1994) finds a change in the beta on the value-weighted portfolio of NYSE and AMEX stocks, and Barberis et al. (2005) shows similar results for the S&P 500 index. Goetzmann and Massa (2003) attributes this effect to the high correlation of in- and outflows of index investment capital. Barberis and Shleifer (2003) and Barberis, Shleifer and Wurgler (2005) develop a theoretical model and empirically confirm two non-fundamental factors driving this excess return comovement. First, changes in sentiment of investors whose trading is based primarily on asset categories can cause large price movements in assets in the same category. The second explanation comes from market frictions, causing assets in different categories to incorporate new information at a different speed than those in the same category.

Evidence on the effects of index inclusion of REITs is relatively new and sparse. The few papers in this area show consistent evidence that pricing efficiency improves with index membership (Aguilar, Boudry, and Connolly, 2015; Huang, Su, and Chiu, 2009; Jirasakuldech and Knight, 2005). On the comovement of REITs, Ambrose, Lee, and Peek (2007) examine the change in beta surrounding the addition of several REITs into the S&P 400, S&P 500 and S&P 600 indices. They show that the returns on non-index REITs become more highly correlated with the returns on the indices after some of their peers join those indices, and attribute this effect to spillovers of investor sentiment and market frictions across firm categories. This finding is further supported in a more recent paper by Pavlov, Steiner and Watcher (2017). However, the long-run correlation between REIT returns remains largely dependent on their similarities in property types and geographic exposure, even after controlling for index inclusion. In fact, index membership enhances the impact of property fundamentals on REIT returns, consistent with improved pricing efficiency.

## 3.3 Data and methodology

### 3.3.1 Background on the Global Industry Classification Standards

Standard & Poor's (S&P) and Morgan Stanley Capital International (MSCI) teamed up in 1999 to create a set of global sector and industry definitions to classify companies based on their primary business activity, known as the Global Industry Classification Standards (GICS). The GICS hierarchy has 4 levels: the most granular tier is sector, followed by industry group, industry and subindustry. At inception, GICS classification system comprised of 10 sectors, 24 industry groups, 67 industries and 156 subindustries. MSCI and Standard & Poor's review it annually and, when significant changes in the marketplace require, update it to ensure that the framework reflects the current state of all industries.

In a press release on Nov 10, 2014, MSCI and S&P announced that they would add real estate as an 11th sector to GICS, in order to reflect the increasingly popular view that real estate was an independent asset class. The announcement also requested feedback from market participants on their proposal to implement the change after the market close on August 31, 2016. This implementation date was subsequently confirmed in another press release on March 13, 2015. Although MSCI and S&P had introduced several adjustments to the GICS system in the past, this was the first time a change at the sector level was made. Prior to this change, real estate was an industry group under the Financials sector, together with 3 other industry groups, namely Banks, Diversified Financials, and Insurance. The new Real Estate sector has only 1 industry group (Real Estate), under which there are two industries: (1) Equity Real Estate Investment Trusts,[20] and (2) Real Estate Management and Development. As its name suggests, the first industry consists of all

[20] Mortgage REITs remain in the Financials sector.

equity REITs, which are further classified into 8 subindustries based on their type of property focus (for example, Residential REITs, Retail REITs…). Equity REITs account for almost 98% of the market capitalization of the whole sector. The second industry makes up the remaining 2% and includes companies engaged in real estate development, management, and service businesses. **Figure 3.1** shows the old and new GICS structures with a breakdown of the Financials and Real Estate sector.

Positive feedback on this change from the marketplace confirmed that real estate was viewed as a distinct asset class. At the time of announcement, market participants predicted that the change will brought about new money into real estate stocks due to improved visibility. In addition, both MSCI and S&P employ the GICS structure to create their indexes, including the widely followed S&P 500, which are commonly used as benchmarks by investors. The new real estate sector was hence predicted to also cause significant portfolio restructuring.

### 3.3.2 Methodology

The main hypothesis in this paper predicts that the correlation between real estate stocks and financials stocks will be reduced following the emancipation of the former. The change in their correlation is estimated using the following regression model:

$$REIT_t = \alpha + \beta_1 Financials_t + \beta_2 Financials_t * Event_t + \gamma Event_t + \varepsilon_t, \quad (1)$$

where *t* denotes date. The study period runs from 10 Nov 2010, covering 4 years before the new sector was announced for the first time, to 8 Jun 2017. I do not include data prior to 2010 in order to avoid any possible contamination effect from the 2008 financial crisis. In addition, lengthening the estimation window also introduces risks of confounding factors. The error terms are clustered by quarter.

The dependent variable is the daily return of equity REITs as measured by the FTSE NAREIT All Equity REIT Index (hereafter NAREIT index), which contains all tax-qualified REITs that meet minimum asset, size and liquidity criteria (see http://www.ftse.com/products/indices/NAREIT for more details). Since they make up the bulk of the new Real Estate sector (98% of total market capitalization), I use equity REITs to represent the whole sector. $Financials_t$ is the daily return of each of the three industry groups in the Financials sector: banks, diversified financials, and insurance. Their daily returns are measured by the S&P 1500 Banks (hereafter Banks), S&P 1500 Diversified Financials (hereafter Diversified), and S&P 1500 Insurance (hereafter Insurance) indexes. These indexes comprise those companies included in the S&P Composite 1500 index[21] that are classified as members of the Banks, Diversified Financials, and Insurance industry groups.

The dummy variable $Event_t$ takes the value of 1 if time *t* is after one of the three events: announcement of the new sector (10 Nov 2014), confirmation of the implementation date (13 March 2015), and implementation date (1 Sep 2016). In order to understand their incremental effects on REIT correlations, I estimate regression (1) separately for each of the three events. The estimation windows for the regressions are depicted in **Figure 3.2**. In particular, to estimate the announcement effect, the first regression covers 11/10/2010 – 03/12/2015 (Period 1), with $Event_t$ equaling 1 for observations from 11/11/2014 to 03/12/2015. The second regression use data between 11/11/2014 and 08/31/2016 (Period 2) to measure the effect of confirming the effective date of the change, and the last regression is estimated for the 03/14/2015 – 06/08/2017 period

[21] The S&P Composite 1500 combines three indices: S&P 500, S&P MidCap 400, and S&P SmallCap 600. It covers approximately 90% of the U.S. market capitalization. It is designed for investors seeking to replicate the performance of the U.S. equity market or benchmark against a representative universe of tradable stocks (see https://us.spindices.com/indices/equity/sp-composite-1500 for more information).

(Period 3). The two coefficients of interests in this test are $\beta_1$ and $\beta_2$. The coefficient $\beta_1$ on the variable $Financials_t$ is the REITs beta with respect to Financials stocks prior to the corresponding event, while $\beta_2$, the coefficient on the interaction term, is the change in beta induced by the event.

## 3.4 Empirical results

### 3.4.1 Descriptive statistics

Panel A of **Table 3.1** presents the summary statistics of the daily returns of equity REITs and the other 3 industry groups under the Financials sector over the full study period, as measured by the NAREIT index and S&P 1500 indexes. REITs exhibits a slightly lower average return (0.034%) than the other groups, while the banking industry registers the highest return (0.053%). In the last column, Augmented Dicker-Fuller tests confirm that all series are stationary at the 1% level, allowing us to performing meaningful regression analysis on them.

In Panel B, I compute simple pairwise correlation coefficients for two sub-periods, using the announcement date 11/10/2014 as the cut-off point. This allows us to compare their comovement before and after the market first learned about the decision to add a new Real Estate sector into the GICS system.[22] As shown in the first column of Panel B, the correlation between REITs and the three industry groups ranged from 0.73 to 0.78 before the announcement date. Consistent with our expectation, these correlation coefficients reduced to only 0.6-0.7 post-announcement. The separation of real estate from financials stocks seems to have significantly reduced the latter's influence on the performance of the former. In comparison, the correlation

---

22 My search in the Factiva database did not indicate that there were any news about the possibility of a new GICS sector prior to the S&P press release on 11/10/2014. Thus, it is reasonable to assume that the market first knew about their decision on that day.

among Banks, Diversified and Insurance remain relatively the same in both periods, ranging between 0.89 and 0.91 (column 2 and 3).

### 3.4.2 Baseline results

Turning to our formal regression analysis, **Table 3.2** presents the estimation results of model (1), estimated separately for each of the three sub-periods corresponding to the announcement, confirmation and implementation dates. Focusing first on the REIT betas with respect to Banks in the first three columns, I find that their pre-event correlation was 0.609, as denoted by the coefficient on the S&P 1500 Banks variable for Period 1 (column 1). The interaction term Announce*SP1500 appears negative, and its magnitude suggests that the beta reduced by 0.401 units after the decision to add a new sector was first announced. This is both statistically and economically significant. The new correlation for the post-announcement period is therefore 0.208, which is essentially the coefficient reported in the first row of Period 2 (column 2). The interaction term in this case is positive but statistically indifferent from 0, implying that confirming the implementation date had no noteworthy incremental effect on the relationship between REIT and bank stocks. This finding is not surprising, given that the confirmation is actually immaterial and should have been expected by the market. In addition, since this implementation date was announced and confirmed well in advance, we would also expect the event itself to have no additional effect beyond the market reaction at the time of announcement. However, interestingly, the results for Period 3 indicate that REIT beta decreased further by 0.304 unit after the change became effective on 1 Sep 2016, which is statistically and economically significant. Nonetheless, this finding can be explained by the fact that many funds, such as pension funds, have strict regulations about the asset classes that they invest in and hence were unable to restructure their portfolios until the new sector is officially in effect.

The results are qualitatively similar when we examine the comovement between REITs and Diversified Financials equities in column 4 to 6, with the exception that the implementation event has the strongest effect. In the case of the Insurance industry group, the effect of the implementation is still negative but statistically insignificant. To summarize the baseline results, I find that REITs stocks became less correlated with Financials stocks when the new sector was first announced in 2014, and their connection further weakened as the new sector came into effect in 2016. These observations support my hypothesis that the elevation of real estate to a standalone sector not only validates but also enhances its important role in a well-diversified portfolio.

### 3.4.3 Correlation with other sectors

In this section, I turn to examine the link between real estate, financials, and a third sector that did not experience any changes to its structure – the Industrials sector. Specifically, I estimate and compare the REIT-Industrials versus Banks-Industrials correlation. For clarity, in all subsequent tests I simplify the reported regression results with the following two simplifications. First, I use Banks as a representative of the Financials sector, but the results are qualitatively similar using the Diversified Financials and Insurance industry groups. Second, I do not differentiate the incremental effects of announcement, confirmation, and implementation; the only event considered is the announcement date, and the variable $Event_t$ will take the value of 1 for all observations after from 11/10/2014 till the end of the study period.

As shown in the first two columns of **Table 3.3**, REIT stocks became less while Banks stocks grew more connected with the Industrials sector following REITs' promotion. Prior to the event, the correlation coefficients of REITs and Banks with Industrials were relatively close to each other, 0.802 and 1.030, respectively. The coefficients on the two interaction terms suggest that the former then reduced to only 0.587 while the latter went up to 1.193 in the post-announcement period. The difference between these two coefficients are significant at the 1% level ($\chi^2$=15.921). Hence, this observation further confirms my earlier result that the real estate sector exhibited improved diversification benefits as a result of its separation from the Financials sector.

As a comparison, I present the correlation between the Information Technology (IT) and the Industrials sector in the last column of **Table 3.3**. The IT sector also did not experience any changes to its classification, and thus can serve as a control group. Consistent with prior expectation, I do not observe any discernible change in its association with the Industrials sector. The difference between the coefficient of the interaction term in this model and that in the REIT model (column 1) is again statistically significant at the 1% level ($\chi^2$=28.68).

### 3.4.4 Return volatility

It follows from the above findings that leaving the Financials sector may also benefit REITs by reducing their return volatility. To test this prediction, I compare the standard deviation of the daily returns of REITs and Banks pre- and post-announcement in **Table 3.4**. The standard deviation is calculated from the beginning of the study period (11/10/2010) till 11/09/2014 for the "Before" period, and from 11/11/2014 till the end (06/08/2017) for the "After" period.

Both industry groups exhibited lower return volatility in the second period, but only the reduction in REIT volatility is statistically significant. In terms of economic magnitude, it is noteworthy that REIT returns became as much as 60% less volatile due to their departure. This finding implies that there were likely volatility spillover from the Financials industries to REITs in the past, but this spillover effect was drastically weakened when the link between the two sectors was severed. Although the standard deviation of Banks also fell by one third, it is not statistically significant.

### 3.4.5 Trading volume

Finally, in this section I examine if separating REITs from the Financials sector has any effect on their trading volume. The regressions reported in **Table 3.5** are similar to model (1) except that the change in the log of daily trading volume is used in place of daily return. Data on trading volume of REITs are provided by SNL Financial – a database by S&P Global Market Intelligence, and the trading volume of Banks is from Factset by Factset Research System. Intuitively, we would expect a reduction in the sensitivity of REIT trading volume to that of the Financials sector as a result of the former's exclusion from the latter. Contrary to this expectation, I do not find evidence that there existed significant changes in their relationship: none of the coefficients on the interaction terms for the three events – announcement, confirmation and implementation – are statistically significant. On their volatility, in **Table 3.6** I find that REITs experienced a small increase that is statistically significant but economically trivial. This is inconsistent with the popular conjecture by REIT professionals that this structure change would lower REITs' trading volatility. On the other hand, banks stocks had slightly lower trading volume volatility. Overall, it appears that becoming an independent sector did not affect trading activities in the REIT market, at least in the short run. When it is possible to extend the study period, future research may be able to detect any changes in the longer run.

## 3.5 Conclusion

When real estate was separated from the Financials sector and promoted to a distinct sector in the GICS structure in 2016, it was warmly welcomed by the market as a validation of the growing prominence and independence of real estate. This paper is the first to empirically study the effect this momentous event on the real estate sector. Using Real Estate Investment Trusts (REITs) to represent the new GICS Real Estate sector, I find that their correlation with the Financials sector fell from 0.568-0.775 to 0.338-0.581 after their departure. The reduction in their connection occurred first at the time of announcement (10 Nov 2014), and again at the time the change came into effect (1 Sep 2016). In addition, for robustness test I also examine the link between real estate, financials, and a third sector that did not experience any changes to its structure – the Industrials sector. Prior to the event, the correlation coefficients of REITs and Banks with Industrials were relatively close to each other, but REIT stocks then became less while Banks stocks grew more connected with the Industrials sector. Together, these findings support my hypothesis that the real estate sector exhibits improved diversification benefits as a result of its separation from the Financials sector.

Turning to return volatility, my results show that REIT returns became as much as 60% less volatile after it became an independent sector. The financials sector also exhibits lower return volatility during this period in but it is not statistically significant. Finally, I find that there was no change in the sensitivity of REIT trading volume to that of the Financials' sector. Contrary to prior expectations by market participants, the daily trading volatility of REITs also did not lower. Taken together, it appears that becoming an independent sector did not affect trading activities in the REIT market, at least in the short post-implementation period covered in this paper.

## Figure 3.1. GICS structures before and after 31 Aug 2016

This figure shows the GICS structures with a detailed breakdown of the Financials and Real Estate sector

| Sector | Industry group | Industry | Sub-industry |
|---|---|---|---|
| Financials | Banks | Banks | Diversified Banks |
| | | | Regional Banks |
| | | Thrifts & Mortgage Finance | Thrifts & Mortgage Finance |
| | Diversified Financials | Diversified Financial Services | Other Diversified Financial Services |
| | | | Multi-Sector Holdings |
| | | | Specialized Finance |
| | | Consumer Finance | Consumer Finance |
| | | Capital Markets | Asset Management & Custody Banks |
| | | | Investment Banking & Brokerage |
| | | | Diversified Capital Markets |
| | Insurance | Insurance | Insurance Brokers |
| | | | Life & Health Insurance |
| | | | Multi-line Insurance |
| | | | Property & Casualty Insurance |
| | | | Reinsurance |
| | Real Estate | Real Estate Investment Trusts | Diversified REITs |
| | | | Industrial REITs |
| | | | Mortgage REITs |
| | | | Hotel & Resort REITs |
| | | | Office REITs |
| | | | Health Care REITs |
| | | | Residential REITs |
| | | | Retail REITs |
| | | | Specialized REITs |
| | | Real Estate Management & Development | Diversified Real Estate Activities |
| | | | Real Estate Operating Companies |
| | | | Real Estate Development |
| | | | Real Estate Services |
| Energy | | | |
| Materials | | | |
| Industrials | | | |
| Consumer Discretionary | | | |
| Consumer Staples | | | |
| Health care | | | |
| Information Technology | | | |
| Telecommunication Services | | | |
| Utilities | | | |

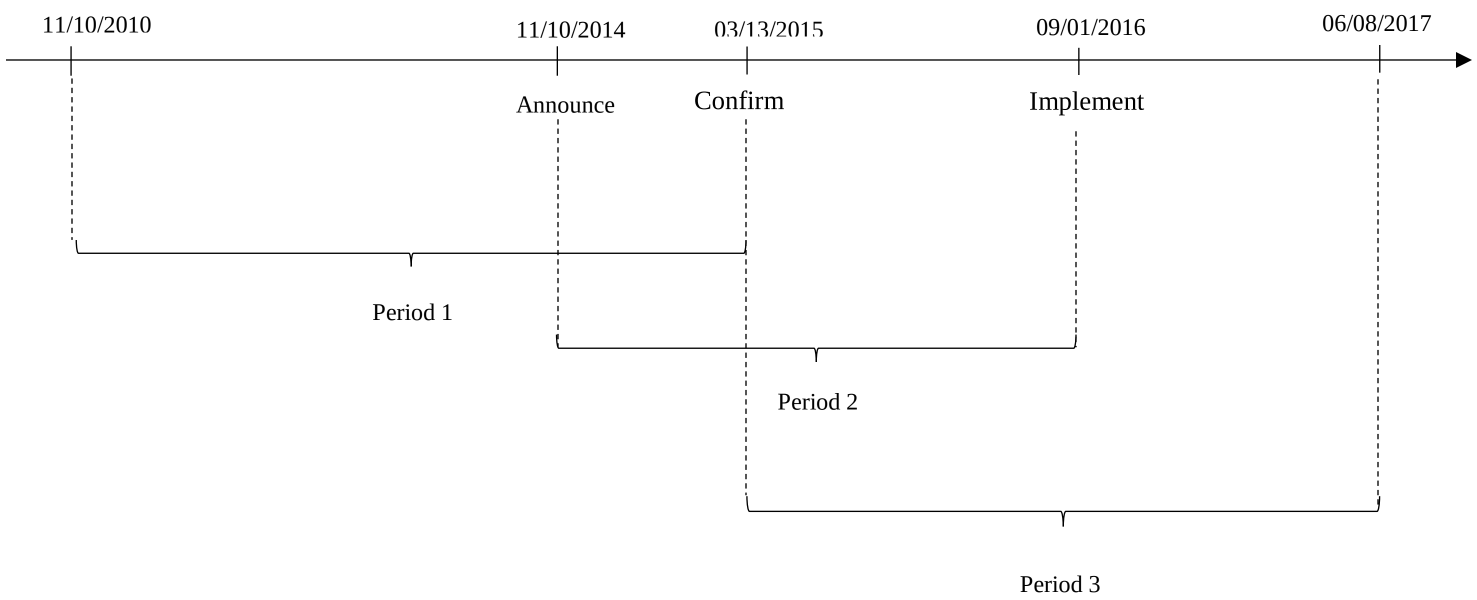
**Panel A. GICS structure effective till 31 Aug 2016**

| **Sector** | **Industry Group** | **Industry** | **Sub-industry** |
|---|---|---|---|
| Financials | Banks | Banks | Diversified Banks |
| | | | Regional Banks |
| | | Thrifts & Mortgage Finance | Thrifts & Mortgage Finance |
| | Diversified Financials | Diversified Financial Services | Other Diversified Financial Services |
| | | | Multi-Sector Holdings |
| | | | Specialized Finance |
| | | Consumer Finance | Consumer Finance |
| | | Capital Markets | Asset Management & Custody Banks |
| | | | Investment Banking & Brokerage |
| | | | Diversified Capital Markets |
| | | | Financial Exchanges & Data |
| | | Mortgage Real Estate Investment Trusts | Mortgage REITs |
| | Insurance | Insurance | Insurance Brokers |
| | | | Life & Health Insurance |
| | | | Multi-line Insurance |
| | | | Property & Casualty Insurance |
| | | | Reinsurance |
| Real Estate | Real Estate | Equity Real Estate Investment Trusts | Diversified REITs |
| | | | Industrial REITs |
| | | | Hotel & Resort REITs |
| | | | Office REITs |
| | | | Health Care REITs |
| | | | Residential REITs |
| | | | Retail REITs |
| | | | Specialized REITs |
| | | Real Estate Management & Development | Diversified Real Estate Activities |
| | | | Real Estate Operating Companies |
| | | | Real Estate Development |
| | | | Real Estate Services |
| Energy | | | |
| Materials | | | |
| Industrials | | | |
| Consumer Discretionary | | | |
| Consumer Staples | | | |
| Healh care | | | |
| Information Technology | | | |
| Telecommunication Services | | | |
| Utilities | | | |

**Panel B. GICS structure effective 1 Sep 2016**

**Figure 3.2. Estimation periods**

This chart explains the time periods used to estimate the baseline model in Table 2.



## Table 3.1. Descriptive statistics

This table presents the descriptive statistics of all variables.

**Panel A: Summary statistics**

| Variable | Mean (%) | S.D. (%) | Min (%) | Max (%) |
|---|---|---|---|---|
| REITs | 0.034 | 1.113 | -8.915 | 9.574 |
| Banks | 0.053 | 1.416 | -9.304 | 6.978 |
| Diversified | 0.044 | 1.427 | -11.568 | 8.503 |
| Insurance | 0.051 | 1.124 | -7.983 | 7.989 |

**Panel B: Pairwise correlation**

| | REITs | Banks | Diversified |
|---|---|---|---|
| *Before Announcement (11/10/2010 - 11/9/2014)* | | | |
| Banks | 0.7332 | 1 | |
| Diversified | 0.7584 | 0.9133 | 1 |
| Insurance | 0.7855 | 0.8914 | 0.9126 |
| *After Announcement (11/10/2014 - 6/8/2017)* | | | |
| Banks | 0.6027 | 1 | |
| Diversified | 0.6797 | 0.9089 | 1 |
| Insurance | 0.7014 | 0.885 | 0.9114 |

## Table 3.2. Change in REIT correlation with Financials

This table presents the estimation results of the following equation:

$$REIT_t = \alpha + \beta_1 Financials_t + \beta_2 Financials_t * Event_t + \gamma Event_t + \varepsilon_t.$$

$REIT_t$ is the daily return of the FTSE NAREIT All Equity REIT Index $Financials_t$ is the daily return of each of the three indexes: S&P 1500 Banks, S&P 1500 Diversified Financials, and S&P 1500 Insurance. $Event_t$ takes the value of 1 if time *t* is after one of the three events: announcement of the new sector (10 Nov 2014), confirmation of the implementation date (13 March 2015), and implementation date (1 Sep 2016). Period 1 covers 11/10/2010 – 03/12/2015; Period 2 covers 11/11/2014 – 08/31/2016; Period 3 covers 03/14/2015 – 06/08/2017 period. Standard errors are clustered at the quarter level.

| | Banks | | | Diversified financials | | | Insurance | | |
|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | Period 1 | Period 2 | Period 3 | Period 1 | Period 2 | Period 3 | Period 1 | Period 2 | Period 3 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| S&P 1500 | 0.609*** | 0.208** | 0.338*** | 0.568*** | 0.339*** | 0.504*** | 0.775*** | 0.312** | 0.581*** |
| | (0.068) | (0.072) | (0.072) | (0.054) | (0.077) | (0.091) | (0.062) | (0.104) | (0.090) |
| Announce (10nov2014) | 0.000 | | | 0.000 | | | 0.000 | | |
| | (0.000) | | | (0.000) | | | (0.000) | | |
| Announce*SP1500 | -0.401*** | | | -0.229** | | | -0.463*** | | |
| | (0.101) | | | (0.094) | | | (0.123) | | |
| Confirm (13mar2015) | | -0.000 | | | -0.000 | | | -0.000 | |
| | | (0.001) | | | (0.000) | | | (0.000) | |
| Confirm*SP1500 | | 0.130 | | | 0.165 | | | 0.269 | |
| | | (0.107) | | | (0.125) | | | (0.146) | |
| Implement (31aug2016) | | | -0.000 | | | -0.001 | | | -0.001 |
| | | | (0.001) | | | (0.001) | | | (0.000) |
| Implement*SP1500 | | | -0.340*** | | | -0.382** | | | -0.307 |
| | | | (0.102) | | | (0.167) | | | (0.196) |
| Constant | 0.000 | 0.000*** | 0.000 | 0.000 | 0.000*** | 0.000 | 0.000 | 0.000*** | 0.000 |
| | (0.000) | (0.000) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Observations | 1,090 | 455 | 564 | 1,090 | 455 | 564 | 1,090 | 455 | 564 |
| Adjusted $R^2$ | 0.514 | 0.212 | 0.169 | 0.553 | 0.326 | 0.261 | 0.591 | 0.340 | 0.287 |

### Table 3.3. Correlation with other sectors

The dependent variables are the daily return of REITs, S&P 1500 Banks, and the S&P 1500 Information Technology in column 1, 2, and 3, respectively. $Announce_t$ takes the value of 1 for all observations after 10 Nov 2014. Standard errors are clustered at the quarter level.

| VARIABLES | REIT | Banks | IT |
|---|---|---|---|
| | (1) | (2) | (3) |
| S&P 1500 Industrials | 0.802*** | 1.030*** | 0.845*** |
| | (0.021) | (0.022) | (0.013) |
| Announce (10 Nov 2014) | 0.000 | 0.000 | 0.000 |
| | (0.000) | (0.000) | (0.001) |
| Announce*SP1500 Industrials | -0.215*** | 0.163*** | 0.073 |
| | (0.039) | (0.041) | (0.067) |
| Constant | -0.000 | -0.000 | 0.000 |
| | (0.000) | (0.000) | (0.000) |
| *Own coefficient minus REIT's coefficient (Announce*SP1500 Industrials):* | | | |
| *Chi-squared* | | *15.921**** | *28.68**** |
| Observations | 1,654 | 1,654 | 1,654 |
| Adjusted $R^2$ | 0.522 | 0.672 | 0.732 |

## Table 3.4. Volatility of daily returns

This table reports the standard deviation of daily returns. "Before" period covers 11/10/2010 – 11/09/2014; "After" period covers 11/11/2014 – 06/08/2017.

| | Observations | REIT | Banks |
|---|---|---|---|
| Before | 1,006 | 0.0442 | 0.0603 |
| After | 648 | 0.0177 | 0.0416 |
| *Difference* | | *-0.0265* | *-0.0187* |
| *F-stat* | | *-1.5224**** | *-1.0904* |

## Table 3.5. Correlation of trading volume

This table presents the estimation results of the following equation:

$$REIT_t = \alpha + \beta_1 Banks_t + \beta_2 Banks_t * Event_t + \gamma Event_t + \varepsilon_t.$$

$REIT_t$ is the change in the log of REITs trading volume. $Banks_t$ is the change in the log of the trading volume of the S&P 1500 Banks Index. $Event_t$ takes the value of 1 if time *t* is after one of the three events: announcement of the new sector (10 Nov 2014), confirmation of the implementation date (13 March 2015), and implementation date (1 Sep 2016). Period 1 covers 11/10/2010 – 03/12/2015; Period 2 covers 11/11/2014 – 08/31/2016; Period 3 covers 03/14/2015 – 06/08/2017 period. Standard errors are clustered at the quarter level.

| VARIABLES | Change in log of daily trading volume | | |
|---|---|---|---|
| | Period 1 | Period 2 | Period 3 |
| Banks' Trading Volume | 0.518*** | 0.557*** | 0.580*** |
| | (0.047) | (0.104) | (0.045) |
| Announce (10 Nov 2014) | -0.004 | | |
| | (0.002) | | |
| Announce*Banks Trading Volume | 0.039 | | |
| | (0.108) | | |
| Confirm (13mar2015) | | 0.004** | |
| | | (0.002) | |
| Confirm*Banks Trading Volume | | 0.023 | |
| | | (0.121) | |
| Effective (31aug2016) | | | -0.000 |
| | | | (0.002) |
| Effective*Banks Trading Volume | | | 0.021 |
| | | | (0.098) |
| Constant | 0.001 | -0.003 | 0.001 |
| | (0.002) | (0.002) | (0.001) |
| Observations | 1,090 | 455 | 562 |
| Adjusted $R^2$ | 0.303 | 0.394 | 0.392 |

## Table 3.6. Volatility of trading volume

This table reports the standard deviation of daily trading volume. “Before” period covers 11/10/2010 – 11/09/2014; “After” period covers 11/11/2014 – 06/08/2017.

| | Observations | REIT | Banks |
|---|---|---|---|
| Before | 1,006 | 18.623 | 12.732 |
| After | 646 | 18.997 | 12.479 |
| *Difference* | | *0.373* | *-0.254* |
| *F-stat* | | *-1.394**** | *-1.770**** |

# Appendix

## Table A1. List of top 10 originators and servicers

| *Panel A. Top 10 originators by number of loans originated* | | | |
|---|---|---|---|
| Originator | Number of loans | Share of sample | Cumulative share |
| New Century | 51,728 | 9.70% | 9.70% |
| Option One | 40,419 | 7.58% | 17.27% |
| Fremont | 36,666 | 6.87% | 24.15% |
| First Franklin | 33,470 | 6.27% | 30.42% |
| BNC | 31,095 | 5.83% | 36.25% |
| WMC Mortgage | 30,032 | 5.63% | 41.88% |
| Argent | 25,948 | 4.86% | 46.75% |
| Wells Fargo | 22,883 | 4.29% | 51.04% |
| Countrywide Home Loans | 16,548 | 3.10% | 54.14% |
| Resmae Mortgage | 11,377 | 2.13% | 56.27% |
| *Panel B. Top 10 servicers by number of loans serviced* | | | |
| Servicer | Number of loans | Share of sample | Cumulative share |
| Ocwen | 149,796 | 28.08% | 28.08% |
| Nationstar | 55,213 | 10.35% | 38.43% |
| Wells Fargo | 53,889 | 10.1% | 48.53% |
| JP Morgan Chase | 48,570 | 9.11% | 57.64% |
| Aurora Loan Services | 24,330 | 4.56% | 62.20% |
| Countrywide Home Loans | 18,416 | 3.45% | 65.65% |
| Litton Loan Servicing | 16,930 | 3.17% | 68.83% |
| Ameriquest | 16,411 | 3.08% | 71.90% |
| Select Portfolio Servicing | 16,395 | 3.07% | 74.98% |
| Homeq | 13,817 | 2.59% | 77.57% |

## Table A2. List of public homebuilders covered in this study

| | Company name | Market cap 2016Q3 ($mil) |
|---|---|---|
| 1 | AV homes, Inc. | 665.8 |
| 2 | Beazer Homes USA, Inc. | 1,717.5 |
| 3 | CalAtlantic Group, Inc. | 7,666.7 |
| 4 | Century Communities, Inc. | 903.0 |
| 5 | Comstock Holding Companies, Inc. | 60.4 |
| 6 | D.R. Horton, Inc. | 14,534.1 |
| 7 | Hovnanian Enterprises, Inc. | 2,272.9 |
| 8 | KB Home | 4,005.7 |
| 9 | Lennar Corporation | 17,402.8 |
| 10 | LGI Homes, Inc. | 1,139.0 |
| 11 | M.D.C. Holdings, Inc. | 2,213.4 |
| 12 | M/I Homes, Inc. | 1,254.6 |
| 13 | Meritage Homes Corporation | 2,528.7 |
| 14 | New Home Company, Inc. | 455.0 |
| 15 | NVR, Inc. | 6,810.3 |
| 16 | PulteGroup, Inc. | 9,948.9 |
| 17 | Taylor Morrison Home Corporation | 5,340.2 |
| 18 | Toll Brothers, Inc. | 8,474.8 |
| 19 | TRI Pointe Group, Inc. | 3,517.0 |
| 20 | UCP, Inc. | 444.8 |
| 21 | WCI Communities, Inc. | 920.3 |
| 22 | William Lyon Homes | 1,841.4 |

# THAO T.T. LE

10 Vairo Blvd Apt 32A, State College, PA 16803
Cell: (814) 321-3833; Email: ttl129@smeal.psu.edu; URL: http://sites.psu.edu/thaole/

## EDUCATION

*Pennsylvania State University*, Ph.D., Business Administration 2017
*National University of Singapore*
- M.Sc., Real Estate & Urban Economics 2012
- B.Sc. (Honors), Real Estate 2009

## PUBLICATIONS

"The Impact of Construction Quality on House Prices", with Joseph Ooi and Nai-Jia Lee. *Journal of Housing Economics*, 2016, vol.26, p. 126-138.

"Explaining House Price Dynamics: Isolating the Role of Non-fundamentals", with David Ling and Joseph Ooi. *Journal of Money, Credit and Banking*, 2015, vol. 47(S1), p.87-125.

"The Spillover Effects of Infill Developments on Local Housing Prices", with Joseph Ooi. *Regional Science and Urban Economics,* 2013, vol. 43, p. 850 – 861.

"Financial Structure of Property Companies and Capital Market Development", with Joseph Ooi. *Journal of Property Investment & Finance,* 2012, vol. 30(6), p. 596 – 611.

"New Supply and Price Dynamics in the Housing Market", with Joseph Ooi. *Urban Studies*, 2011, vol. 49(7), p. 1435-1451.

## CONFERENCE PRESENTATIONS

"House Price Dynamics: Isolating the Role of Non-fundamentals", American Real Estate and Urban Economics Association Annual Conference 2014, Philadelphia, PA.

"Sentiment, Housing Prices and Housing Bubbles", American Real Estate and Urban Economics Association Mid-year Conference, 2011, Washington D.C.

## TEACHING EXPERIENCE

| | |
|---|---|
| Instructor, Contemporary Issues in Real Estate (RM450) - Penn State University | Spring 2015, 2017 |
| Teaching Assistant, Real Estate and Capital Markets (RM/FIN470) - Penn State University | Spring 2014 |
| Teaching Assistant, Contemporary Issues in Real Estate (RM450) - Penn State University | Fall 2013 |

## SCHOLARSHIPS AND AWARDS

*Scholarships*

Peter E. Liberti and Judy D. Olian Scholarship – Penn State University (2016)
Institute of Real Estate Research Scholarship – Penn State University (2016)
Student Travel Grant – Real Estate Research Institute Conference (2015)
Jeanne and Charles Rider Graduate Fellowship - Penn State University (2014, 2015)
Frank P. and Mary Jean Smeal Endowment Fund Scholarship - Penn State University (2013)
Pennsylvania State University Scholarship (2013-2015)
Student Travel Grant – National University of Singapore (2011)

*Research Awards*

Best Paper Award - RE Valuation category, American Real Estate Society Annual Meeting 2013, Kohala Coast, HI.
Best Paper Award - Housing category, American Real Estate Society Annual Meeting 2010, Naples, FL.