T A Begley and A Purnanandam

Design of financial securities: empirical evidence from private-label RMBS deals

Article

Begley, T A and Purnanandam, A

(2016)

Design of financial securities: empirical evidence from private-label RMBS deals.

The Review of Financial Studies, 30 (1). pp. 120-161. ISSN 0893-9454

DOI: https://doi.org/10.1093/rfs/hhw056

Oxford University Press

https://academic.oup.com/rfs/article/30/1/120/2669...

This is a pre-copyedited, author-produced version of an article accepted for publication in Review of Financial Studies following peer review. The version of record: Taylor A. Begley, Amiyatosh Purnanandam; 'Design of financial securities: empirical evidence from private-label RMBS deals', Review of Financial Studies (2017) 30(1) pp.120-161, is available online at:
https://academic.oup.com/rfs/article/30/1/120/2669961/Design-of-Financial-Securities-Empirical-Evidence and at:
https://doi.org/10.1093/rfs/hhw056

Design of Financial Securities: Empirical Evidence from Private-label RMBS Deals*

Taylor Begley† and Amiyatosh Purnanandam‡

March 28, 2016

Review of Financial Studies, forthcoming

Abstract

We study the key drivers of security design in the residential mortgage-backed security (RMBS) market during the run-up to the subprime mortgage crisis. We show that deals with a higher level of equity tranche have a significantly lower delinquency rate conditional on observable loan characteristics. The effect is concentrated within pools with a higher likelihood of asymmetric information between deal sponsors and potential buyers of the securities. Further, securities that are sold from high-equity-tranche deals command higher prices conditional on their credit ratings. Overall, our results show that the goal of security design in this market was not only to exploit regulatory arbitrage, but also to mitigate information frictions that were pervasive in this market.

Keywords: Security design, Mortgage-backed securities, Equity tranche, Subprime mortgage crisis.

JEL Classification: G20, G30.

*We are grateful to Tim Adam, Michael Barr, Sugato Bhattacharyya, Charles Calomiris, John Griffin, Charlie Hadlock, Chris James, Ravi Jagganathan, Pab Jotikasthira, Larry Harris, Jay Hartzell, Han Kim, Andrey Malenko, Gustavo Manso, Gregor Matvos, Stefan Nagel, Greg Nini, Uday Rajan, Amit Seru, Chester Spatt, Alexei Tchistyi, James Vickery, Joao Santos, and Nancy Wallace. We thank seminar participants at Duke, University of Florida, the Federal Reserve Bank of Cleveland, University of Michigan, Northeastern University, NYU, and USC; and conference participants at CAREFIN-Milan, Corporate Finance Conference at Washington University in St. Louis, Federal Reserve Bank of Philadelphia, Financial Intermediation Research Society, NBER Summer Institute, SFS Cavalcade, the USC Doctoral Finance Conference, the Utah Winter Finance Conference, and the Western Finance Association for helpful comments on the paper.

†London Business School; tbegley@london.edu

‡Ross School of Business, University of Michigan; amiyatos@umich.edu

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

1 Introduction

Security design cannot create value in the absence of market frictions. In practice, however, security design is a central feature of many large financial markets. In the market for residential mortgage-backed securities (RMBS), the pooling of many mortgage loans followed by the tranching of their cash flows takes center stage. This suggests that there are important market frictions that create a demand for these securities by some investors. Regulatory policies that constrain or strongly incentivize some investors to hold highly-rated securities is one such key friction. For example, differences in regulatory capital treatment across loans or securities of similar risk can be a key motivation behind the creation of mortgage-backed-securities. This regulation-driven view of security design suggests that securitization is simply a vehicle for creating highly-rated securities out of relatively illiquid mortgage loans to serve some regulated segments of investors. In contrast, information-based theories of security design suggest that securitization can also be used as a tool to mitigate informational frictions that are pervasive in these markets. Was the creation of mortgage-backed-securities in the pre-financial crisis period simply a tool to exploit regulatory distortions, or was it also used as a device to alleviate informational frictions between the buyers and sellers? A clear understanding of this question is important not only for improving our understanding of economic theory behind this market, but also for ongoing policy debates on securitization reforms such as risk retention rules in the RMBS market.[1]

RMBS securities can be broadly classified into three groups: senior debt (the AAA-rated tranche), junior debt (the mezzanine tranche), and the residual equity tranche. Security design is at the very core of the existence of this market. The regulation-based view suggests that the key goal this security design is to maximize the share of the AAA tranche. Thus,

[1] For example, issues surrounding the equity tranche of securitization deals form an important part of the Dodd-Frank Reform Act. In discussing the effects of risk retention requirements pursuant to the Section 946 of the Dodd-Frank Wall Street Reform and Consumer Protection Act, the treasury secretary stresses the importance of this tool in mitigating some contracting frictions and notes that: "...the academic literature on risk retention with respect to asset-backed securitization is limited." Scharfstein and Sunderam (2011) examine some other recent policy proposals and provide suggestions for the more broad reform of the housing finance system.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

it focuses on the division between the AAA tranche and the rest of the deal, and provides no meaningful prediction on how the remaining portion of the deal gets split further into the mezzanine and equity tranches. We take this as the symmetric information benchmark. However, it is well documented that this market is fraught with asymmetric information and conflicts of interest among contracting parties.[2] Potentially, informed sellers can use the design of these securities to convey their private information to uninformed buyers (Leland and Pyle, 1977; DeMarzo, 2005). Specifically, in a separating equilibrium, sellers of pools with relatively better collateral can differentiate themselves from other sellers by creating and retaining a larger residual interest via equity tranche in the pool.[3] Thus, the information-based theories focus on the division of cash flows between the equity tranche and sold tranches (AAA + mezzanine) and predicts a positive relationship between the level of equity tranche in a deal and the quality of the underlying pool. These theories have little to say about the division of cash flows among the sold claims. In this paper, we empirically examine the determinants of tranche structure, and the relationship between the level of equity tranche and pool quality in light of the predictions of these differing motivations of security design.

Gathering information on pool quality, tranching structure, and ex-post performance of the loans in the pool – all of which are needed to conduct our analysis – is a non-trivial exercise involving hand collection of data and manual matching of differing data sources. Thus, we adopt a representative sampling approach and carefully assemble a sample that comprises about 500,000 loans bundled into private-label RMBS deals from 2001-02 and 2005. We combine tranche-level security data with the underlying pool characteristics at the time of RMBS issuance, and track the default performance of each loan in these pools through December 2011. In addition to our main empirical analyses, our study provides some

[2]See, for example, Keys, Piskorski, Seru, and Vig (2012), Gorton and Metrick (2012) for recent surveys; Keys, Mukherjee, Seru, and Vig (2010), Mian and Sufi (2009), Purnanandam (2011), Demyanyk and Van Hemert (2011), He, Qian, and Strahan (2012), Loutskina and Strahan (2011), Acharya, Richardson, et al. (2009) for work related to the subprime mortgage crisis; and Ashcraft and Schuermann (2008) for a detailed analysis of the securitization process.

[3]See Gorton and Pennacchi (1990), Boot and Thakor (1993), Riddiough (1997), DeMarzo and Duffie (1999), DeMarzo (2005), Hartman-Glaser, Piskorski, and Tchistyi (2011), and Chemla and Hennessy (2014) for a rigorous theoretical treatment of this issue.

of the first descriptive statistics on important variables such as the level of equity tranche in this market.

We first examine the cross-sectional determinants of the tranche structure between AAA, mezzanine, and equity tranches. The equity tranche in a deal provides a reasonable approximation to the economic exposure of the sponsors at issuance. Even if the sponsors were to sell these pieces at a later date, it still locks up their capital and therefore imposes carrying costs on them in the interim period.[4] The RMBS market during our sample period experienced a very high rate of growth, lending support to the argument that the sellers had significant opportunity costs of locking up their capital during this period. Moreover, potential buyers of equity tranche are highly-sophisticated investors such as hedge funds and CDO managers who are likely to be better informed than the AAA-tranche clientele. Therefore, sponsors of a poor-quality pool are likely to incur a higher cost if they eventually trade a portion of the equity tranche with these agents at a later date. Hence, even with the possibility of subsequent sale, when sponsors have deals with higher quality underlying pools, they are likely to keep higher fraction of equity tranche just as in the standard separating equilibrium predictions.

In our regression models, we use the percentage of no-documentation loans in a pool as a cross-sectional measure of information asymmetry concerns between the deal sponsors and investors. Since no-documentation loans are not accompanied by verifiable documents such as tax filings of the borrower, it leaves a great degree of discretion with the loan originator, which in turn is likely to increase the adverse selection concerns of the ultimate investors. In addition, we include a number of pool characteristics such as FICO scores, Loan-to-Value (LTV) ratio, and geographical diversity of the collateral as proxies for the observable credit risk of the deal. We show that FICO, LTV, and geographical diversification of the pool are the key determinants of the level of AAA-tranche in a deal. Consistent with Ashcraft, Goldsmith-Pinkham, and Vickery (2010), we also find that deals created in later cohort of

[4] We present a detailed discussion of the construction of the equity tranche and sponsors' economic exposure to it in Section 2.

our sample (2005) have a lower proportion of AAA-rated tranche relative to the earlier cohort (2001-2002). However, the proportion of no-documentation loans in the pool plays no role in explaining the division of a deal between AAA-rated and non-AAA-rated tranches. In contrast, deals with higher proportion of no-documentation loans have significantly higher levels of equity tranche. In fact, FICO, LTV, and geographical diversification do not have any meaningful relationship with the level of the equity tranche in a deal. This finding is consistent with the key idea that investors are likely to have higher adverse selection concerns in relatively opaque deals, which in turn motivates the sponsors to create a larger equity tranche. Broadly, these sets of results provide support for both regulation-based and information-based views of security design: concerns about asymmetric information explain the division between sold and initially unsold (equity) tranches, whereas credit risk concerns better explain the division between AAA-rated and non-AAA-rated tranches. We next turn to our main question that allows us to separate the two hypotheses regarding the role of the equity tranche: does the size of the equity tranche convey the sponsor's private information about the pool quality?

In a purely regulation-based motivation of securitization, we expect no meaningful association between the level of equity tranche at the time of deal origination and the pool's ex-post performance conditional on observable characteristics. In an extreme version of this view, it has also been argued that higher level of equity tranche is simply a result of the sponsor's inability to sell a larger portion of poor-quality pools to the investors. Thus, the regulation-based view predicts zero or a negative association between equity tranche and pool quality. In contrast, the information-based view predicts a positive association between the two.[5] By investigating the relationship between the level of the equity tranche and the future default performance of the pools, we are able to evaluate these sharply different predictions. It is important to note that to tease out the implications of these theories, we are interested

[5]There is also a possibility of pooling equilibrium within the information-based view. Under pooling equilibrium, we should again not find any association between the level of equity tranche and pool quality since sponsors with low pool quality mimic the ones with high pool quality in such equilibrium. We focus on the predictions from a separating equilibrium.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

in relating equity tranche to the "abnormal" default performance – the default performance conditional on the collateral's observed characteristics. Our measure of "abnormal default" below maps nicely to this notion.

We first compute a measure of expected default, defined as the 3-month delinquency in the base model, for every pool in our sample based on observable risk characteristics of the pools and macroeconomic shocks by using a standard default model. The difference between the actual default rate and expected default rate of a pool is labeled as "abnormal default" in the rest of the paper. We relate the level of equity tranche in the pool to its abnormal default rate to test the above predictions. We find that deals with higher proportion of equity tranche have significantly lower abnormal defaults. Further, this effect is concentrated within deals that are relatively opaque (i.e., deals that contain above median no-documentation loans). Said differently, a higher level of equity tranche predicts better performance, and especially for pools with severe adverse selection concerns. In economic terms, pools with above-median level of equity tranche have a 1.8 percentage points lower delinquency rate that cannot be explained away by observable credit risk characteristics and macroeconomic conditions. Compared to the sample average delinquency rate of 28%, this is an economically meaningful effect. In contrast, we find no meaningful relation between the proportion of mezzanine tranche in a deal and future loan performance. Thus, it is not the level of AAA-subordination, but the level of equity tranche in a deal that explains the ex-post default performance of the deal. These results are consistent with the information-based motivation of security design: sponsors create larger equity tranche in deals with positive information on unobservable dimensions.

To further examine the drivers of these results, we repeat the above analysis for a sub-sample of pools where sponsors are likely to have a more significant private information advantage over potential investors. Prior research and economic intuition suggest that sponsors are more likely to possess private information about the underlying loan pools when they are also the loan originator (e.g., see Keys et al., 2010). For these deals, the seller is

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

likely to possess more soft information about the deal such as the underwriting standards at origination as compared to loans that are purchased from other institutions. Consistent with the separating equilibrium prediction, we find that the relationship between the level of the equity tranche and ex-post loan performance is stronger in deals where the sponsors are also the top loan originators, again particularly so for opaque pools. Overall, these results show that the level of the equity tranche best predicts future abnormal default for precisely the deals where information asymmetry between buyers and sellers is likely most severe: deals with more opacity, and deals where the sponsor has a relatively greater information advantage over investors.

We provide further evidence in support of private information content of equity tranche by exploiting the passage of Anti-Predatory Lending laws across several states during our sample period ("APL states"). These laws put stricter requirements on the lenders in terms of their lending practices and disclosure policies which, on the margin, made it more difficult for the lenders to originate poor-quality loans. Such a government regulation should reduce the lemons problem in the market, making the use of private-information-based contracting mechanisms less important. Therefore, prior to the passage of this law, the equity tranche is likely to serve as a more important signal of private information for loans originated in APL states. At the same time, the states that do not pass such laws should experience no systematic change in the relationship between the equity tranche and abnormal default rate. Consistent with this idea, we show that loans originated in APL states in the pre-passage period default at disproportionately lower rate if they are backed by a higher level of equity tranche. Since regulations that drive demand for AAA-rated securities are at the national level and vary only in the time series, this test that exploits variations across states is especially powerful in identifying the information-based role of the equity tranche.

We next turn to the pricing implications of equity tranche. If a higher level of equity tranche conveys the sponsors' positive private information about the pool, then investors should respond to this information by paying a higher price for the sold tranches of the

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

deal. To separate out the mechanical leverage effect of a higher level of equity tranche, we condition our analysis on the credit ratings of sold tranches. We find that sold tranches command higher prices (i.e., lower yield spread) for the same credit rating class if they are backed by a higher level of equity tranche. Again, the effect is concentrated within opaque pools. Together, these results show that opaque pools with a higher level of equity tranche have lower abnormal default rate ex post, and ex ante, they command a higher price. These findings are consistent with the idea that the equity tranche serves as a mechanism to convey the sponsor's private information to potential buyers.

Our study connects to several strands of literature in banking, securitization, and real estate finance. Griffin and Tang (2012) study rating inflation in a large sample of CDOs from 1997 to 2007 and conclude that rating agencies used their subjective assessment to increase the size of AAA-rated tranche beyond the model-implied objective level. Our study is related to Ashcraft et al. (2010) who report a significant decline in RMBS subordination levels between 2005 and mid-2007 and show that the ratings contain useful information about the deal's credit risk. Our results broadly support these findings. Specifically, we also find that ex-ante measures of credit risk explain the variation in the extent of AAA-tranche in the deal. However, they do not investigate the role of equity tranche in conveying the sponsors' private information to the buyers, which is the key focus of our study. Demiroglu and James (2012) show that linkages between syndicate members can result in better ex-post performance of the securitization deals. Hartman-Glaser (2012) studies the effect of seller's reputation capital in these contracts. He et al. (2012) show the influence of large sponsors on credit rating agencies. An, Deng, and Gabriel (2011) study the role of conduit lenders in mitigating informational problems in CMBS deals. Our work also relates to a growing and large literature regarding the conflicts of interest in the securitization market (see Keys et al., 2010; Purnanandam, 2011; He et al., 2012).[6] Unlike these studies, our paper does not study the motivations behind and differences in securitized versus retained loans, or the possibility

[6]See Benmelech, Dlugosz, and Ivashina (2012) on securitization in the case of collateralized loan obligations and Nadauld and Weisbach (2012) for the effect of securitization on the cost of debt.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

of originator moral hazard that comes with securitization.[7] Instead we highlight the effect of informational frictions within the set of securitized deals and the RMBS contract's ability to mitigate some of these frictions.

There has been a renewed interest in the theoretical literature on security design, especially in the RMBS market (see, e.g., Hebert, 2015; Williams, 2015; Hartman-Glaser et al., 2011; Hartman-Glaser, 2012; Chemla and Hennessy, 2014). In addition to providing support for some of the key predictions of extant theories, our results also have important implications for the next generation of theoretical models in the area. First, we highlight the need for developing theoretical models of multiple tranching that can pin down the differences in motivations behind the creation of junior debt claims (mezzanine tranche) and equity tranche.[8] As our empirical results show, these pieces play very different roles in the market. Second, our results show that despite the possibility of the sale of equity tranche in this market, it still served as a mechanism to convey the sponsor's private information. These results call for a need to jointly model the ex-ante tranching decision of the sponsor and the ex-post liquidity of these securities in future theoretical work. We leave some of these topics for future research.

The rest of the paper is organized as follows. Section 2 develops the hypotheses and describes the data. Section 3 presents the results. Section 4 presents robustness tests, and Section 5 concludes the paper.

[7] An originate-to-hold model of lending can be viewed as a limiting case of an RMBS deal where the entire stake is kept by the originating bank. From that perspective, our empirical findings are consistent with the basic idea of this literature: as the sellers stake in the deal increases, the underlying loans perform better in future.

[8] A recent paper by Friewald, Hennessy, and Jankowitsch (2015) develops a theoretical model of security design in which secondary market liquidity plays a key role in obtaining an equilibrium with multiple tranching of the cash flows backed by a security.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

2 Hypothesis Development and Data

2.1 Hypotheses

We want to empirically analyze whether the goal of security design in the pre-financial crisis period was simply to exploit regulatory distortions, or to also alleviate informational frictions. We do so by testing three key predictions of information-based models of security design, and by contrasting them with the alternative view based on regulatory frictions. We adapt the predictions of standard signaling models, such as Leland and Pyle (1977), to our empirical setting. In the model, informed sellers benefit from selling their assets to outside investors either because they have higher risk aversion or they are more impatient compared to the buyers. However, the uninformed buyers face the standard lemon's problem. Either pooling or separating equilibria can emerge in this model. In the pooling equilibrium, sellers of both good and bad assets sell the entire asset at a price that reflects the average quality of the assets in the market. In contrast, in the separating equilibrium, sellers of better assets send a costly signal to the buyers by retaining a higher fraction of the asset's risk with themselves. This allows them to separate themselves from the sellers of bad asset and thus get a better price from the buyers.

A key requirement of the separating equilibrium is the "no-mimicking condition" that ensures that the sellers of bad assets do not find it in their interests to mimic the actions of the good type. Specifically, the good type keeps such a large stake in the asset that the bad type simply does not find it incentive compatible to imitate. In equilibrium, therefore, the buyers are able to infer the private information of the sellers by looking at the extent of risk retained by the seller.

Of course, our empirical setting has several other real world features that are not captured by this simple model. However, the model is rich enough to provide motivation for three main empirical tests for the RMBS market. The first relates the level of equity tranche to concerns about asymmetric information. We hypothesize that the use of equity tranche

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

as a signaling device is higher when there is a larger gap between the sellers' and buyers' information set. For example, if there is no information asymmetry between the two agents, there is no need to incur this cost. On the other hand, when the lemon's problem is severe, the benefit from separation is likely to be higher. Our second hypothesis relates the level of equity tranche to the pool's future performance. This test is a central feature to information-based models with a separating equilibrium. If sponsors are using the level of equity tranche to convey their private information to the market, then we expect to observe better ex-post performance for deals with larger equity tranche. To be precise, we expect to see better ex-post performance of such deals after parsing out the effects of observable risk characteristics, such as FICO score and LTV ratio, of the pool at the time of deal creation. Further, the relationship between equity tranche and future performance is likely to be stronger for deals where information concerns are likely to be severe.[9] Our third hypothesis tests the pricing implication of equity tranche. Reflecting the positive information gleaned from the higher level of equity tranche, we hypothesize that tranches from such deals obtain higher prices from investors.

The regulation-based view suggests that the sponsor created as large a piece of AAA-rated tranche as possible based on the risk factors of the underlying collateral, but it provides no concrete prediction on how the remaining tranche gets split between the mezzanine and equity piece. For expositional simplicity, we contrast the predictions of the two views as we present our findings of the specific tests in the later parts of the paper.

2.2 Data

We construct a novel dataset of RMBS pools and tranches using data from relevant SEC filings and matching them with loan-level data obtained from CoreLogic, a private data vendor. We hand collect the security-level data from the SEC filings to ensure that we do not miss any tranches in a specific deal. It is an important step in computing the level of

[9]In the limiting case of full transparency (symmetric information), signaling is meaningless.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

equity tranche in a deal since commonly available databases such as SDC Platinum database of Thomson Reuters do not provide information on the unsold equity portion of the deal, our main variable of interest. In addition, we hand-collect several important pieces of information such as the proportion of no-documentation loans in a pool and the identity of key players in the securitization chain from the SEC filings that are also not easily available from other sources. Since originators use different labels to classify loans into different documentation classes (for example, "stated documentation," "LITE," and "stated income, stated assets" loans), we carefully read through all the deal prospectuses to ensure consistency in the definition of "NoDoc" loans across deals.

A natural trade-off of this detailed approach to data collection is that we have a relatively smaller sample size as compared to the universe of all deals. Therefore, we take special care in ensuring that our sample is representative. We use a stratified random sampling method to collect private-label RMBS deals covering a wide cross-section of banks and borrowers. We provide detailed description of sample selection criteria and data collection exercise in the Appendix A.1. The Appendix also shows that our sample is comparable to other large sample studies in the literature in terms of key sample characteristics such as size, average FICO scores, and LTV ratios.

Figure 1 presents a schematic diagram of a representative deal and the relevant data sources. Our random sample begins with 234 RMBS deals from 2001-02 and 2005 covering a wide range of sponsors, originators, and servicers. Our main empirical tests are based on a slightly smaller samples that have all the necessary information needed for the analysis. These deals have about 3000 tranches issued against cash flows from over 500,000 loans. The sample is approximately equally balanced between early and late periods (defined as 2001-02 and 2005, respectively). Our sample represents about 14% of the dollar volume of securities issued in the market during the sample period. Thus, we have a representative as well as an economically meaningful sample of deals from the pre-crisis period.

Our loan-level data contain information on characteristics such as FICO scores and LTV

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

ratios at the time of the deal as well as each loan's ex-post performance. In particular, we have information on whether the property entered into delinquency or foreclosure any time from the origination date through December 31, 2011 when our data end.[10] Table 1 presents summary statistics at the loan-, pool-, and tranche-level. We winsorize all variables at 1% from both tails to remove any outlier effects. Based on 509,757 loans that enter our full sample, the average loan's FICO score is 658 with an LTV ratio of 77%, which are in line with other studies of this market during these periods. There is considerable cross-sectional heterogeneity in these two key measures of credit risk across loans. About 67% of the loans are classified as adjustable-rate mortgages (ARM) and 90% of loans are owner occupied residences. Turning to pool-level statistics, the average pool has $775 million in principal amount and is backed by 3,064 loans.

We measure geographical diversification as the complement of one-state concentration of the loan. We first compute the percentage of loans in a pool that comes from each state and then identify the state with maximum share of loans in the pool. Our measure of geographical diversification (*GeoDiverse*) is simply one minus this share.[11] The average pool in our sample has *GeoDiverse* score of 59, representing one-state concentration of 41%. Our sample contains a wide variety of institutional players covering commercial banks, investment banks, and mortgage companies. The full sample contains 26 unique sponsors and 35 unique top originators. We present the list of institutions that are most frequently involved in the deals in our sample in Table A.1 of the Appendix.

The key measure of future performance of these loans is their delinquency status, which takes a value of one if the borrower is behind scheduled payments by over 90 days. 37% of the loans in the sample enter delinquency, and 16% are eventually foreclosed upon. The dollar-weighted average pool-level delinquency and foreclosure rates are 28% and 11%, respectively.[12] We now describe the construction of our key variables that measure information

[10]These CoreLogic-specific data items are available for 162 deals in our sample.

[11]Our results are robust to using alternative measures of geographical diversification such as a Herfindahl index across states and concentration in top-three states (*GeoHerfindahl*), which is presented in later tests.

[12]The default data are available for a slightly lower number of deals because it is based on the sample

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

asymmetry and the level of the equity tranche.

No-documentation loans

Our motivation behind using the proportion of *NoDoc* loans in a pool as a measure of asymmetric information is simple. While all parties are able to observe this proportion, loan originators are likely to have relatively better information about these loans simply by being close to the borrower. This in turn is likely to create higher adverse selection concerns for the buyers of the securities. We obtain the percentage of no-documentation (*NoDoc*) loans in a pool directly from the deal prospectus. No-documentation loans are defined as loans that document neither the income nor the assets of the borrowers. Since different originators label these loans differently, we read through all the deal prospectuses to ensure consistency in our definition across deals. Originators classify these loans under various categories such as "stated documentation," "LITE," and "stated income, stated asset." The prospectus provides further details on the originator-specific underwriting criteria and terminologies, including the details on the various documentation classifications and verification undertaken by the originator. Based on this disclosure, we classify a loan under the no-documentation category if the originator has not verified both the borrower's income and assets.[13] Based on this classification scheme, *NoDoc* loans make up about 17% of all loans in the average pool. There is significant variation in this measure as it ranges from about 0.55% of the pool in the 25th percentile to 34% of the pool in the 75th percentile.

Tranche Measurement

Our main variable of interest is the level of the equity tranche in a deal. We collect this information from the deal prospectuses that provide detailed security-level data on the notional amount of each tranche in the deal, their credit ratings, and the offered yield spread. We

formed by the intersection of our hand-collected data with CoreLogic default data.

[13] We provide an example of the classification of *NoDoc* loans in Appendix A.2.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

combine all tranches that are AAA-rated by at least two rating agencies as the AAA-rated tranche. The equity tranche is the junior-most portion of the deal, which the deal prospectus indicates as not being offered to public. We classify tranches that are rated below AAA but above the equity tranche as the mezzanine tranche.

Table 1 provides descriptive statistics on the tranche structure. Overall, 90.75% of the average deal is tranched into AAA-rated security, while 1.25% of the average deal is in the equity tranche. We find that the size of the average AAA-rated tranche drops from 93.13% in 2001-02 to 88.58% in 2005 (not tabulated). The level of equity tranche more than doubled from 0.74% to 1.71% over the same time period. To give these numbers some perspective, Benmelech and Dlugosz (2009) find that about 71% of CLO pools are rated AAA and 11% are unrated while Stanton and Wallace (2011) find about 84-87% of CMBS pools are rated AAA and 3-4% are unrated equity tranche. Not surprisingly, RMBS tranching structure is closer to the numbers reported by Stanton and Wallace (2011) as compared to the summary statistics of Benmelech and Dlugosz (2009), who include several other types of assets in the pool. Another useful metric of comparison for the equity tranche is the extent of equity capital in a typical bank's overall balance sheet, which is typically in the range of 4-8%.

Sponsors' Economic Exposure to the Equity Tranche

We use the level of equity tranche at the time of security sale as the measure of the sponsor's retained interest in the pool. Ideally, we would like detailed data on the amount of securities retained by the sponsors for a long time after the initial deal creation as a measure of retained interest. Unfortunately, these data are not available due to limited disclosure requirements. In the absence of this proxy, the unsold equity tranche at the time of security sale provides the most natural alternative measure. There are several economic reasons to support the use of equity tranche for our empirical exercise. First, anecdotal evidence suggests that banks often retained part of this exposure on their balance sheet. For example, the Financial Crisis Inquiry Commission's Report presents a case study of an MBS deal issued by Citi

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Bank in 2006 called CMLTI 2006-NC2. They provide details on the identity of the holders of different tranches of this deal (see page 116 of the report) and show that Citi Bank retained the equity tranche along with Capmark Financial Group, a real-estate investment firm. Similarly, Demiroglu and James (2012) provide an example from a deal sponsored by Bear Stearns that shows the sponsor's commitment to initially hold the residual interest: *"The initial owner of the Residual Certificates is expected to be Bear Stearns Securities Corp."*

Second, while it seems sensible to think that the possibility of future sale of equity tranche will render it useless as a signaling device in a standard theoretical model, this is not necessarily true. Even though the sponsors can subsequently offload this risk in the secondary market in the medium to long run, in the immediate aftermath of the deal the risk remains with the sponsor and entails important carrying costs. Indeed, there have been numerous commentaries on the role of warehousing risk in this market during the sub-prime mortgage crisis. Thus, the extent of equity tranche at the time of security sale provides a clean proxy for risk exposure during the initial period. Moreover, typical buyers of the AAA-rated tranche are often regulated institutions such as insurance companies who buy these securities due to regulatory frictions. On the other hand, the typical potential buyer of an equity tranche is a hedge fund or CDO manager, whose primary motivation is to profit from this risky trade. Therefore, the potential future buyers of a share in the equity tranche are likely to be relatively more sophisticated and informed agents as compared to the passive buyers of AAA-tranche. For example, when there is some delay between the deal creation and the potential future sale of the equity tranche, the buyers are able to observe the performance of loans in the pool in the interim period. Incidence of early defaults can provide meaningful new information about the pool quality to the sophisticated buyers of equity tranche that the original hard information alone may not have captured. Second, a key source of information advantage in this market is the better understanding of complex valuation models used to price these securities. Bernardo and Cornell (1997) provide evidence

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

in support of this argument using data from the Collateralized Mortgage Obligations market. Hedge funds are likely to have relatively better understanding of RMBS valuation models and information on crucial inputs such as loan correlations, as compared to relatively passive investors of AAA-rated securities. The stricter market discipline means that sponsors of poor quality pools are likely to incur relatively higher costs when they sell their equity tranche to these buyers. This force, in turn, may support the existence of a separating equilibrium. Hebert (2015) provides further discussion on the importance of such segmentation in RMBS markets.

Finally, as shown by Acharya, Schnabl, and Suarez (2013), there are several instances of securitization where the residual credit risk stayed with the sponsors.[14] Overall, these arguments suggest that equity tranche created at the time of RMBS issuance imposes significant costs on the sponsor consistent with the underlying theoretical assumption of the signaling models. Ultimately, the relationship between the level of the equity tranche and pool quality remains an empirical question that we address in the rest of the paper.

3 Empirical Results

3.1 Cross-Sectional Determinants of Tranche Structure

In our first test, we examine the cross-sectional determinants of tranche structure by relating the fraction of AAA-rated and equity tranches in a deal to the underlying pool characteristics

[14] Finally, we check the annual reports of major sponsors in our sample and find significant retention on their balance sheets. For example, Lehman Brothers had approximately $2 billion of non-investment grade retained interests in residential mortgaged-backed securitization as of November 30, 2006. We obtain similar evidence from the annual reports of Goldman Sachs and Merrill Lynch during this period. For example, Goldman Sachs' 2005 annual report states, "During the years ended November 2005 and November 2004, the firm securitized $92.00 billion and $62.93 billion, respectively, of financial assets, including $65.18 billion and $47.46 billion, respectively, of residential mortgage-backed securities." The report also shows the value of their retained interests in mortgage-backed securities to be $2.928 billion and $1.798 billion, respectively, for those time periods. A back of the envelope calculation suggests that (2.928-1.798)/62.93 = 1.73% was retained during this time period. While this is only a rough approximation, it clearly shows that deal sponsors did retain at least a piece of these securities. A similar computation using information from Merrill Lynch's annual reports gives an estimate of 2.84%.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

using the following model estimated with pool-level data:

$$[\%AAA_p \text{ or } \%Equity_p] = \alpha + \beta(InfoAsym_p) + \Gamma(Credit_p) + \delta(GeoDiverse_p) + \theta(Late_p) + \epsilon_p \tag{1}$$

The dependent variable in the above regression is the proportion of either AAA or equity tranche in a deal as a percentage of the total deal value. Explanatory variables include pool characteristics such as the average credit risk of the borrowers and the concern about information asymmetry. As discussed earlier, we use the percentage of *NoDoc* loans in the pool as the proxy for the extent of asymmetric information ($InfoAsym_p$), or opacity of the underlying pool. $Credit_p$ includes the weighted-average FICO score, the weighted-average LTV ratio, and the fraction of adjustable rate mortgages (ARM) in the pool. FICO and LTV directly measure the credit risk and leverage of the deal, and hence are predictors of future default by the borrower. We include percentage of ARM in the pool as an additional control variable for both credit risk and interest-rate risk of the pool. We include a measure of geographical diversification ($GeoDiverse_p$) of the pool as an additional variable to capture the effect of correlations of loans within the pool. We control for the time effect by including an indicator variable *Late* that equals one for deals from the 2005 cohort, and zero for the earlier 2001-2002 cohort.[15] Inclusion of this variable in the regression model allows us to separate the effect of aggregate macroeconomic shocks such as the level of interest rate and the demand of such securities from investors.

The regulation-based view of securitization predicts that a primary objective of security design is to maximize the AAA-rated tranche in a deal, which can subsequently be sold to investors such as retirement funds and insurance companies that are either explicitly mandated or strongly incentivized by capital charges to only invest in AAA-rated securities. Under this view, adverse selection concerns of the buyers plays no role in security design and thus it has little to say on how the non-AAA rated portion of the deal is further divided

[15]Our results are not sensitive to using even finer time periods such as the month or quarter of the deal.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

into the mezzanine and equity tranche. On the other hand, one of the key predictions of information-based models is that the level of retained interest in the deal should increase with the asymmetric information concerns about the underlying pool. In such deals, debt security buyers are more likely to demand a higher level of equity tranche to mitigate their concerns about adverse selection.

Columns (1)-(3) of Table 2 present the results for the AAA-rated tranche regressions, whereas columns (4)-(6) present results for the equity tranche. Column (1) begins with only *Late* and $\%NoDoc$ as regressors. These estimates confirm that the level of AAA-rated tranche decreased between the two cohorts, and also indicates that deals with more no-documentation loans have a lower level of AAA before controlling for other credit risk characteristics. Once these credit risk factors are included in column (2), the relationship between $\%NoDoc$ and $\%AAA$ disappears. As expected, pool characteristics such as FICO scores, LTV ratio and the fraction of ARM are strongly correlated with the fraction of AAA-rated tranche. Pools with higher geographical diversification also have higher levels of AAA-rated tranches. Column (3) presents the same model with sponsor fixed effects. Our results remain similar. These results suggest that observable credit risk factors of the pool and the extent of geographical diversification of the collateral help explain the AAA tranche in the deal.

Columns (4)-(6) examine the variation in the equity tranche. In column (4), which only includes *Late* as a control variable, we find a positive and significant ($p < 0.01$) coefficient of 0.030 on the $\%NoDoc$ variable. In economic terms, one standard deviation increase in no-documentation loans (17.5 percentage points) is associated with an increase of about 0.53 percentage points, or a 40% (70%) increase in the equity tranche level for the mean (median) deal. The coefficient estimate on *Late* shows that the level of equity tranche increased in later period. In column (5), we include all the control variables and find that the estimate on $\%NoDoc$ is virtually unaffected. In column (6) we include sponsor fixed effects in the model. This specification ensures that our results are not driven by unobserved sponsor characteristics such as its reputation in the market. Our results remain robust

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

to this specification. This result states that within a given sponsor, the deals with more opaque pools have a higher level of equity tranche. Observable credit risk characteristics of the pool such as FICO score and LTV ratio do not explain significant variation in equity tranche across deals. In addition to the slope coefficients, the R^2 of the models provides an interesting insight. For the $\%Equity$ regression, inclusion of observable credit risk variables only improves the model's R^2 from 33% to 35% (columns 4 and 5), whereas the corresponding R^2 improves from 39.6% to 81.4% for the $\%AAA$ regression (columns 1 and 2). Overall, these estimates show that the opacity of the loan pool, and thus asymmetric information concerns, is a key driver of the size of the equity tranche.

Taken together, these results suggest that concerns about private information drive the cross-sectional dispersion in the division between the equity tranche and the sold tranches (mezzanine + AAA), whereas hard pieces of information such as FICO score, LTV ratio, and geographical diversification drive the division between the AAA-rated and non-AAA-rated (mezzanine + equity) tranches. Consistent with the regulation-based view of security design, therefore, issuers were issuing significantly higher fraction of their pool as AAA-rated tranche when observable credit risk of the pool was better. However, the results on the equity tranche cannot be explained by this view alone. Concerns about information asymmetry are harder to price, and as per the information-based view of security design, the level of the equity tranche emerges as an additional contracting tool to mitigate informational frictions. Overall, these results are consistent with both the regulation-based view (the division of AAA and non-AAA) and the information-based view (the division between equity and sold) of security design.

Our inference is robust to computing standard errors clustered at the sponsor level, which we present in robustness tests later. We recognize, however, that it is important to have a sufficiently large number of clusters to obtain consistent standard errors using this method, and our sample has only 26 clusters. As an alternative estimation technique, we estimate a seemingly unrelated regression model for the proportion of AAA, mezzanine, and equity

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

tranche in a deal (not tabulated). Our results are stronger for this specification.

So far, we have focused on the cross-sectional determinants of tranche structure. While these results provide new insights into the motivations behind tranche structure of these deals, they do not yet show that sponsors used the level of equity tranche to convey their private information to the market. We explore this issue in depth in the next section.

3.2 Ex-Post Loan Pool Performance

Does the creation of a larger equity tranche indicate deal sponsors' favorable private information about the underlying loans in the pool? Are these effects mainly concentrated in pools with higher concerns about asymmetric information? We exploit the cross-sectional variation in equity tranche levels along with data on ex-post loan performance to address these questions. If sponsors with favorable private information about the underlying pool create a larger equity tranche, then we expect to observe relatively better ex-post default performance by such pools after conditioning on observable pool characteristics. In other words, we expect *abnormal* default performance of high-equity-tranche pools to be better, where abnormal default performance measures the actual default rate of the pool against a benchmark default rate based on ex-ante observable information. However, if the security design was solely driven by regulation-based motivations, then we should not expect to see any effect of equity tranche on future default performance. Further, if deals with higher levels of equity tranche are simply poorer-quality deals on either observable or unobservable dimensions, then we expect to see *worse*, not better, ex-post default performance of high-equity-tranche deals.

Thus, the information-based view of the equity tranche predicts a negative association between the level of equity tranche and ex-post delinquency or foreclosure rates of loans in the underlying pool, whereas other arguments predict either no relationship or a positive relationship. In light of these opposing empirical predictions, our simple research design

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

can separate the information-based motivation behind equity tranche variation from other competing channels. We first describe our abnormal default performance measure, and then proceed to the empirical results.

Abnormal Default Model

Our goal is to separate the effect of observable loan characteristics from the default performance of the loans in a pool. We do so by first fitting a model of default for the entire sample of loans in our dataset (separately for each cohort), and then taking the residual of the default model as the abnormal default performance of the loan. In the first stage, we fit the following standard model:

$$Pr(default_i = 1) = \frac{e^{(\beta X_i + \gamma X_{i,p})}}{1 + e^{(\beta X_i + \gamma X_{i,p})}} \tag{2}$$

The dependent variable in this logit model takes a value of one if a loan enters into delinquency status, defined as a 3-month delay in payments on the mortgage, any time during our sample period.[16] As an alternative measure of default, we also use the foreclosure status of the loan as the dependent variable, and obtain similar results. We include both loan-level characteristics (denoted by X_i) and pool-level characteristics for the pool that loan i belongs to (denoted by $X_{i,p}$) in the model. X_i includes a comprehensive set of observable borrower characteristics including FICO score, LTV ratio, combined LTV ratio of the loan if there is a second loan on the property, an indicator variable for negative amortization loans, year of origination, indicators for the type of interest rate on the loan (e.g., ARM, balloon, or fixed rate) and the geographical location (i.e., state) of the property. We also include indicators for the loan purpose (e.g., purchase, refinance, refinance-cash out), and the property type (e.g., single family residence). We choose these variables based on economic

[16] Our results remain similar if we use a linear probability model, or if we estimate the default models using the full sample rather than estimating the model separately for each cohort. Our main tests using these alternative models are presented in the Appendix in Table A.3.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

intuition and previous research in the area (see e.g., Demyanyk and Van Hemert, 2011). Pool-level variables include the weighted-average FICO and LTV ratio of the loan, as well as the standard deviation of these measures within the pool. The inclusion of pool-level variables in the model allows us to control for pool specific information that maybe relevant for the loan's default risk.

We estimate the model separately for the early and late cohort pools to allow for differential effect of the covariates on future loan default across these periods. The estimation results of this model are provided in the Appendix in Table A.2, and are similar to other results in this literature. Our aim is to take the predicted values of this regression model as a baseline measure of "normal" or expected default probability of the loan. We compare this predicted default rate ($\widehat{default}_{ip}$) with the actual default rate ($default_{ip}$) in the pool to compute our measure of abnormal default for the entire pool. Specifically, we compute the dollar-weighted abnormal default rate for the pool ($AbDefault_p$) with N_p loans as follows:

$$AbDefault_p = \sum_{i=1}^{N_p} w_i(default_{i,p} - \widehat{default}_{ip})$$

Our measure computes the dollar-weighted difference (weights w_i with $\sum_{i=1}^{N_p} w_i = 1$) in the predicted and actual default rate of the pool. A higher value of abnormal default indicates that the pool experienced a higher-than-expected default rate as compared to the modeled benchmark.[17] When this number is higher, investors of tranches backed by the pool experience larger losses. Hence abnormal default closely measures the economic losses experienced by the investors of securities.

Table 3 provides the distribution of abnormal default for the entire sample, as well as its distribution across pools with relatively high and low levels of no-documentation loans in them (i.e., opaque versus transparent pools). This is presented graphically as a kernel

[17]Alternatively, we compute our measure of abnormal default based on the ratio of actual and predicted default rate and find similar results for our tests. Our measure is dollar weighted to best capture the economic losses borne by the security holders.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

density in Figure 2. The abnormal default (delinquency rate) rate for the entire pool is centered around zero (mean of 0.24% and median of -0.21%) with a standard deviation of 5.47%. Thus, a pool that is one standard deviation higher than the mean has 5.47% excess or abnormal default. The difference ranges from -3.64% to 3.24% as we move from the 25th to the 75th percentile of the distribution. These are large cross-sectional differences in economic terms since the mean dollar-weighted delinquency rate for our entire sample is 28% (Table 1). Our empirical tests explain this cross-sectional variation in abnormal default rate as a function of the level of the equity tranche in the deal.

The summary statistics on abnormal default across opaque and transparent pools (i.e., pools that are above or below the sample median in terms of proportion of no-documentation loans) provides comfort to our basic assumption that no-documentation loans create higher concerns for asymmetric information. As we can see, the *variability* in abnormal default is considerably higher for the opaque pool. The difference in the variance across the two sub-samples is statistically significant at 1%. Said differently, conditional on observable risk metrics, loans in opaque pool have much more volatile outcomes compared to the relatively transparent pool. Thus, investors are likely to be more concerned about the true deal quality for such pools.

Empirical Results

We estimate both the unconditional effect of the level of equity tranche on abnormal default as well as its effect across opaque and transparent pools. We expect equity tranche to have relatively more meaningful impact for pools for which investors have higher concerns about adverse selection. Said differently, when adverse selection concerns are higher, the benefit that the sellers of good pools obtain by separating themselves from the sellers of bad pools is likely to be higher as well. We do so by estimating the following empirical model:

$$AbDefault_p = \beta_0 + \beta_1(Opaque_p) + \beta_2(HighEq_p) + \beta_3(Opaque_p \times HighEq_p) + \epsilon_p \tag{3}$$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

AbDefault is the abnormal default rate of pool *p*. *Opaque* equals one for pools that have above-median percentage of no-documentation loans within each cohort (early or late), and zero otherwise. Similarly, *HighEq* equals one for pools that have above-median level of equity tranche within each cohort, and zero otherwise.[18]

We present the estimation results in Table 4. Columns (1) and (2) present our main results from our parsimonious models. In these models, we only include our main variables of interest, namely *HighEq*, *Opaque*, and their interaction. Column (1) shows that deals with higher equity tranche have significantly lower abnormal default rate and more opaque deals have a higher abnormal default rate. Column (2) shows that the effect of equity tranche on abnormal default is mainly concentrated within deals that are opaque. Unconditionally, deals with above-median equity tranche have 1.84% lower abnormal default, whereas this difference widens to approximately 3.6% (sum of coefficients *HighEq* and *HighEq*Opaque*) within the opaque pools (significant at the 5% level). These results remain stable as we add more control variables in column (3). Recalling that the sample average delinquency rate is 28%, these results show that pools with higher level of equity tranche have both statistically and economically significant lower abnormal default rates, especially within opaque pools.

In column (4), we further separate the effect of AAA-subordination from the effect of the equity tranche on abnormal default. In addition, we include sponsor fixed effects and more pool-level variables in the model to ensure that our results are not explained away by these variations across pools.[19] Specifically, we create a variable *HighMezz* that equals one for deals with above-median level of mezzanine tranche, and include both this variable and its interaction with *Opaque*. While the effect of the equity tranche on abnormal default remains large and significant, we find no evidence that deals with higher levels of mezzanine

[18]Defining variables by cohort is not only economically sensible, but it also ensures that we have a good balance in the number of observations that fall in each of the four buckets based on these two dimensions (i.e., our estimates are not driven by small subset of observations in any one of these four groups).

[19]Note that these pool-level variables have been used to measure the expected default rate in the first stage. Their inclusion in the model is to ensure that our results are not driven by any residual explanatory power left in these variables due to imperfect model fit.

tranche have lower abnormal default. In column (5), we show that our results are robust to clustering standard errors at the sponsor level.

These results show that it is the level of the equity tranche, and not simply the subordination level of AAA tranche, that explains abnormal performance in the future. This provides a clear distinction between the regulation-based and information-based views of the equity tranche. While the regulation-based view predicts either no relationship or a positive relationship between the equity tranche and abnormal default rates, our evidence is consistent with the information-based prediction that a higher level of equity tranche conveys positive information on the unobserved quality of the pools. If the equity tranche contained no information about unobserved deal quality, we would find no results.

We now consider an even stricter measure of equity tranche. Some deals acquire a credit rating for the equity tranche at the time of deal creation along with the other tranches, whereas most of them do not. Unrated portions of the equity tranche are likely to be even more difficult and costly to sell in the future as compared to rated portions and so may send a stronger signal. Based on this idea, we re-estimate our regression model (3) with only the unrated portion of the equity tranche as our main variable in the abnormal default regressions. Columns (1)-(3) of Table 5 present the results. We find similar or stronger evidence of a negative effect of equity tranche on deal performance using this measure.

Columns (4) and (5) of Table 5 provide results based on two alternative measures of abnormal default. In column (4) we present estimation results based on foreclosure rate, rather than delinquency rate, as the measure of default. Our results remain similar. In column (5), we present results where we compute abnormal default as the ratio of actual default to predicted default instead of the difference. Specifically, we construct an abnormal default ratio based on the actual and predicted default rates ($\sum_{i=1}^{N_p} w_i(default_{i,p}/\widehat{default}_{ip})$). Again our results remain similar. Overall, there is a clear pattern in our analysis. Deals with a higher level of equity tranche have better abnormal default performance, and the results are concentrated in opaque deals, where information concerns are likely to be the greatest.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

3.3 A Key Source of the Information Asymmetry Problem

If information asymmetry concerns were indeed a key driver of the results, then the effects should be most prominent in deals where the sponsor is more likely to have the greatest information advantage over potential investors. This notion is consistent with the earlier findings that the effects are stronger in opaque deals. Where do sponsors get private information about the underlying loan quality? As a part of the securitization chain, sponsors are likely to have access to a much more detailed knowledge about and documents from the originators as compared to potential investors. Sponsors are also more likely to have several other informal channels of information exchange with the originators. If the loans are originated by the sponsors themselves, the information advantage over potential RMBS investors is likely even greater. We use this insight as a motivation for our next tests.

We collect the identity of top originators for each pool in our sample from the deal prospectuses. In almost half the cases, deal sponsors are also the top originators of the loan pool. In such cases, sponsor's information advantage over the buyers is likely to be higher, and we expect the equity tranche to play an even more meaningful role. While establishing an important economic channel of private information, these tests further allow us the separate a purely regulation-based view of the equity tranche (which makes no differential prediction on the effect of equity tranche across these sub-samples) from the information-based view.

Table 6 reports the estimation results. We provide estimation results for our entire sample for reference, followed by subsamples based on the sponsor's role in origination. We present our results both with and without the control variables. The results are striking. When the sponsor is also the top originator, we find a large negative effect of equity tranche on future abnormal default rate, where the result is again concentrated within opaque deals. While the estimated coefficient on *Opaque∗HighEq* remains negative for both sub-samples, it is statistically and economically significant only in the subsample where sponsors are also the top originators (columns 3 and 6). Relative to the full-sample estimates in column (4), the

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

economic magnitude of the coefficients of interest increase by a considerable amount for the subsample of pools where sponsors are also the originators in column (6), with the estimate on the interaction term *Opaque∗HighEq* increasing by about 80% from -3.7% to -6.7%.

These results show that when the information advantage of sponsors over investors is likely to be higher, the magnitude of the relationship between the level of the equity tranche and future default rate is substantially larger. This is consistent with view that the equity tranche plays a key role in conveying the sponsor's private information in RMBS deals.

3.4 Identification Using Anti-Predatory Lending Laws

A potential concern with our analysis may be that pools with a higher level of equity tranche are systematically *better* on observable dimensions that we, as econometricians, are unable to control for. If that be the case, we would find lower ex-post default for such pools even without any private information component of the equity tranche. Note that we have already controlled for some of the most important observable loan characteristics such as FICO score, LTV ratio, the nature of interest rate, year of origination, and geographical location of the property in our benchmark default model. Earlier research has shown that these variables explain most of the variation in ex-post default of mortgages. Therefore, it is unlikely that our results are simply an artifact of missing observable characteristics. To further support this claim, we exploit the passage of state-level Anti-Predatory Lending Laws (APLs) as a source of exogenous variation in concerns about lenders' private information.

Several states passed these laws during our sample period to protect homeowners from predatory lending practices. These laws are structured along the lines of Federal Home Ownership and Equity Protection Act (HOEPA), and they typically impose more stringent restrictions on lending practices at the state level as compared to the Federal Act. APLs vary across states in terms of the type of loans they cover and the restrictions they impose on the lenders in terms of required lending practices and information disclosure rules. For example,

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

some of these restrictions include limits on allowable prepayment penalties and balloon payments, borrower counseling requirements, and restriction on mandatory arbitration. Ho and Pennington-Cross (2005, 2006) provide detailed explanations of these laws and the timing of their passage by different states.

The passage of the law is likely to decrease the lenders' ability to originate and package predatory or abusive loans at the margin (see Agarwal, Amromin, Ben-David, Chomsisengphet, and Evanoff, 2012). Such a government regulation should make the use of private contracting mechanisms less important. Therefore, prior to the passage of this law, the equity tranche is likely to serve as a more important signal of private information. Said differently, if the equity tranche indeed conveys private information, then it should have a higher impact in the pre-APL period where there was relatively less government regulation on information disclosure rules. At the same time it is unlikely that regulation-based drivers of security design change in any differential manner across APL and non-APL states around the passage of this law. Indeed, the regulation-driven demand for AAA-rated securities are at the national or international levels, not at the state levels, since these demands are driven by national regulations and so vary purely in the time series.

Ho and Pennington-Cross (2005, 2006) provide an index of the strength of APLs across states as well as the date of the law passage. Their index varies from 4 to 17 with a median score of 10, where a higher index level indicates stronger laws in the state. Based on this measure, we classify all states with index value of 10 or above as the states with strong APL. These states are California, Colorado, Connecticut, Georgia, Illinois, Indiana, Massachusetts, New Jersey, New Mexico, North Carolina, and Washington DC. Of these states, all but Massachusetts and Connecticut, passed their law during our sample period (i.e., between 2002 and 2004), providing us with data on both before and after the law passage. For our test, we create an indicator variable *APL* that takes a value of one for states with strong APL, and zero otherwise. We create an indicator variable *Before* that equals one for loans that belong to states before the passage of law, and zero after that. As

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

in our earlier tests, we create an indicator variable for high equity tranche (*HighEq*) based on the median level of this variable for each cohort in our sample. With *APL*, *HighEq*, and *Before*, we estimate a triple-difference model to estimate the difference in the effect of high equity tranche on future loan delinquency rate for the APL states before and after the passage of the law as compared to the corresponding difference for states without the law. In this estimation strategy, we separate out the unconditional level effects of each one of these variables on the delinquency rate as well as all the double-interaction effects. The coefficient on the triple-interaction term presents us with the estimate of interest.

Because changes in APL laws are at the state-time level, the variation is at the loan level and not the pool level. Thus, we perform this analysis at the loan level. We fit a loan-level linear regression model with delinquency status as the dependent variable. Column (1) of Table 7 presents results. We find a negative and significant coefficient on *APL * HighEq * Before* indicating that equity tranche conveys stronger information about the future loan performance for APL states before the passage of the law. Column (2) includes sponsor fixed effect in the model and shows that our results are similar. In column (3) we use logistic model, instead of the linear probability model, to estimate the regression, and obtain similar results. Finally, column (4) shows that our results are robust to the use of foreclosure status as the measure of default. Overall, these findings are consistent with the level of equity tranche being an indicator of sponsors' favorable private information.

3.5 Pricing Effect of Equity Tranche

An important prediction of signaling models is the presence of a downward sloping demand curve: as sponsors sell more of their assets, investors demand lower prices. Sponsors trade off the resulting liquidity discount from selling more of their assets with the cost of retaining higher equity tranche. Since pricing data for sold tranches is unavailable, following the prior literature we use yield spread on these securities to test this prediction (He et al., 2012). It is relatively straightforward to compute yield spread for floating rate coupons. It

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

is estimated as the spread over LIBOR benchmark reported in the deal prospectus. For the fixed rate tranches, we need to know the duration of these securities to be able to compute the benchmark rate more precisely. Absent this information, we only focus on floating rate tranches for this part of the analysis. Despite this limitation, we are able to cover about 70% of tranches in our sample.

We want to estimate the effect of the equity tranche on the pricing of sold tranches it supports in that same deal. An immediate implication of a higher level of equity tranche is that there is less leverage in the deal. In such deals, more senior tranches that are sold to the investors are safer and therefore they should command attractive prices. This effect is independent of any information revelation via the equity tranche that we are interested in. To separate out the leverage effect, we condition our analysis on the credit rating of sold tranches. We compare the pricing of two similarly rated tranches coming from deals with different levels of equity tranche. The test raises an immediate concern: if the credit rating agencies were able to fully incorporate the effect of equity tranche in their rating, then we should find no effect of equity tranche on tranche yield conditional on credit rating. Our test allows us to isolate the price effects that are driven by information conveyed by the size of the equity tranche beyond any information that the credit rating agencies incorporate into the rating methodologies. Fitch specifically notes in a 2008 report that, "Fitch does not currently make specific adjustments in structured finance rating analysis specifically based around whether or not the risk of the first loss piece has been retained by key transaction parties." Our results are consistent with this assertion. Further, literature is replete with evidence that credit rating agencies did not fully reflect all the information available at the time of deal creation, and perhaps they suffered from incentive problems (see, e.g., Griffin and Tang, 2012; He et al., 2012).

We maintain our basic empirical design that estimates the effect of equity tranche separately across opaque and transparent pools. If the effect of equity tranche on prices come entirely due to the leverage effect, then we should find no difference across opaque and

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

transparent pools. On the other hand, if the effect comes via the revelation of private information, then we expect to see higher prices for tranches backed by higher equity tranche only in the opaque pools. We divide all tranches into broad credit rating classes: AAA, AA, A, and BBB.[20] For deals with multiple tranches within one rating class, we compute a dollar-weighted average yield spread and consolidate them into one observation. This aggregation leads to 549 sold tranches in our sample, out of which 379 are floating rate. We break all pools into two categories based on whether they have above or below the median level of equity tranche. Table 8 presents the cross-tabulation of the average yield spread of sold tranches across high and low equity tranche groups for every credit rating category. There is a clear pattern in the data: within each credit rating class, the yield spread is lower for pools with higher equity tranche. As sponsors sell more of their pool's cash flows to outside investors, the price decreases (yield spread increases).

We estimate a regression model relating yield spread to level of the equity tranche in the deal after controlling for the credit rating fixed effects. Column (1) of Table 9 present our base results. The significant negative coefficient on *HighEq* indicates that after controlling for the credit rating class, high-equity-tranche deals have 25 basis points lower yield spread. More important, the effect comes entirely from the *Opaque* deals as shown in column (4). This is precisely the group where we find a considerably lower abnormal default rate in our earlier tests. Column (7) includes sponsor fixed effect in the model and obtains similar results. We further break our analysis down to AAA-rated and non-AAA-rated securities and report the results across these classes of securities in the remaining columns of the Table. The results show that the effect is concentrated among the non-AAA rated tranche backed by opaque deals (columns 2, 5, and 8), highlighting their inherently higher informational sensitivity. Taken together with the abnormal default rate results, our results show that the level of the equity tranche did contain the sponsor's private information, and cross-sectional variation in at-issuance market prices reflected this. The result is consistent with our information-based

[20]There are a very small number of sold tranches below the BBB rating. We include them in the BBB category.

interpretation of security design.

A potential concern with this analysis is that credit ratings are coarse measures of default risk. Specifically, if tranches backed by a higher level of equity tranche have better creditworthiness on unobserved dimensions within a given credit rating compared to a similarly-rated tranche backed by lower level equity tranche, then we will observe yield differences across such deals even without any signaling. However, we show that our results mainly come from *Opaque* deals, i.e., from deals with higher asymmetric information concerns. Thus, while not conclusive our research design mitigates this concern to a large extent.

Our results are cross-sectional in nature, and they do not directly speak to the issue of the optimality of the level of equity tranche in this market. Needless to say, the level of equity tranche proved inadequate to cover actual losses during the subprime mortgage crisis. However, a simple back-of-the envelop calculation shows that the average level of equity tranche was not insignificant compared to the expected losses. The extent of losses depends on the probability of default (PD) and the losses given default (LGD). We use a very simple framework to assess expected pool losses based on these two dimensions as discussed below.

We first consider a default rate (PD) of 11% that closely matches the realized foreclosure rate in our sample (pre-crisis foreclosure rate in our data is about 3.3%). This can be taken as the perfect foresight scenario for default rate and a very conservative upper bound on default expectations. Prior studies have estimated LGDs in the broad range of 15-25% of the mortgage balance at the time of default (see, e.g., Lekkas, Quigley, and Van Order, 1993; Mason, 2007; Qi and Yang, 2009). We consider a range of recovery rates (1-LGD) from 50% to 100% to be conservative on this dimension. We compute the expected loss by simply multiply the outstanding balance at the time of default by the LGD for each loan that actually defaulted in our sample, and compare that number to the initial principal pool amount. We find the expected loss to be 5.38% of the pool amount based on a recovery rate of 50%. The loss decreases to 1.08% if we assume a very high recovery rate of 90%. We repeat this analysis for a lower default rate assumption as well. Assuming that expected PD

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

was one-fifth of the realized default rate in our sample period (PD=2.2%), we find expected losses to vary between 1.07% and 0.22% of the pool amount as we increase the recovery rate from 50% to 90%. Compared to these numbers, the average equity tranche of 1.2% in our sample is not insignificant. A detailed analysis of the optimality of the level of equity tranche requires a more serious analysis of the beliefs of the investors at the time of deal creation, correlation of loans within the pool, and the proper modeling of recovery rate. We leave these issues for future research work.

4 Robustness Tests

4.1 Alternative Estimation Model

In our base case, we estimate the effect of equity tranche on abnormal default model using a two-step procedure, first by computing the abnormal default measure and then relating it to cross-sectional variation in the level of equity tranche in the second step. The procedure allows us to link the pool-level equity tranche to pool-level default performance. This most closely maps to the economic relationships we have in mind since investors are concerned about losses at the pool level, which is where their security payoffs are based. The pool-level regression also allows us to account for the relative weight of a loan in the pool. Again, this feature helps us develop a measure of loss that is closer in spirit to the economic losses experienced by investors. As a robustness exercise, we also estimate the base model in a one-step framework where we estimate the following models:[21]

$$Pr(default_i = 1) = \frac{e^{(\beta_1 X_i + \gamma_1 X_{ip} + \beta_2 HighEq_{ip} + \beta_3 Opaque_{ip} + \beta_4 HighEq_{ip} \times Opaque_{ip})}}{1 + e^{(\beta_1 X_i + \gamma_1 X_{ip} + \beta_2 HighEq_{ip} + \beta_3 Opaque_{ip} + \beta_4 HighEq_{ip} \times Opaque_{ip})}} \tag{4}$$

X_i and X_{ip} are same as what are used in the earlier default model (2). In addition, we now introduce the pool-level equity tranche and opacity variables directly in the model.

[21] We also run the same test using a linear probability model and find similar results.

Thus the estimated coefficient on *HighEq* measures the abnormal default rate of deals with higher equity tranche.

The results are reported in Table A.4 of the Appendix. Column (2) includes the standard group of controls, and estimates the effect of equity tranche without introducing its interaction with *Opaque* in the model. We find a coefficient of -0.098 that is significant at the 1% level. Column (3) further shows that the effect is concentrated within opaque pools, just as we found in our base case pool-level analysis. The coefficient of -0.171 on the interaction term *Opaque*HighEq* is significant at 1% level. Note that our coefficients are estimated more precisely for the loan-level regressions mainly because we have over 500,000 observations for these regressions. As a point of reference, in column (1) of the Table, we estimate the effect of *HighEq* on loan default without controlling for any loan or pool characteristic such as FICO score or LTV ratio and find a positive and significant coefficient. Contrasting the regression results of columns (1) and (2), it is evident that the relationship between equity tranche and future default mainly comes from the portion of default that is not explained away by observable risk characteristics such as FICO and LTV. The results show that the information content in the size of the equity tranche is all about *abnormal* default rate, and not about the unconditional default rate. This is consistent with the theoretical prediction of signaling models – equity tranche should work as a signal of *unobserved* loan quality.

Column (4) presents the full model including variables relating to the size of the mezzanine tranche, and column (5) presents the corresponding marginal effects evaluated at the mean. The point estimate of interest on the interaction *Opaque*HighEq* is -0.182, with marginal effect at the mean of -0.030 (p-value < 0.01). The marginal effects allow us to directly compare the economic magnitudes of regression coefficients in the loan-level regression with the corresponding effects in the pool-level regressions. In the loan-level regression, we find that loans in high-equity-tranche opaque pools have 3.0% lower abnormal default probability. For the comparable pool-level specification, we find that such pools have 3.7% lower abnormal default (see column (4) of Table 4). Overall, these results show that our

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

estimates are similar, both in statistical and in economic terms, across the two estimation methods.

4.2 Alternative Channels

It has been recognized in the literature that concerns such as sponsor's reputation, servicing rights, and influence over credit rating agencies can play important roles in the way participants contract in this market. These considerations could potentially interact with the structuring of RMBS design. While we do not explore these interactions in detail, this section presents several tests to establish the robustness of our analysis even in the presence of these competing influences. We first consider the possibility that our results are driven by deals where sponsors and originators have more "skin in the game" by holding servicing contracts (e.g., Piskorski, Seru, and Vig, 2010; Demiroglu and James, 2012). In addition to earning fees from the origination of loans, lenders sometimes retain servicing rights on loans that provide them with an additional stream of income for the life of the loan. If the sponsors hold servicing rights on the loans, this implicit equity stake may provide stronger incentives for them to ensure that the pool is populated with higher quality loans. If deals with higher servicing "skin in the game" coincide with those with higher equity tranche, then our inferences maybe contaminated. To empirically separate out this alternative channel, we collect data on the identity of primary servicer for the loans in the pool. We create a dummy variable that indicates if the sponsor is also the servicer (*SellAndService*) and a dummy variable that indicates if the top originator for the pool is also the servicer (*TopOrigAndService*).[22]

Another mechanism that can potentially confound our results is the reputational concerns of the members of the syndicate (Hartman-Glaser, 2012). As shown earlier, our results are not materially affected by the inclusion of sponsor fixed effects in the estimation exercise. This ensures that we are able to separate out time-invariant reputational effect of the sponsors.

[22] We perform the same tests using a dummy variable that indicates if the servicer is any of the top four originators and get similar results.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Since we consider a short time-period (2001-2005) for our analysis, it is reasonable to assume that a sponsor's reputation remained practically constant during the sample period. As an alternative test in a similar spirit, we consider the heterogeneity in the sponsor-type to control for the reputational concerns. We expect that long-lived and established commercial banks such as JP Morgan Chase have different concerns about protecting their franchise values as compared to specialized mortgage originating institutions such as Ameriquest. Also, large commercial and investment banks may be able to exert more influence over the credit rating agencies to receive inflated ratings relative to smaller stand-alone mortgage lenders (He et al., 2012). To address these issues, we classify each sponsor as a commercial bank, investment bank, savings and loan institution, or mortgage lender and then include dummy variables for these categories in the regression model. These tests do not affect our main results, and are subsumed by our specifications using sponsor fixed effects, so we do not tabulate the results.

Table 10 reproduces the main results from earlier sections of the paper alongside a specification that includes the variables mentioned above as well as *GeoHerfindahl*, which is a Herfindahl index across states and concentration in top-three states as alternative measure of geographical diversification. For each specification, we also include sponsor fixed effects and cluster standard errors at the sponsor level. All of our results remain qualitatively similar. Among the additional control variables, we do find some effect consistent with "skin in the game" hypothesis as deals where the top originator is also the servicer have better ex-post performance (column 6). However, inclusion of this control variable does not substantially affect any of our results. Overall, our results are unlikely to be affected by these alternative channels.

5 Conclusion

This paper empirically examines the motivations behind security design in residential mortgage-backed securities during the run-up to the subprime mortgage crisis. We show that the goal

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

of securitization was not only to exploit incentives generated by the preferential regulatory treatment of AAA-rated tranches, but also to mitigate information frictions that were pervasive in this market. Specifically, we document that the level of the equity tranche conveys the sponsor's private information, particularly in deals with severe adverse selection concerns. Further, investors responded to this signal by paying higher prices for deals backed by higher equity tranche. These pieces of evidence provide support for some of the fundamental predictions of security design models based on asymmetric information (e.g., Leland and Pyle, 1977; DeMarzo, 2005).

Our findings show that the design of mortgage-backed securities was able to mitigate some of the contracting frictions as predicted by extant theoretical models in the literature. By design, our study is cross-sectional in nature. Therefore, we are able to comment on the ability of equity tranche in explaining economic outcomes only in a relative sense. Our study does not rule out the possibility that the absolute level of equity tranche supporting these deals was too low during the sample period. Indeed, Stanton and Wallace (2011) show that in the period leading up to the crisis, the rating agencies allowed subordination levels in CMBS markets to fall to suboptimal levels. The key contribution of our paper is to show that cross-sectional pattern in securitization design does follow the predictions of asymmetric information models. This finding has important implications for the development of future theoretical models in this area as well as for informing policy debates surrounding this market.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

References

Acharya, Viral, Matthew Richardson, et al., 2009, *Restoring Financial Stability: How to Repair a Failed System*, volume 542 (Wiley).

Acharya, Viral V., Philipp Schnabl, and Gustavo Suarez, 2013, Securitization without risk transfer, *Journal of Financial Economics* 107, 515 – 536.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas Evanoff, 2012, Predatory lending and the subprime crisis, *Working Paper* .

An, Xudong, Yongheng Deng, and Stuart A Gabriel, 2011, Asymmetric information, adverse selection, and the pricing of CMBS, *Journal of Financial Economics* 100, 304–325.

Ashcraft, Adam, Paul Goldsmith-Pinkham, and James Vickery, 2010, *MBS Ratings and the Mortgage Credit Boom* (DIANE Publishing).

Ashcraft, Adam B, and Til Schuermann, 2008, *Understanding the securitization of subprime mortgage credit* (Now Pub).

Benmelech, Efraim, and Jennifer Dlugosz, 2009, The alchemy of CDO credit ratings, *Journal of Monetary Economics* 56, 617–634.

Benmelech, Efraim, Jennifer Dlugosz, and Victoria Ivashina, 2012, Securitization without adverse selection: The case of CLOs, *Journal of Financial Economics* .

Boot, Arnoud WA, and Anjan V Thakor, 1993, Security design, *The Journal of Finance* 48, 1349–1378.

Chemla, Gilles, and Christopher A Hennessy, 2014, Skin in the game and moral hazard, *The Journal of Finance* 69, 1597–1641.

DeMarzo, Peter, and Darrell Duffie, 1999, A liquidity-based model of security design, *Econometrica* 67, 65–99.

DeMarzo, Peter M, 2005, The pooling and tranching of securities: A model of informed intermediation, *Review of Financial Studies* 18, 1–35.

Demiroglu, Cem, and Christopher James, 2012, How Important is Having Skin in the Game? Originator-Sponsor Affiliation and Losses on Mortgage-backed Securities, *Review of Financial Studies* 25, 3217–3258.

Demyanyk, Yuliya, and Otto Van Hemert, 2011, Understanding the subprime mortgage crisis, *Review of Financial Studies* 24, 1848–1880.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Friewald, Nils, Christopher Hennessy, and Rainer Jankowitsch, 2015, Secondary market liquidity and security design: Theory and evidence from abs markets, *Working Paper* .

Gorton, Gary, and Andrew Metrick, 2012, Securitization, Working Paper 18611, National Bureau of Economic Research.

Gorton, Gary, and George Pennacchi, 1990, Financial intermediaries and liquidity creation, *Journal of Finance* 49–71.

Griffin, John M, and Dragon Yongjun Tang, 2012, Did subjectivity play a role in CDO credit ratings?, *The Journal of Finance* 67, 1293–1328.

Hartman-Glaser, Barney, 2012, Reputation and signaling, *Working paper* .

Hartman-Glaser, Barney, Tomasz Piskorski, and Alexei Tchistyi, 2011, Optimal securitization with moral hazard, *Journal of Financial Economics* 104, 186–202.

He, Jie, Jun Qian, and Philip Strahan, 2012, Are all ratings created equal? the impact of issuer size on the pricing of mortgage-backed-securities, *Journal of Finance* .

Hebert, Benjamin, 2015, Moral hazard and the optimality of debt, *Working Paper* Stanford University.

Ho, Giang, and Anthony Pennington-Cross, 2005, The impact of local predatory lending laws on the flow of subprime credit, *Federal Reserve Bank of St. Louis Working Paper No. 2005-049B. Available at: http://research.stlouisfed.org/wp/2005/2005-049.pdf* .

Ho, Giang, and Anthony Pennington-Cross, 2006, The impact of local predatory lending laws on the flow of subprime credit, *Journal of Urban Economics* 60, 210–228.

Keys, Benjamin, Tomasz Piskorski, Amit Seru, and Vikrant Vig, 2012, Mortgage financing in the housing boom and bust, in *Housing and the Financial Crisis* (University of Chicago Press).

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? evidence from subprime loans, *The Quarterly Journal of Economics* 125, 307–362.

Lekkas, V, JM Quigley, and R Van Order, 1993, Loan loss severity and optimal mortgage default, *Journal of the American Real Estate & Urban Economics Association* 21, 353–371.

Leland, Hayne Ellis, and David H Pyle, 1977, Informational asymmetries, financial structure, and financial intermediation, *Journal of Finance* 32, 371–87.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Loutskina, Elena, and Philip E Strahan, 2011, Informed and uninformed investment in housing: The downside of diversification, *Review of Financial Studies* 24, 1447–1480.

Mason, Joseph R, 2007, Mortgage loan modification: Promises and pitfalls, *Available at SSRN 1027470* .

Mian, Atif, and Amir Sufi, 2009, The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis, *The Quarterly Journal of Economics* 124, 1449–1496.

Nadauld, Taylor D, and Michael S Weisbach, 2012, Did securitization affect the cost of corporate debt?, *Journal of Financial Economics* 105, 332–352.

Piskorski, Tomasz, Amit Seru, and Vikrant Vig, 2010, Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis, *Journal of Financial Economics* 97, 369–397.

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

Qi, Min, and Xiaolong Yang, 2009, Loss given default of high loan-to-value residential mortgages, *Journal of Banking & Finance* 33, 788–799.

Riddiough, Timothy J, 1997, Optimal design and governance of asset-backed securities, *Journal of Financial Intermediation* 6, 121–152.

Scharfstein, David, and Adi Sunderam, 2011, The economics of housing finance reform, in Martin Neil Baily, ed., *The Future of Housing Finance*, 146–198 (Brookings Institution Press).

Stanton, Richard, and Nancy Wallace, 2011, CMBS subordination, ratings inflation, and regulatory-capital arbitrage, *Working paper, UC Berkeley* .

Williams, Basil, 2015, Search, liquidity, and retention: Signaling multidimensional private information, *Working paper* Duke University.

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Figures

Figure 1: **Example Deal: Fremont Home Loan Trust Series 2002-1**

This figure provides an example deal from our sample to illustrate the construction of a typical deal and the sources of our data. Loan specific characteristics such as FICO score, loan amount, loan type, LTV, etc. are from CoreLogic. Aggregate deal statistics, including the tranche structuring of the deal, were hand collected from the Form 424(b)(5) filings to the SEC.

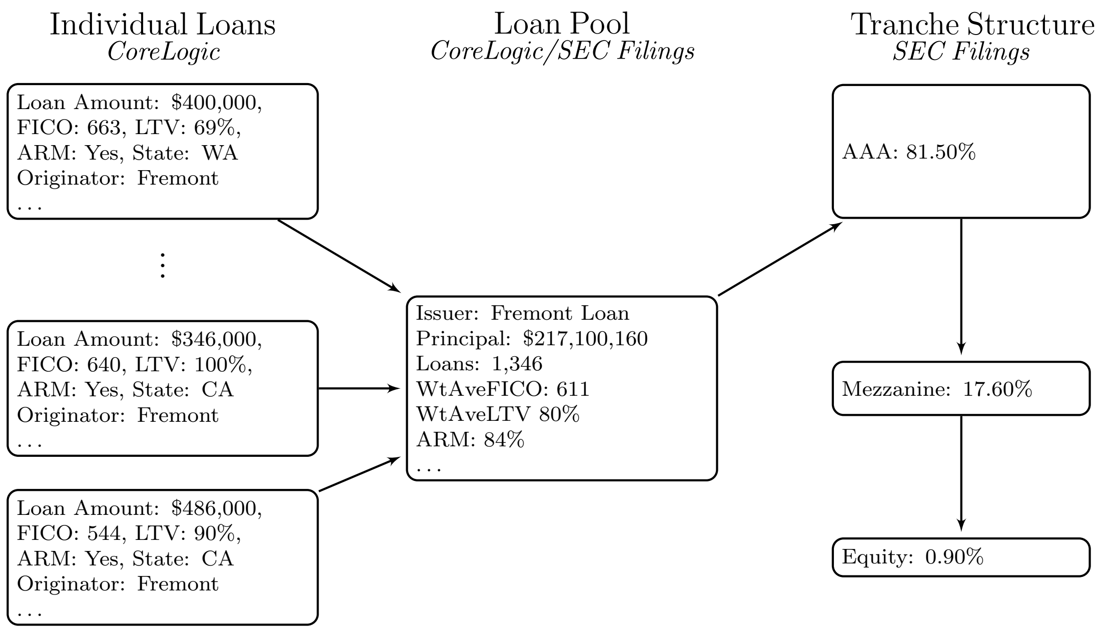
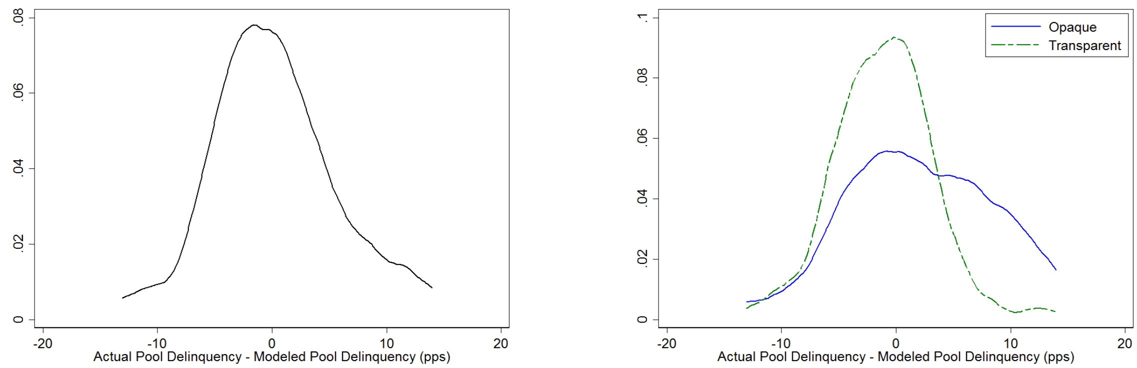


Figure 2: **Distribution of Abnormal Default**

This figure presents kernel densities of our measures of abnormal default. Panel 2a presents our primary measure of abnormal default, which we calculate as the difference between the actual ex-post pool delinquency rate minus the model-predicted delinquency rate. Panel 2b presents the same measure separately by opaque (deals with above-median number of no-documentation loans) and transparent pools.



(a) Full Sample

(b) Opaque versus Transparent

Main Tables

Table 1: **Full Sample Summary Statistics**

This table presents summary statistics for our sample. The loan-level characteristics and ex-post default data are based on the sample of loans from CoreLogic. Pool-level characteristics including tranche structure are from the random sampling of SEC filings. We present pool-level default rates on a dollar-weighted basis.

	Mean	Std Dev	Min	25%	50%	75%	Max	N
Loan-Level Characteristics								
Loan Amount	266278.05	210186.99	4950.00	112200.00	202500.00	375000.00	4350000.00	509757
FICO	657.98	78.05	465.00	601.00	659.00	719.00	800.00	509757
LTV	77.05	13.63	31.05	71.43	80.00	85.00	100.00	509757
CLTV (if 2nd lien present)	95.81	8.55	9.68	95.00	100.00	100.00	100.00	96566
ARM	0.67	1.41	0.00	0.00	1.00	1.00	100.00	509752
Single Family Residence	0.76	0.43	0.00	1.00	1.00	1.00	1.00	509757
Owner Occupied	0.90	0.30	0.00	1.00	1.00	1.00	1.00	509757
Negative Amortization	0.05	0.21	0.00	0.00	0.00	0.00	1.00	509757
Future Loan Delinquency	0.37	0.48	0.00	0.00	0.00	1.00	1.00	509757
Future Loan Foreclosure	0.16	0.37	0.00	0.00	0.00	0.00	1.00	509757
Pool-Level Characteristics								
Principal Pool Amount	775.23	505.17	165.97	402.52	655.21	977.38	2633.28	234
Number of Loans	3064.23	2503.26	340.00	1346.00	2183.50	3987.00	12202.00	234
Late (Year=2005)	0.52	0.50	0.00	0.00	1.00	1.00	1.00	234
NoDoc	16.87	17.54	0.00	0.55	11.05	33.71	62.35	224
FICO	689.35	50.89	586.00	631.00	713.86	732.93	747.00	230
LTV	74.21	6.23	58.26	69.38	74.84	79.00	90.50	230
ARM	58.34	43.01	0.00	0.00	78.61	100.00	100.00	229
GeoDiverse	59.30	17.59	16.17	49.51	60.56	74.27	86.30	234
GeoHerfindahl	20.94	16.29	2.33	8.41	16.57	26.49	100.00	234
Future Pool Delinquency	0.28	0.17	0.02	0.10	0.28	0.40	0.67	162
Future Pool Foreclosure	0.11	0.10	0.00	0.01	0.08	0.18	0.41	162
Tranche Structure:								
% AAA Tranche	90.75	7.14	75.55	83.11	94.00	96.68	98.57	234
% Mezzanine Tranche	8.00	6.56	0.00	2.61	5.00	15.25	20.96	234
% Equity Tranche	1.25	1.27	0.00	0.50	0.75	1.75	6.27	234

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 2: **Cross-Sectional Determinants of Deal Structure**

This table presents OLS estimates from regressions of *%AAA Tranche* (columns (1)-(3)) and *%Equity Tranche* (columns (4)-(6)) on loan-pool characteristics. *%AAA Tranche* is the percent of the principal pool amount that is AAA-rated, *%Equity Tranche* is the percent of the principal pool amount that is not publicly offered, *Late* is a dummy variable equal to 1 for deals from 2005, *% NoDoc* is the percent of the loan pool with no documentation, *FICO* is the pool's weighted average FICO score, *LTV* is the pool's weighted average loan-to-value ratio, *% ARM* is the percent of the loan pool with adjustable rate mortgage loans, *GeoDiverse* measures the geographic diversity and is 100 - (percent of largest one state origination concentration) in the mortgage pool. All standard errors are heteroskedasticity robust.

	%AAA			%Equity		
	(1)	(2)	(3)	(4)	(5)	(6)
Late	-3.372***	-3.453***	-3.305***	0.832***	0.845***	0.738***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)
% NoDoc	-0.221***	-0.019	-0.024	0.030***	0.028***	0.024***
	(<0.01)	(0.22)	(0.21)	(<0.01)	(<0.01)	(0.01)
FICO		0.099***	0.094***		-0.002	0.002
		(<0.01)	(<0.01)		(0.40)	(0.64)
LTV		-0.276***	-0.276***		-0.011	0.017
		(<0.01)	(<0.01)		(0.57)	(0.44)
% ARM		0.013**	0.010*		0.003*	0.004***
		(0.02)	(0.06)		(0.05)	(0.01)
GeoDiverse		0.058***	0.048***		-0.003	0.001
		(<0.01)	(<0.01)		(0.52)	(0.83)
Constant	96.158***	40.729***	42.597***	0.318***	2.660	-1.378
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(0.32)	(0.68)
Sponsor FE	No	No	Yes	No	No	Yes
Observations	215	215	215	215	215	215
R^2	0.396	0.814	0.829	0.330	0.350	0.503

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 3: **Abnormal Default**

This table presents descriptive statistics for the pool-level abnormal delinquency rate (*AbDefault*) winsorized at the 1% level. This variable is the actual pool-level default rate minus the model-predicted default rate, and is measured in percentage points (see Section 3.2 for further details). In addition to statistics for the full sample, this table presents statistics across *Opaque* (above-median no-documentation loans with each cohort) and transparent pools.

	Mean	Std Dev	Min	10%	25%	50%	75%	90%	max	N
Transparent	-1.04	4.34	-13.04	-5.59	-4.04	-0.93	1.42	3.79	13.96	96
Opaque	2.11	6.38	-13.04	-5.65	-2.41	2.38	7.05	11.02	13.96	66
Full Sample	0.24	5.47	-13.04	-5.59	-3.64	-0.21	3.24	7.31	13.96	162

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 4: **Ex-Post Outcomes: Abnormal Default**

This table presents OLS estimates from regressions of *AbDefalut* on loan-pool characteristics, where *AbDefault* is the actual pool-level delinquency rate minus the model-predicted default rate, and is measured in percentage points (see Section 3.2 for further details). *Opaque* is a dummy variable equal to 1 for deals with *%NoDoc* greater than that of the median deal in its cohort, *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal in its cohort, *HighMezz* is a dummy variable equal to 1 for deals with *%Mezzanine Tranche* greater than that of the median deal in its cohort, *Late* is a dummy variable equal to 1 for deals from 2005, *FICO* is the pool's weighted average FICO score, and *LTV* is the pool's weighted average loan-to-value ratio. All standard errors are heteroskedasticity robust, and clustered at the sponsor when noted at the bottom of the table.

	(1) AbDefault	(2) AbDefault	(3) AbDefault	(4) AbDefault	(5) AbDefault
Late			1.321	1.958**	1.958*
			(0.15)	(0.04)	(0.07)
WtAveFICO			0.008	-0.013	-0.013
			(0.53)	(0.51)	(0.44)
WtAveLTV			-0.023	-0.130	-0.130
			(0.85)	(0.33)	(0.32)
Opaque	3.660***	5.043***	5.220***	4.426***	4.426***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(0.01)
HighEq	-1.844**	-0.312	0.181	2.430**	2.430**
	(0.03)	(0.70)	(0.83)	(0.03)	(0.04)
Opaque*HighEq		-3.306*	-3.282*	-3.698*	-3.698*
		(0.05)	(0.06)	(0.05)	(0.05)
HighMezz				-1.617	-1.617
				(0.33)	(0.40)
Opaque*HighMezz				1.111	1.111
				(0.60)	(0.69)
Sponsor FE	No	No	No	Yes	Yes
Observations	162	162	162	162	162
R^2	0.106	0.126	0.142	0.385	0.385
SE Type	robust	robust	robust	robust	Sponsor

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 5: **Ex-Post Outcomes: Abnormal Default – Alternative Measures**

This table presents OLS estimates from regressions of *AbDefault*, measured in various ways (see Section 3.2 for further details), on loan pool characteristics. *AbDefault* (columns 1-3) is the actual pool-level delinquency rate minus the model-predicted delinquency rate, and is measured in percentage points. *AbForeclosure* (column 4) computed in the same way, only using foreclosure in the place of delinquency. *AbDelinqRatio* (column 5) is the actual pool-level delinquency rate divided the model-predicted delinquency rate. *Opaque* is a dummy variable equal to 1 for deals with *%NoDoc* greater than that of the median deal in its cohort, *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal in its cohort, *HighMezz* is a dummy variable equal to 1 for deals with *%Mezzanine Tranche* greater than that of the median deal in its cohort, *Late* is a dummy variable equal to 1 for deals from 2005, *FICO* is the pool's weighted average FICO score, and *LTV* is the pool's weighted average loan-to-value ratio. Variables that are noted with *Unrated (UR)* represent similar measures to those above, only with the *%Equity Tranche* measured as the portion of the deal that was not rated by a credit rating agency. All standard errors are heteroskedasticity robust, and clustered at the sponsor when noted at the bottom of the table.

	(1) AbDefault	(2) AbDefault	(3) AbDefault	(4) AbForeclosure	(5) AbDelinqRatio
Opaque	3.416***	4.617***	4.573**	1.910***	0.154***
	(<0.01)	(<0.01)	(0.01)	(<0.01)	(<0.01)
HighUnRated	-1.668**	-0.336	2.670**		
	(0.05)	(0.69)	(0.03)		
Opaque*HighUnRated		-3.002*	-4.475**		
		(0.09)	(0.02)		
HighMezz_UR			-1.734		
			(0.40)		
Opaque*HighMezz_UR			1.082		
			(0.67)		
HighEq				0.969	0.078*
				(0.24)	(0.06)
Opaque*HighEq				-2.498**	-0.134**
				(0.02)	(0.02)
HighMezz				-0.413	-0.025
				(0.82)	(0.60)
Opaque*HighMezz				1.203	0.020
				(0.47)	(0.82)
Controls	No	No	Yes	Yes	Yes
Sponsor FE	No	No	Yes	Yes	Yes
Observations	162	162	162	162	162
R^2	0.102	0.119	0.397	0.223	0.384
SE Type	robust	robust	Sponsor	Sponsor	Sponsor

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: **Ex-Post Outcomes: The Channel of Private Information**

This table presents OLS estimates from regressions of *AbDefault* on loan pool characteristics. *AbDefault* is the actual pool-level delinquency rate minus the model-predicted delinquency rate, and is measured in percentage points (see Section 3.2 for further details). *Opaque* is a dummy variable equal to 1 for deals with *%NoDoc* greater than that of the median deal in its cohort, *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal in its cohort, *HighMezz* is a dummy variable equal to 1 for deals with *%Mezzanine Tranche* greater than that of the median deal in its cohort, *Late* is a dummy variable equal to 1 for deals from 2005, *FICO* is the pool's weighted average FICO score, and *LTV* is the pool's weighted average loan-to-value ratio. *Sponsor is Top Originator* indicates deals where the deal sponsor originated more loans in pool than any other originator. All standard errors are heteroskedasticity robust.

		Sponsor is Top Originator			Sponsor is Top Originator	
	All	No	Yes	All	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)
Late				1.958**	3.532**	3.789**
				(0.039)	(0.040)	(0.010)
WtAveFICO				-0.013	-0.057*	0.025
				(0.512)	(0.079)	(0.313)
WtAveLTV				-0.130	-0.064	-0.320
				(0.326)	(0.733)	(0.101)
Opaque	5.043***	4.788***	4.250**	4.426***	4.285**	2.214
	(<0.001)	(0.002)	(0.026)	(0.001)	(0.021)	(0.304)
HighEq	-0.312	-0.846	-0.023	2.430**	0.620	3.687***
	(0.698)	(0.561)	(0.981)	(0.027)	(0.724)	(0.002)
Opaque*HighEq	-3.306*	-1.700	-5.738**	-3.698*	-1.011	-6.687***
	(0.053)	(0.457)	(0.032)	(0.055)	(0.707)	(0.007)
HighMezz				-1.617	-5.183**	3.097**
				(0.332)	(0.040)	(0.038)
Opaque*HighMezz				1.111	-0.649	5.698*
				(0.598)	(0.797)	(0.092)
Sponsor FE	No	No	No	Yes	Yes	Yes
Observations	162	72	90	162	72	90
R^2	0.126	0.147	0.099	0.385	0.492	0.427

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 7: **Anti-Predatory Lending Laws, Equity Tranche, and Ex-Post Outcomes**
This table presents loan-level estimates from regressions of *LoanDefault* on loan, pool, and state characteristics using OLS (columns 1, 2, and 4) and logit (column 3). Columns (1-3) measure *LoanDefault* as future delinquency status (*Delinquency*), and column (4) used future foreclosure status (*Foreclosure*). *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal, *APL* is a dummy variable equal to 1 for loans from states that enact anti-predatory lending laws, *Before* is a dummy variable equal to 1 for the time period prior to the passage of APL laws. All regressions include the loan and pool control variables including FICO, LTV, Combined LTV, Negative Amortization, property type, loan purpose, year of origination, and state. Standard errors are computed as heteroskedasticity-robust (parenthesis) and clustered at the Sponsor level [brackets].

	(1) Delinquency	(2) Delinquency	(3) Delinquency	(4) Foreclosure
APL	0.039	0.026	0.552***	0.012
	(0.14)	(0.31)	(<0.01)	(0.41)
	[0.18]	[0.33]	[<0.01]	[0.59]
HighEq	-0.012***	0.016***	0.011	0.004
	(<0.01)	(<0.01)	(0.54)	(0.18)
	[0.46]	[0.45]	[0.91]	[0.77]
Before (Early)	-0.011	0.007	0.015	-0.007
	(0.11)	(0.29)	(0.73)	(0.11)
	[0.61]	[0.66]	[0.90]	[0.48]
APL * HighEq	0.013***	0.002	0.025	0.015***
	(<0.01)	(0.62)	(0.24)	(<0.01)
	[0.39]	[0.88]	[0.72]	[0.15]
HighEq * Before	0.008	0.006	0.149***	0.002
	(0.17)	(0.31)	(<0.01)	(0.56)
	[0.71]	[0.74]	[0.08]	[0.89]
APL * Before	0.005	-0.001	-0.149***	0.013***
	(0.28)	(0.83)	(<0.01)	(<0.01)
	[0.65]	[0.92]	[0.03]	[0.14]
APL * HighEq * Before	-0.059***	-0.040***	-0.172***	-0.041***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
	[<0.01]	[0.03]	[0.02]	[0.01]
Sponsor FE	No	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Prop Type FE	Yes	Yes	Yes	Yes
OrigYear FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Purpose FE	Yes	Yes	Yes	Yes
Observations	509757	509757	509735	509757
Pseudo R^2			0.150	
R^2	0.157	0.163		0.116

p-values in parentheses (robust) and brackets [clustered]
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 8: **Equity Tranche and Yield Spreads Cross-tabulation**

This table presents the mean yield spread for variable rate tranches in the sample according to the size of the equity tranche and the tranche's rating class. For deals with multiple tranches within a rating class, the observation is the dollar-weighted average of the coupons. *High Equity* indicates that the pool under consideration has *%Equity Tranche* greater than that of the median deal in its cohort.

	Tranche Rating			
Equity Tranche Size	AAA	AA	A	$\leq$ BBB
Low Equity	0.42	1.21	1.45	2.39
	(0.05)	(0.28)	(0.23)	(0.21)
High Equity	0.35	0.78	1.27	2.25
	(0.03)	(0.10)	(0.11)	(0.12)

Standard errors in parentheses

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 9: **Price Response to Equity Tranche**

This table presents OLS estimates from regressions of the yield spread (in percentage points) on loan pool characteristics. Each observation represents a *Pool* × *Rating Class* dollar-weighted spread for variable rate tranches, where we define *Rating Class* as *AAA*, *AA*, *A*, and *BBB and below*. *Late* is a dummy variable equal to 1 for deals from 2005, *Opaque* is a dummy variable equal to 1 for deals with *%NoDoc* greater than that of the median deal in its cohort, and *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal in its cohort. All standard errors are heteroskedasticity robust.

	(1) All	(2) nonAAA	(3) AAA	(4) All	(5) nonAAA	(6) AAA	(7) All	(8) nonAAA	(9) AAA
Late	-0.50***	-0.63***	-0.24***	-0.59***	-0.77***	-0.26***	-0.25**	-0.36**	-0.14*
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(0.05)	(0.02)	(0.07)
High Equity	-0.25**	-0.34**	-0.08						
	(0.02)	(0.02)	(0.17)						
Opaque				-0.18	-0.30	-0.07	0.26	0.19	0.12
				(0.43)	(0.42)	(0.45)	(0.19)	(0.53)	(0.42)
High Equity * Opaque				-0.32***	-0.44***	-0.08	-0.36**	-0.58**	-0.09
				(<0.01)	(<0.01)	(0.35)	(0.04)	(0.03)	(0.46)
High Equity * Not Opaque				-0.04	-0.10	-0.05	-0.17	-0.45	-0.06
				(0.85)	(0.77)	(0.50)	(0.47)	(0.30)	(0.67)
Rating FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sponsor FE	No	No	No	No	No	No	Yes	Yes	Yes
Observations	379	262	117	379	262	117	379	262	117
R^2	0.43	0.30	0.15	0.45	0.33	0.17	0.66	0.65	0.58

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table 10: **Robustness – Alternate Channels**

This table presents our main results from earlier tables alongside specification that include other variables that capture the roles and connections of the various agents in the securitization chain. *Late* is a dummy variable equal to 1 for deals from 2005, *%NoDoc* is the percent of the loan pool with no documentation loans, *FICO* is the pool's weighted average FICO score, *LTV* is the pool's weighted average loan-to-value ratio, *%ARM* is the percent of the loan pool with adjustable rate mortgage loans, *GeoDiverse* measures the geographic diversity and is computed as 100 - (percent of largest one state origination concentration) in the mortgage pool, *GeoHerfindahl* measures the geographic Herfindahl concentration and is scaled to range from 0 (maximum diversification) and 100 (all loans from a single state), *Opaque* is a dummy variable equal to 1 for deals with *%NoDoc* greater than that of the median deal, *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal, and *SellAndService* is a dummy variable equal to 1 for deals where the issuer is also the primary servicer. *TopOrigAndService* is a dummy variable equal to 1 for deals where the top originator in the pool is also the primary servicer. All specifications include Sponsor fixed effects, and all standard errors are clustered at the Sponsor level.

	%AAA		%Equity		AbDefault	
	(1)	(2)	(3)	(4)	(5)	(6)
Late	-3.305***	-3.114***	0.738***	0.698***	1.958*	1.640*
	(<0.01)	(0.01)	(<0.01)	(<0.01)	(0.09)	(0.09)
NoDoc	-0.024	-0.024	0.024***	0.024***		
	(0.32)	(0.33)	(<0.01)	(0.01)		
FICO	0.094***	0.089***	0.002	0.003	-0.013	-0.011
	(<0.01)	(<0.01)	(0.66)	(0.43)	(0.54)	(0.59)
LTV	-0.276**	-0.267**	0.017	0.014	-0.130	-0.203
	(0.01)	(0.02)	(0.44)	(0.55)	(0.28)	(0.10)
ARM	0.010*	0.008*	0.004*	0.005*		
	(0.06)	(0.09)	(0.06)	(0.05)		
GeoDiverse	0.048**		0.001			
	(0.02)		(0.86)			
GeoHerf		-0.041***		0.000		
		(0.01)		(0.96)		
Opaque					4.426***	4.556***
					(<0.01)	(<0.01)
HighEq					2.430*	2.636*
					(0.08)	(0.05)
Opaque*HighEq					-3.698*	-4.564*
					(0.09)	(0.06)
HighMezz					-1.617	-1.711
					(0.47)	(0.44)
Opaque*HighMezz					1.111	1.683
					(0.71)	(0.59)
SellAndService		0.980		-0.486		0.898
		(0.36)		(0.18)		(0.75)
TopOrigAndService		0.513		-0.302		-4.167***
		(0.44)		(0.29)		(0.01)
Sponsor FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	215	215	215	215	162	162
R^2	0.829	0.828	0.503	0.517	0.385	0.431

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

A Appendix

A.1 Sample Construction and Data Collection

We use a stratified random sampling method to select private-label (i.e., non-agency backed) RMBS deals for inclusion in our study. We choose two time periods for our sample selection: an "early period" that covers deals from 2001-02 and a "late period" that covers deals from 2005. This stratification strategy allows us to separate out time-specific effects from our main cross-sectional results. It also allows us to investigate the time variation in the functioning of this market and exploit changes in anti-predatory-lending laws. Ashcraft and Schuermann (2008) report that the issuance of non-agency mortgage-backed securities increased eight-fold from $99 billion in 2001 to $797 billion in 2005 in the sub-prime and Alt-A segment. Thus our sample covers both an early/nascent period and a relatively matured period of RMBS market. We also stratify the sample along the prime-subprime dimension, slightly over-sampling the subprime pools to make sure that portion of the sample is large enough to make statistically meaningful inference. Our random sample begins with 234 deals. Due to variation in the data items included in the filings, our main regression specifications include 215 deals (cross-sectional determinants)and 163 deals (ex-post performance) that have full data on all variables of interest.

We collect data on mortgage pools and their tranches from Form 424(b)(5) filings which are submitted to the SEC pursuant to SEC Rule 424(b)(5). While the detail of the information provided varies slightly from deal to deal, the form typically contains data on all the major participants in the deal (e.g., sponsor, originators), pool-level characteristics and tranche-level data. Among other items, these data specifically include the loan originators and the share of the deal they originated, weighted average loan-to-value (LTV) ratio, weighted average FICO score, and a breakdown of loan types, geography and loan documentation levels within the pool.

Form 424(b)(5) also provides a listing of each tranche in the pool along with its principal

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

amount and credit rating. For our analysis, we aggregate the tranches into three bins: AAA-rated tranches, mezzanine tranches and equity tranches. We present a detailed discussion of the equity tranche in Section 2. The AAA tranche is self-explanatory and the mezzanine tranche is simply the subordinated tranche that lies between the AAA and equity tranches. The publicly offered tranches (AAA and mezzanine) include ratings from at least two major credit rating agencies. While disagreements in ratings among the ratings agencies are rare for the senior tranches, we use the lower of the ratings when conflicts occur.

We match these deals with detailed loan-level data obtained from CoreLogic. Pools in our sample cover over 500,000 individual mortgages. We obtain key information for each loan in a given pool from CoreLogic such as the loan amount, FICO score, LTV ratio, and loan type along with location of the property and various other characteristics. Finally, we obtain the ex-post performance of these loans from CoreLogic as well. We obtain information on the incidence of delinquency and foreclosure anytime from the origination of the deal through December 2011. This information allows us to conduct our test relating tranche structure to ex-post loan performance.

To illustrate the representativeness of our random sampling strategy, we compare some key summary statistics below to Ashcraft et al. (2010), whose target sample population is similar to ours.

		Deal Characteristics			Loan Characteristics		
Sample	Time Period	Deals	Size	Loans/Deal	Loans	CLTV	FICO
Our Random Sample	2001-2002, 2005	234	\$775m	3064	509,757	80%	657
Ashcraft et al. (2010)	2001-2007	3,144	\$749m	3840	12,074,103	84%	656

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

A.2 Example Documentation Description from a Deal Prospectus

Series Name: ABFC Mortgage Loan Asset-Backed Certificate, Series 2002-WF2

> The Originator's subprime mortgage loan programs include a full documentation program, a "stated income, stated asset" program and a "lite" documentation program. Under the full documentation program, loans to borrowers who are salaried employees must be supported by current employment information in the form of one current pay-stub with year-to-date information and W-2 tax forms for the last two years (a complete verification of employment may be substituted for W-2 forms). The Originator also performs a telephone verification of employment for salaried employees prior to funding. In some cases, employment histories may be obtained through V.I.E., Inc., an entity jointly owned by the Originator and an affiliated third party, that obtains employment data from state unemployment insurance departments or other state agencies. Under the full documentation program, borrowers who are self-employed must provide signed individual federal tax returns and, if applicable, signed year-to-date income statements and/or business federal tax returns. Evidence must be provided that the business has been in existence for at least one year. If the business has been in existence less than two years, evidence must be provided that the applicant had previously been in the same line of work for at least one year. Under the full documentation program, at certain loan-to-value ratio levels and under certain circumstances not all sources of funds for closing are verified as the borrowers.
>
> Under the Originator's "Stated Income, Stated Asset" program, the applicant's employment, income sources and assets must be stated on the initial signed application. The applicant's income as stated must be reasonable for the applicant's occupation as determined in the discretion of the loan underwriter; however, such income is not independently verified. Similarly the applicant's assets as stated must be reasonable for the applicant's occupation as determined in the discretion of the loan underwriter; however, such assets are not independently verified. Except under the Stated Asset Program, verification of funds sufficient to close the mortgage loan is performed. Under

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

the "LITE" Documentation program, the Originator reviews the deposit activity reflected in the most recent six or twenty-four consecutive months of the applicant's bank statements as an alternative method of establishing income. Maximum loan-to-value ratios within each credit level are lower under the stated income, stated asset program than under the full documentation program.

A.3 Appendix Tables

Table A.1: **Institutions and their Various Roles**

This table presents the most common institutions in the sample and the frequency in which they participated in various roles.

Institution	Seller	Top Originator	Type
Ace	5	0	Mortgage Lender
Ameriquest	16	16	Mortgage Lender
Bear Stearns	22	0	Investment Bank
Bank of America	29	24	Commercial Bank
Citi	11	4	Commercial Bank
Credit Suisse	19	11	Investment Bank
Countrywide	6	13	Savings and Loan
Deutsche Bank	6	0	Commercial Bank
Goldman Sachs	21	0	Investment Bank
HSBC	3	0	Commercial Bank
IndyMac	11	12	Savings and Loan
JP Morgan	9	5	Commercial Bank
Lehman Brothers	7	4	Investment Bank
Merrill Lynch	8	1	Investment Bank
Option One	8	13	Mortgage Lender
Stanwich	3	0	Mortgage Lender
UBS	7	0	Commercial Bank
Washington Mutual	13	15	Savings and Loan
Wells Fargo	17	31	Commercial Bank
Other	13	85	

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table A.2: **Default Model**

This table presents the results of the default model. We use the estimated coefficients of this model to predict the loan-by-loan probability of delinquency to construct our measure of *AbDefault* used in the paper. Following prior literature (e.g., see Demyanyk and Van Hemert, 2011), we include the borrower's FICO score, the loan-to-value ratio, loan purpose (e.g., Refinancing with Cash-Out), loan type, (e.g., 5-year Interest Only), state fixed effects, and year fixed effects. The results below show the key drivers of default risk, with the point estimates on the other variables in the estimation omitted in the interest of space. The main specification is a logit model that is separately estimated for each cohort (early and late periods).

	Early Cohort		Late Cohort	
	$\hat{\beta}$	Std Error	$\hat{\beta}$	Std Error
Loan-to-Value	0.008***	(<0.001)	0.012***	(<0.001)
Combined LTV	0.006*	(0.051)	0.013***	(<0.001)
FICO	-0.009***	(<0.001)	-0.007***	(<0.001)
Negative Amortization	0.019	(0.788)	0.404***	(<0.001)
Pool StDev(LTV)	-0.028***	(0.008)	-0.044***	(<0.001)
Pool WtAveLTV	0.063***	(<0.001)	0.032***	(<0.001)
Pool StDev(FICO)	-0.016***	(<0.001)	-0.003***	(<0.001)
Pool WtAveFICO	0.001***	(<0.001)	-0.001***	(<0.001)
Product = Fixed (omitted)				
Product = ARM	0.181***	(<0.001)	-0.297***	(<0.001)
Product = Balloon 15/30	0.127***	(0.006)	-0.585***	(<0.001)
Product = 3yr IO ARM	0.130	(0.616)	-0.099***	(<0.001)
Product = 5yr IO ARM	-0.046	(0.622)	-0.087***	(<0.001)
Product = 7yr IO ARM	-0.136	(0.392)	-0.204***	(<0.001)
Product = 10yr IO ARM	0.152	(0.683)	0.105***	(<0.001)
Property = Single Family (omitted)				
Property = Condo	-0.190***	(<0.001)	-0.029**	(0.044)
Property = Coop	0.244	(0.386)	-0.268***	(0.006)
Property = Multi-Unit	0.215***	(<0.001)	0.208***	(<0.001)
Property = Townhouse	0.230	(0.365)	0.144	(0.317)
Property = PUD	-0.148***	(<0.001)	-0.005	(0.685)
Property = Manufactured	0.319***	(<0.001)	0.512***	(<0.001)
Purpose = Purchase (omitted)				
Purpose = Refi (cashout)	-0.004	(0.811)	-0.043***	(<0.001)
Purpose = Refi (no cashout)	-0.092***	(<0.001)	-0.057***	(<0.001)
State FE	Yes		Yes	
Origination Year FE	Yes		Yes	
Observations	143072		366640	
Pseudo R^2	0.181		0.090	

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table A.3: **Abnormal Default with Alternative Default Models**

This table presents OLS estimates from regressions of *AbDefault* on loan-pool characteristics, where *AbDefault* is the actual pool-level delinquency rate minus the model-predicted delinquency rate, and is measured in percentage points (see Section 3.2 for further details). This table present results where the model-predicted default rate is computed using different models (logit or linear probability model) and estimation period (cohort-by-cohort or full sample). *Opaque* is a dummy variable equal to 1 for deals with *%NoDoc* greater than that of the median deal in its cohort, *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal in its cohort, *HighMezz* is a dummy variable equal to 1 for deals with *%Mezzanine Tranche* greater than that of the median deal in its cohort, *Late* is a dummy variable equal to 1 for deals from 2005, *FICO* is the pool's weighted average FICO score, and *LTV* is the pool's weighted average loan-to-value ratio. All standard errors are clustered at the sponsor level.

Default Model:	logit		LPM	
Estimation Window:	By Cohort (1)	Full Sample (2)	By Cohort (3)	Full Sample (4)
Late	1.958*	3.392**	2.163**	2.850**
	(0.07)	(0.01)	(0.03)	(0.02)
WtAveFICO	-0.013	-0.024	-0.013	-0.024
	(0.44)	(0.16)	(0.53)	(0.25)
WtAveLTV	-0.130	-0.166	-0.226	-0.297*
	(0.32)	(0.24)	(0.15)	(0.08)
Opaque	4.426***	4.715***	4.420***	4.595***
	(0.01)	(<0.01)	(<0.01)	(<0.01)
HighEq	2.430**	2.549**	2.321*	2.578**
	(0.04)	(0.04)	(0.06)	(0.04)
Opaque*HighEq	-3.698*	-3.646*	-2.853*	-3.465**
	(0.05)	(0.07)	(0.09)	(0.05)
HighMezz	-1.617	-1.980	-1.370	-1.795
	(0.40)	(0.30)	(0.43)	(0.30)
Opaque*HighMezz	1.111	0.907	0.937	1.458
	(0.69)	(0.77)	(0.67)	(0.53)
Sponsor FE	Yes	Yes	Yes	Yes
Observations	162	162	162	162
R^2	0.385	0.436	0.380	0.434

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056

Table A.4: **Ex-Post Outcomes: Loan-Level Default**

This table presents loan-level estimates from regressions of *LoanDefault* on loan, pool, and state characteristics using logit regression, with column (5) reporting the marginal effects evaluated at the means. *LoanDefault* equals 1 for loans that experience future delinquency. *Opaque* is a dummy variable equal to 1 for deals with *%NoDoc* greater than that of the median deal in its cohort, *HighEq* is a dummy variable equal to 1 for deals with *%Equity Tranche* greater than that of the median deal in its cohort, *HighMezz* is a dummy variable equal to 1 for deals with *%Mezzanine Tranche* greater than that of the median deal in its cohort, All regressions include the loan and pool control variables including FICO, LTV, Combined LTV, Negative Amortization, property type, loan purpose, year of origination, and state. Standard errors are computed as heteroskedasticity-robust (parenthesis) and clustered at the Sponsor level [brackets].

	Full Sample					Sponsor is Top Originator No	Sponsor is Top Originator Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HighEq	0.548***	-0.098***	0.136***	0.144***	0.024***	-0.021	0.196***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(0.58)	(<0.01)
	[<0.01]	[0.04]	[0.03]	[0.03]	[0.03]	[0.76]	[<0.01]
Opaque		0.239***	0.292***	0.289***	0.048***	0.164***	0.107***
		(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)
		[<0.01]	[<0.01]	[<0.01]	[<0.01]	[0.22]	[0.07]
Opaque*HighEq			-0.171***	-0.182***	-0.030***	-0.001	-0.295***
			(<0.01)	(<0.01)	(<0.01)	(0.99)	(<0.01)
			[<0.01]	[<0.01]	[<0.01]	[0.99]	[<0.01]
HighMezz				-0.061***	-0.010***	-0.354***	0.193***
				(0.01)	(0.01)	(<0.01)	(<0.01)
				[0.62]	[0.62]	[0.08]	[<0.01]
Opaque*HighMezz				0.018	0.002	-0.075*	0.372***
				(0.50)	(0.50)	(0.09)	(<0.01)
				[0.91]	[0.91]	[0.45]	[0.02]
Sponsor FE	No	No	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Prop Type FE	No	Yes	Yes	Yes	Yes	Yes	Yes
OrigYear FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Purpose FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	509735	509735	509735	509735	509735	215501	294213
Pseudo R^2	0.066	0.144	0.151	0.151	0.151	0.147	0.147

p-values in parentheses (robust) and brackets [clustered]

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The version of record Rev. Financ. Stud. (2016) is available online at: http://dx.doi.org/10.1093/rfs/hhw056