Correlated Collateral

Taylor Nadauld, Shane Sherlund, and Keith Vorkink[1]

Brigham Young University,

Federal Reserve Board of Governors,

and Brigham Young University

December 8, 2011

[1]We thank Mike Anderson, Effi Benmelech, Brian Boyer, Josh Coval, Andrew Karolyi, Craig Merrill, Todd Mitton, Phil Strahan, Rene Stulz, Mike Weisbach, and seminar participants at BYU, The Ohio State University, and the University of Utah for helpful comments. Vorkink acknowledges financial support from the Harold F. and Madelyn Ruth Silver Fund and his Ford Research Fellowships. Contact Information: Nadauld: Marriott School, Brigham Young University, Provo, UT, 84602. Phone: 801-422-6033. Email: taylor.nadauld@byu.edu. Sherlund: shane.m.sherlund@frb.gov. Vorkink: Marriott School, Brigham Young University, Provo, UT, 84602. Phone: 801-422-1765. Email: keith_vorkink@byu.edu. The analysis and conclusions contained in this paper are those of the authors and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System, its members, or its staff.

Abstract

The prioritized cash flow rules that govern structured finance essentially guarantee that senior tranches will only default in the worst states of the world. In this paper we present empirical evidence which suggests that the impact of economic catastrophe on a structured finance instrument also depends critically on the degree of correlation in the underlying collateral generating the cashflows. Subprime mortgage-backed securities with highly geographically concentrated mortgage collateral (our proxy for correlation) experienced substantially higher deal-level default rates and more significant credit rating downgrades during the financial crisis. Not all deals had geographically concentrated collateral. In fact we document considerable cross-sectional variation. Collateral concentration did not appear to be priced in the subprime bond market.

1 Introduction

The prioritized cash flow rules that govern structured finance effectively guarantee that senior tranches will only default in times of "economic catastrophe" (Coval, Jurek, and Stafford (2009a)). In this paper we argue that the exposure of a structured finance instrument to economic catastrophe is heavily influenced by the nature of a deal's collateral. Structured finance securities with correlated collateral are more exposed to economic catastrophe than securities with well-diversified collateral. For example, in a cross-section of subprime residential mortgage-backed securitization deals, we document that deals with highly correlated collateral had substantially higher default rates. Such a result suggests that while the prioritized cash flow rules of structured finance create securities that are buffered against most bad economic outcomes, another critical element in structured finance is correlation in the underlying collateral.[1]

A novel empirical observation provides the fundamental motivation for this paper. We document a considerable amount of cross-sectional variation in the geographic concentration of mortgage collateral across RMBS deals, including many deals with very little diversification.[2] For example, in a sample of over 1,200 subprime RMBS deals collateralized by an average of over 5,000 loans per deal, we find that deals in the 90th percentile of geographic concentration at the time of deal origination had 47.0% of a deal's mortgage collateral originated in the state of California, compared to only 13.6% for deals in the 10th percentile of geographic concentration. Consistent with the notion of diversification, we find that cross-sectional variation in the geographic concentration of collateral has significant power in explaining deal-level default rates during the financial crisis. As an

[1] This statement requires a more careful description of the term catastrophe. Diversification is obviously no defense against an all-encompassing economic catastrophe. Given that we focus specifically on subprime RMBS, which are collateralized by residential mortgage pools, our use of the term catastrophe more specifically refers to geographically-localized economic catastrophe. We address this issue in more detail in Section 3.

[2] Throughout the paper we use the terms correlation and geographic concentration synonymously. In developing economic intuition we prefer the term correlation. Empirically we measure correlation by constructing variables which capture geographic concentration in mortgage collateral. Thus, when discussing empirical tests and results, we generally use the term geographic concentration.

example, 12-, 24-, and 36-month default rates for the top quintile of the 60 deals originated in June 2006 were 4.5, 9.9, and 15.5 percentage points higher, respectively, than default rates in the bottom quintile of geographically concentrated deals.[3]

Our analysis is focused on two important questions. First, given the prioritized cash flow rules which govern non-agency RMBS, what is the impact of correlation on the performance of RMBS mortgage collateral? Second, we ask why some deals were well diversified while others were not? In answering the first question, we show that correlation in deal collateral is unambiguously harmful to senior tranches in a securitization structure, consistent with theoretical predictions. This result is driven by the fact that correlated defaults represent the only situation in which deals would experience large enough default rates to erode the principal of a senior tranche. We show that deals with more geographically concentrated collateral (our proxy for correlation) have substantially higher rates of default during the financial crisis. On average, a deal in the 75th percentile of geographic concentration experienced more than two percentage point larger default rates compared to a deal in the 25th percentile of concentration. In terms of economic magnitude, measures of geographic concentration are behind only FICO scores and loan-to-value ratios in their impact on deal-level default rates. We also find that the probability of experiencing a ratings downgrade is positively correlated with geographic concentration.

In answering the second question, why were some deals well diversified while others were not, our paper provides three pieces of evidence. First, cross-deal variation in collateral concentration cannot be explained by geographic concentration in the *supply* of mortgages offered to the secondary mortgage market.[4] The data suggest that at the deal level, collateral is geographically concentrated at nearly twice the intensity of geographic concentration in subprime mortgage originations. For example, between the years 2004-2006, nearly 18% of all subprime loans were originated in California. In contrast, over the

[3] Plots of deal level default rates as a function of geographic concentration are included in Figures 4 and 5.

[4] Throughout the paper we refer to mortgage supply and demand. Since our analysis is at the deal-level, mortgage demand refers to the demand for mortgage collateral from deal underwriters in the secondary mortgage market. Mortgage supply refers to mortgage originations in the primary mortgage market.

same time period, California-originated loans represented almost 30% of deal collateral, on average. This evidence suggests that correlated collateral appears to be a secondary market demand-driven phenomenom.

Second, cross-sectional variation in collateral concentration can be explained by *expected* rates of house price appreciation. Using housing supply elasticity measures created by Saiz (2009) as an instrument for expected house price appreciation, we find that when deals are geographically concentrated, they tend to be concentrated in areas with low housing supply elasticity. This result highlights the importance of *expected* collateral values in the *expected* performance of a securitized pool of mortgages.

Third, we find that variation in collateral concentration appears to be influenced by the means by which mortgage collateral is obtained. Collateral pools are more geographically concentrated when the majority of securitized loans were originated by an entity affiliated with the deal arranger. In addition, affiliated loan pools are even more concentrated in geographies with high expected price appreciation. This finding indicates that the industrial organization of the mortgage securitization process influences the composition of collateral pools in important ways.

In a concluding section of the paper we investigate whether investors were adequately compensated for correlation risk at deal origination. Our analysis suggests that the answer is no. We fail to find any evidence that the coupon payments of senior bonds with geographically concentrated collateral were systematically higher than the coupon payments promised senior bonds with little geographic concentration in their collateral. Our pricing tests are able to identify differences in coupons with regards to differences in deal-level FICO scores, suggesting that the lack of evidence for pricing differences with regards to correlation may not be a power issue.

This paper contributes to our understanding of the economics of securitization in the following ways. First, while the economic causes of increased default rates through time have been documented in the literature, little attention has been paid to how the structure of securitization deals (i.e., how deals are put together) contributed to the subprime mortgage crisis. For example, Keys, Mukherjee, Seru, and Vig (2010), and Keys, Seru, and Vig (2010) carefully demonstrate that securitization caused a reduction in screening at loan origination, particularly with regards to soft information, resulting in higher default rates than otherwise similar loans.[5] Our paper contributes to the literature investigating default rates by demonstrating that, conditional on having been originated, the manner in which loans are pooled can have a significant impact on pool-level default rates, which is ultimately what matters most for the performance of securities arising from securitization. Stated differently, while it is true that securitization caused the origination of poor quality loans, the pooling of the loans after having been originated and securitized is also of central importance to the performance of subprime securities.

This paper also contributes to a fairly nascent empirical literature on the role of correlation in structured finance. Longstaff and Rajan (2008) demonstrate that the expected clustering of corporate defaults explains 27% of the CDX spread. Coval, Jurek, and Stafford (2009a) model the payoff of highly-rated structured finance instruments and demonstrate that while such bonds default during the worst possible states of the world (on account of the way deals are structured), the highly-rated instruments offer little compensation for this risk. In a companion paper, Coval, Jurek, and Stafford (2009b) discuss the importance of collateral correlation in the context of a broader review of the economics of structured finance. Our paper demonstrates that, even in the presence of economic catastrophe, the composition of a deal's collateral can influence the severity of the impact of catastrophe. This is an important point to make given the commonly held view that the housing market collapse and subsequent jump in default rates simply represented an unfortunate draw from a distribution of potential outcomes. Deal-level default rates did not explode solely on account of a bad housing market draw. The explosion in defaults in some deals appears to have been exacerbated by correlated collateral.

[5] The rise in subprime default rates has been attributed to a host of factors, including an increase in the supply of credit made available to subprime borrowers (Mian and Sufi 2009), increased demand from active secondary mortgage markets (Nadauld and Sherlund 2009), and a general decline in underwriting standards through time (Demyanyk and Van Hemert 2008).

We proceed as follows. We briefly describe the data in Section 2 but leave the details to a data appendix. We describe the economics of correlation in the context of structured finance in Section 3 and demonstrate empirically the relationship between geographic concentration and deal-level default rates. Section 4 contains empirical tests which identify factors influencing cross-sectional variation in the geographic concentration of deal-level collateral. Section 5 investigates whether differences in geographic concentration are priced in the coupon payments offered investors. Section 6 concludes.

2 Data Description

Our analysis consists of three sets of empirical tests. Each of the tests relies generally on the same underlying source of data, but the unit of analysis differs across the tests. In this section we provide a brief overview of the construction and features of each estimation sample. We include greater detail on the construction of the sample and specific variables in a data appendix.

2.1 RMBS Deal-Level Data

ABSNet, a subsidiary of Standard & Poor's contains summary information on the universe of subprime RMBS deals.[6] ABSNet data is at the deal level. LoanPerformance, a subsidiary of First American Trust, provides detailed data on the universe of subprime loans. LoanPerformance data is at the loan level. We match the loan-level data from LoanPerformance to the deal level data provided by ABSNet. The matching allows us to include deal level attributes and performance as a function of the attributes of the individual loans that make up the deal collateral. Our final sample includes loan vintages originated in 1997-2007 and includes 1,234 deals collateralized by over 6 million loans. In evaluating deal-level default rates, our sample runs from 1998-2009. The data appendix

[6] Deal-level summary information from ABSNet includes the original credit rating, original principal amount of each tranche, and tranche CUSIPs (each tranche, or bond, has a unique CUSIP).

describes the details of the matching and aggregation of the loan-level attributes to the deal level.

2.2 Sampling Deal-Level Default Rates

The loan-level data from LoanPerformance tracks the payment history of subprime loans each month, beginning in the month of loan origination. Because we know which loans belong to which deals, we can calculate deal-specific default rates each month from deal inception. Deal-level default rates are a function of many factors, perhaps none more important than time itself. In order to compare default rates across deals originated over different time periods, we compute deal-level default rates 12, 24, and 36-months after deal origination. Our data tracks the performance of the subprime loans in our sample through September 2009. Thus, we can calculate 48-month default rates for any deal originated before September 2005, 36-month default rates on deals originated before September 2006, 24-month default rates for deals originated before September 2007, and 12-month default rates for all deals originated before September 2008.

2.3 The Bond Pricing Sample

A subprime securitization deal produces about 17 bonds, on average, but the bonds share the same collateral. In the pricing analysis we focus on specific bonds arising from a deal, each with a unique CUSIP. Bond-level transaction data are compiled and provided by Thomson Reuters EMaxx services. The observations represent open-market transactions primarily of insurance funds, but include transactions of government state pension funds and mutual funds. We match the bond-level transaction data with deal-level securitization data by CUSIP. We restrict our sample to include only transactions on floating rate bonds purchased at par either at deal origination, or at par within 12 months of deal origination. The at par and at-or-within 12 months of origination sample includes 1,234 unique bond-transactions, while the at-origination only sample includes 745 unique bond-transactions.

The data appendix includes more specific details of the construction of this data.

2.4 Empirical Measures of Collateral Correlation and Summary Statistics

Our first measure of collateral correlation is a Herfindahl index measuring the geographic concentration in a deal. We calculate the percent of each deal's principal concentrated in a given state. The deal-level Herfindahl index is then calculated as the sum of the squared weights, expressed as $\sum_{i=1}^{n} w_i^2$, where i indexes 51 states (Washington D.C. is counted as a separate state). We report summary statistics on deal-level Herfindahl measures in Table 1. Our second measure of collateral correlation is simply the percentage of a deal's loans that are originated in the state of California. We feel this measure is intuitive and relevant given that the average deal has 28% of its loans concentrated in California. By construction, the California concentration measure is highly correlated with the deal-level Herfindahl index. The two measures are different in that the Herfindahl index is a value-weighted measure while the California measure is equally weighted.

What does our proxy for correlation actually measure? It is reasonable to assume that subprime mortgage loans originated within a given geography are exposed to similar economic conditions that directly impact the likelihood of default. Idiosyncrasies in housing markets, employment opportunities, and the legal environment are examples of locally determined factors that impact mortgage defaults. In a regression framework we can explicitly control for factors known to impact mortgage defaults, such as FICO scores, LTV ratios, and even rates of house price appreciation. We cannot control for locally-determined idiosyncratic factors that are unobserved. We believe our geographic proxy should capture the influence of locally-determined factors that explain default rates above and beyond the influence of house price patterns, FICO scores, loan-to-value ratios, etc.

We believe that measures of geographic concentration are reasonable proxies for collateral correlation for another, more practical reason. Market participants often rely on the collateral description from trustee's reports and analytical packages (i.e. Bloomberg) when making investment decisions. Trustee's reports and analytical packages often report the top 5 states most represented in a given deal structure as a measure of collateral correlation.

Table 1 reports summary statistics on the attributes of securitization deals for the entire sample as described in Section 2. The sample contains deals originated between 1998-2008, but we calculate deal-level default rates through September 2009. We report the mean and standard deviation of key deal attributes. The average subprime deal in the sample has a Herfindahl measure of 13.8%, and an average of 28% of deal collateral by loan count in California. The loan-weighted average FICO score is 629.6, with an average combined loan-to-value ratio of 85.4%. The average house price appreciation in the year preceding deal origination was 13.7%.

Table 1 reveals a substantial increase in the number of securitization deals originated over the last decade. More deals were originated in 2005 and 2006 alone than in the entire preceding seven years combined. Though not reported in the table, the principal included in the deals was also substantially larger, having increased from a median size of $478 million in 2000 to $1.02 billion in 2006. Table 1 also highlights an important trend in the structure of securitization deals through time. Collateral concentration increased each year in the early portion of the sample, peaking in 2004, and then declined slightly each year until 2007. Deal-level default rates, measured 24-months after deal origination, remained steady in the 8-12% range for the 2002-2005 vintages until rising to over 20% in the 2006 vintage, eventually reaching as high as 24% for 2007 vintage deals. Rates of house price appreciation grew monotonically through 2005, but turned negative in 2007. FICO scores increased through the sample period, as did combined loan-to-value ratios and the percentage of adjustable-rate loans.

3 The Economics of Correlated Collateral

In this section we evaluate the role, if any, correlated collateral plays in the performance of RMBS. We begin the section with a brief conceptual discussion of the economics of correlation and structured finance. We then turn our attention to the empirical analysis, with a focus on the relationship between collateral concentration and deal-level defaults.

The impact of correlation on a tranche within a securitization deal is subtle. This is because the relationship between collateral correlation and expected default depends critically on the prioritized allocation of cash flows arising from deal collateral. Recent treatments on the subject include Lando (2004) and O'Kane (2008). In the spirit of the models presented in these papers, as well as Coval, Jurek, and Stafford (2009). We present a simple mixed-binomial model of securitization in the Appendix A2 for the interested reader. The model makes simple assumptions describing the default on underlying mortgages and then builds a standard capital structure of tranches based on the underlying pool of mortgages. The model allows for a straightforward investigation into the ex-ante default probabilities of tranches as they relate to a variety of model inputs, including correlation of the underlying mortgages. A key parameter in the model is ρ, which governs the role of correlation. Intuitively, deals that are geographically concentrated (diversified) would be modeled with high (low) values of ρ.[7]

Figure 1 provides a plot of the expected probability of (any) default for a variety of tranches and over a range of correlation values. In Figure 1, higher values along the 'Tranche' axis correspond to more senior tranches within a deal. Figure 1 demonstrates the key relationship used to motivate our empirical study. Specifically, the figure shows the existence of a *positive* relationship between expected probability of default and correlation. For senior tranches the intuition is fairly straightforward. More senior bonds are insulated against idiosyncratic default shocks (an individual borrower loses their job

[7] The intuition may also be understood by thinking about a common decomposition of risk in valuing equities that includes a market component, an industry component, and a firm-specific component. We equate the geographic concentration in mortgage pools in our study to something akin to industry risk in equities, which is diversifiable.

and defaults on their loan) given that junior tranches are exposed to the first defaults of mortgages within the pool. So, if the risk of a typical loan within the pool is primarly idiosyncratic, or stated differently, if the typical mortgage has a low default correlation with other mortgages, then senior tranches have very little chance of experiencing default because most of the loans would need to experience an idiosyncratic shock for defaults to make their way up the capital structure sufficient to hit senior tranches. On the other hand, as the pool of underlying mortgages becomes more and more correlated, mortgage defaults are likely to coincide with other mortgage defaults (this case could be explained by some bad macroeconomic shock leading to widespread default on mortgages within a pool).[8]

Given that Figure 1 indicates correlation increases default risk for senior tranches we would anticipate that underwriters of deals would, ceteris peribus, create deals with low correlation between the individual mortgages. What we find in the data is surprising in this regard. Not only do many deals include highly correlated mortgages but we also find that the variation in correlation across deals is surprisingly high.

3.1 Differences in Default Rates for High and Low Correlation Deals

In this section we report the results of empirical tests designed to identify the impact of collateral correlation on deal-level default rates. We begin by analyzing simple time-series averages of deal-level default rates. Mortgage analysts typically report the performance of deal collateral by loan vintage. Analyzing deals by vintage controls for the deal's seasoning, an important determinant in deal-level default rates. In our analysis, we classify

[8]One other relationship in Figure 1, not studied in this paper, is that correlation is *negatively* related to expected defaults for junior tranches within a pool. Indeed, collateral correlation and a positive macroeconomic shock provides the only hope of survival for a junior bond in a structured finance deal. Our decision to focus on senior bonds is driven primarily from data restrictions. In our pricing study we are only able to observe transactions on more senior (AAA-rated) bonds and are so constrained in our analysis.

a vintage as all the deals originated within the same month.

Figure 3 plots the average default rate through time for deals in the July 2003 vintage. In order to identify the influence of correlation, we stratify the sample into high correlation and low correlation portfolios and plot default rates for each portfolio. High correlation portfolios are comprised of deals with a Herfindahl measure greater than or equal to the 75th percentile of Herfindahl measures within the July 2003 vintage. Low correlation portfolios are comprised of deals with a Herfindahl lower than or equal to the 25th percentile within the July 2003 vintage. Deal-level default rates in the high correlation portfolio and low correlation portfolio remain within a few percentage points of each other until the beginning of 2007, when the difference begins to widen. Differences in default rates widen to almost 15 percentage points by the 3rd quarter of 2009, a considerable margin.

Figures 3b and 3c provide analogous plots for the July 2004 and July 2005 vintages respectively. The path of default rates for the two portfolios track each other very closely until the beginning of 2007, at which point they begin to diverge. The plots indicate that high correlation portfolios have default rates about 5 percentage points higher than low correlation portfolios through the heart of the financial crisis.

Figure 4 plots the results of a more comprehensive test of the difference in default rates between high correlation portfolios and low correlation portfolios. Within each month-vintage in our sample with at least 8 deals, we sort deals into correlation quartiles. Within each vintage we calculate the difference between average default rates in the highest correlation portfolio and average default rates in the lowest correlation portfolio. The solid line in Figure 4 represents the average difference in within-vintage default rates, averaged across vintages. As of January 2003, 7 vintages meet the 8-deal criteria, so the plotted difference of .07% in January 2003 represents the average high-correlation minus low-correlation within-vintage difference in default rates for the 7 vintages. By January 2007, 64 vintages enter into the calculation. The results suggest that differences in default rates between high and low correlation portfolios peaked at 2.7% in May of 2008. The plotted confidence bands provide evidence that differences in default rates between high and low correlation portfolios became statistically different from zero around the first quarter of 2007.

3.2 Comprehensive Analysis of Deal-level Defaults

The next sets of empirical tests are designed to test implications presented in Section 2 and developed more fully in Appendix 2. Correlation is unambiguously harmful to senior tranches in a securitization deal, but perhaps more importantly, the impact of correlation is amplified in the presence of a negative economic "factor" draw. Measuring a negative "factor" draw empirically is difficult due to the fact that the "factor" is unobservable. The plots provided in Figure 3 and 4 suggest that the impact of correlation on default rates began to be most pronounced near Q1 2007. This time period corresponds roughly with the beginning of a period where borrowers began to fall behind on house payments, likely on account of declines in rates of house price appreciation which began in late 2005. As such, we choose January 2007 as the time period which we believe corresponds to the beginning of a period which represents a large negative factor draw.

We estimate the determinants of deal-level default rates and report results in Tables 2 and 3. Table 2 reports the results of an OLS regression of pre-2007 deal-level default rates on a set of deal-level attributes. The pre-2007 default rates estimated in Table 2 meet two criteria. First, they are default rates sampled between the years 1999 and 2006, and second, they are sampled from deals that have seasoned for 12, 24, or 36-months.[9] For example, the 36-month default rate sample could include deals from the July 2003 vintage with deal-level defaults sampled in July of 2006, or deals from the January 2001 vintage sampled in January 2004. This approach allows us to compare default rates across vintages while controlling for deal seasoning.

The regression includes several deal-level attributes to control for other sources of

[9] We begin sampling defaults in 1999 to allow the 1998 deals to have seasoned for at least 12 months.

variation in deal-level defaults. Specifically, we control for the influence of the housing market in two ways. First, we control for the rate of deal-level house price appreciation in the year prior to deal origination. This variable should capture whether deal collateral is concentrated in “hot” housing markets at origination. Second, we control for the deal-level cumulative rate of house price appreciation since a deal’s inception, thus capturing any influence of house prices on default rates over the life of the deal. We control for deal-level FICO scores and borrower leverage in the form of combined loan-to-value ratios. We also control for the percentage of loans in the deal with adjustable rates, that are owner-occupied, second mortgages, low or no documentation, and loans that are refinances. We report results when estimated with and without month fixed-effects (e.g. fixed-effect for all default rates sampled in July 2005, August 2005, etc.)

The results presented in Columns (1)-(4) of Table 2 suggest that the Herfindahl measure of collateral correlation has no statistical impact on deal-level default rates when default rates are sampled any time before 2007. Not surprisingly, higher deal-level FICO scores are associated with lower default rates. Loan-to-value ratios are positively associated with default rates. Deals with higher rates of house price appreciation prior to deal close are associated with lower rates of default. The remaining control variables have the expected sign.

Columns (5)-(8) of Table 2 report results with the California measure of correlation. After controlling for month fixed-effects, deals with a larger percentage of loans in California appear to default at slightly higher rates. A one standard deviation increase in the percent of a deal’s collateral in California (13%) is associated with modest increases in default rates of between .31% - 1%. Other control variables demonstrate the expected sign.

Table 3 tabulates results when 12, 24, and 36-month default rates are sampled after January 2007, the period we believe adequately represents an environment with negative, correlated economic shocks.[10] The results indicate that, as predicted in Figure 1, col-

[10] For example, 12-month default rates sampled any time after Jan. 2007 will include all deals originated

lateral correlation is associated with substantially higher default rates during a negative shock. Each of the 12, 24, and 36-month default rates, estimated with or without month fixed-effects, are influenced significantly by correlation in deal collateral. One standard-deviation increases in the Herfindahl measure of correlation (8%) contribute to 1.2-1.5% higher deal-level default rates, about one-fourth to one-eighth standard deviation in the distribution of default rates sampled during and after 2007. The impact of collateral correlation using the California measure is of a similar magnitude; one standard deviation increase in the percent of collateral in California is associated with a 1.5% increase in the 24-month default rate. In terms of economic magnitude, collateral correlation is behind only FICO scores and loan-to-value ratios in its impact on deal-level default rates during the financial crisis.

3.3 Credit Ratings and Collateral Concentration

The results of Tables 2 and 3 demonstrate that collateral concentration plays an economically important role in deal-level default rates. A second way to calibrate the economic effects of collateral concentration is through bond credit ratings. Credit ratings of subprime RMBS play an important role in the real economy for several reasons. First, financial institutions, particularly banks and insurance companies, rely on credit ratings in calculating capital requirements. Second, a substantial fraction of highly-rated RMBS ended up on the balance sheet of commercial and investment banks as well. Third, highly-rated subprime RMBS served as collateral in repurchase agreements (REPO), an important source of funding for broker/dealer investment banks (Gorton and Metric (2010)).

We analyze the relationship between collateral concentration and credit rating outcomes. Seniority within a deal plays an important role in determining bond credit ratings. As a control for bond seniority, we include in our sample only the most senior

between 2006 and Sep. 2008. Twenty-four month default rates will include deals originated between 2005 and Sep. 2007. Thirty-six month default rates sampled after Jan. 2007 will include all deals originated between 2004 and September 2006.

tranche from each RMBS deal in our full sample. Bloomberg reports the at-origination and current (as of June 2011) credit rating for each bond. As expected, every bond in this sample was rated AAA at origination. We group the bond's current credit rating into one of three categories. The "no ratings change" category consists of bonds that have maintained at least a A-rating during it's life. The "ratings downgrade" category consists of bonds that have been downgraded to a BBB, BB, B, CCC, CC, or D credit rating. Finally, the "not currently rated" category applies to bonds that do not have a current rating because the bond has matured or prepaid. The three categories of credit rating outcomes are estimated using a multinomial logit framework. We categorize the "no ratings change" category as the baseline.

We report estimates of the multinomial logit model in Table 4. We include the standard set of deal-level collateral controls. We also include fixed effects for the year in which a deal was originated in order to capture any unobserved variation through time. Standard errors are clustered by year to capture correlated ratings changes within a given year. The estimates reported in Column (1) indicate that the probability of experiencing a ratings downgrade, as compared to no change in ratings, is positively correlated with the Herfindahl measure of collateral concentration. Likewise, the probability of a bond pre-paying, as compared to no ratings change, is correlated with collateral concentration, though the estimated impact of collateral concentration on pre-payment is not as large as its impact on the likelihood of a ratings downgrade. We estimate a similar model with the percent of loans in California measure of concentration and report the qualitatively similar results in Columns (4) and (5). In terms of economic magnitude, the estimates predict that, holding all other variables constant, moving from the 25th percentile of concentration to the 75th percentile of concentration decreases the likelihood of maintaining at least a A-rating by almost 17 percentage points.

4 Explaining Cross-Sectional Variation in Collateral Correlation

This section of the paper addresses why were some deals well diversified while others were not. We begin with a discussion of summary statistics which document the nature of the cross-sectional variation. We then focus on three potential explanations of the observed cross-sectional variation, beginning with a consideration of primary-market mortgage supply versus secondary-market demand. Second, we investigate the importance of expected house price appreciation in a portfolio of subprime mortgages. Finally, we analyze the industrial organization of loan origination and securitization that could influence the composition of securitized loan pools.

4.1 Documenting Cross-Sectional Variation in Concentration

Tables 2 and 3 indicate that geographic concentration, our measure for correlation, within a pool of mortgages in a RMBS deal, signficantly impacts the default possibilities of tranches within the deal. The importance of geographic concentration in predicting defaults appears to be of first-order importance, trailing only measures such as FICO in predicting variations in defaults. In Table 5 we report cross-sectional distribution statistics of both our Herfindahl measure (Panel A) as well as our Percent of Deal in California measure (Panel B). The table reports the cross-sectional statistics for each year. From the results of Tables 2 and 3 that indicate a negative relationship between concentration and default, we might expect the cross-sectional distribution (across deals) of geographic concentration to be centered on low values of geographic concentration and with little variation across deals. What we find is quite the opposite. Table 5 shows that the cross-sectional variation in concentration can be quite large and exhibit substantial time-series variation. For example, in the early years of the data the distribution of the concentration measures is consistent with the predicted relationship, low mean values with little variation. However, in the mid-2000s both the center and standard devation of concentration increase dramatically. In 2004, the mean deal has a Herfindahl measure of .165 (% in California is 33.5%) and the 90th percentile is .276 (49.90 % in California). This indicates that ten percent of the deals constructed in 2004 had at least half of their mortgages coming from the state of California.

In the following sections we investigate explanations for such large variation in geographic concentration and find that mortgage demand, expected house price appreciation, and the industrial structure of the mortgage market help explain how so many geographically concentrated deals could be constructed.

4.2 Is Deal-Level Concentration Driven by Mortgage Supply or Mortgage Demand?

It is important to note that the geographic distribution of the supply of subprime mortgages from the primary market to the secondary market is itself not random. Mian and Sufi (2009) provide evidence of a causal relationship between the extension of subprime credit and house prices, arguing that the extension in credit caused the housing market boom. Yet, as mentioned previously, Gorton (2010) argues that, in their contract design, subprime mortgages relied heavily on the future path of house prices. Consistent with these arguments, Mayer and Pence (2008) document that areas with the highest rates of house price appreciation also had the highest rates of subprime loan origination. Thus, while easy credit fueled house price appreciation, e.g. Mian and Sufi (2009), it is also likely the case that high house price appreciation itself subsequently fueled a further easing of credit in the form of larger volumes of house price appreciation-dependent subprime mortgages. The result of the house price appreciation-credit supply feedback is geographic clustering in subprime originations.

We attempt to disentangle whether deal-level concentration is simply the result of concentrated supply by constructing measures of geographic concentration in the *supply* of subprime mortgages using mortgage origination data made available under the Home Mortgage Disclosure Act (HMDA).[11] The HMDA data reports mortgage originations in the primary mortgage market. We measure supply concentration at the origination level using "higher-cost" mortgages, a proxy for subprime loans, as classified in the HMDA data. For each of the years 2004, 2005, and 2006, we calculate a state-level Herfindahl measure of subprime mortgage originations. We also construct a "percent of loans originated in California" measure of supply concentration. As a second measure of available supply, we calculate concentration measures of the total securitization market in a given year using the LoanPerformance data.

Table 6 tabulates the mortgage supply, total securitization market, and deal-level concentration summary statistics. Panel A reports statistics using the Herfindahl measure and Panel B reports results under the percent California metric. The results indicate that geographic concentrations increase substantially in the securitization process. The primary market mortgage supply Herfindahl averages .064 during the years 2004-2006, compared to .091 in the total securitization market over the same period. This difference suggests that loans originated in the primary market are not securitized at equal rates across states. More striking, however, is the difference between mortgage supply measures of concentration and deal-level measures of concentration, particularly under the California measure of concentration. On average, from 2004 through 2006, HMDA data indicate that 17.5% of all subprime loans were originated in California. In constrast, over the same period, for the average securitization deal, 29.7% of a deal's securitized loans were originated in California. The table also reports the total percent of deals in a given year that had higher percentages of geographically concentrated collateral then the geographic concentration in the mortgage origination market. This calculation would be close to zero, on average, if loans were securitized and deals were constructed in perfect proportion to the geographies in which they were originated. The data suggest this is not the case. On average, 64.0% and 82.1% of all deals originated between 2004 and

[11] We discuss the HMDA dataset in the data appendix.

2006 exhibited higher rates of geographic concentration than the degree of concentration in the primary market using the Herfindahl and California measures of concentration, respectively. Taken together, the results presented in Table 6 provide evidence that a reasonably large fraction of the geographic concentration at the deal level is a secondary market demand phenomenom, and not simply the reflection of concentration in mortgage originations.

4.3 Expected House Price Appreciation and Collateral Concentration

In order to estimate the risk associated with a subprime RMBS, a rating agency must estimate both the probability of default and the expected loss given default. The probability of default in a loan pool is traditionally estimated as a function of loan attributes such as FICO scores, loan-to-value ratios, mortgage type, income documentation, loan purpose (refinance vs. purchase), and macroeconomic conditions.[12] Conditional on estimated default rates, expected recovery rates on the mortgage collateral, i.e., *expected* house prices, determine expected loss. Also, Gorton (2010) argues that by their very contract design, the value of subprime mortgages (particularly those with adjustable rates) relied heavily on the future path of house prices. For these reasons we expect conditions in the housing market to have an important influence on the composition of loan pools. In particular, we hypothesize that loan pools, when geographically concentrated, will be concentrated most in areas with high expected rates of house price appreciation.

4.3.1 An Instrument for Expected House Price Appreciation

Using realized rates of house price appreciation as a proxy for expected rates of appreciation is problematic because realized rates of house price appreciation are themselves

[12] The impact of specific loan attributes on loan default rates is documented by Sherlund (2008), Deng, Quigley, and Van Order (2000), and Pennington-Cross and Ho (2006). Loans with high FICO scores, low loan-to-value ratios, and low debt-to-income ratios default less frequently.

influenced by securitization activity. In developing a suitable proxy for expected house price appreciation we rely on the central argument of Glaeser, Gyourko, and Saiz (2008) and Saiz (2009), namely, that housing markets which have a high elasticity of supply are able to respond to housing market demand shocks through an increase in the housing stock. In contrast, municipalities with low elasticity of supply are less able to increase the housing stock in response to demand shocks. As such, prices in low elasticity municipalities rise more dramatically than prices in high elasticity areas in the presence of a demand shock.[13]

MSA-specific housing supply elasticity measures are calculated by Saiz (2008) as a function of topology. As stated in his study, "a relative scarcity of developable land can indeed be caused by topographic factors."[14] Time-invariant supply elasticity measures are computed for 95 separate MSA's. We employ the MSA-specific housing supply elasticity measures calculated by Saiz (2008) to create deal-level housing supply elasticity measures in the following way. We assign each individual loan the elasticity measure for the MSA in which the loan resides.[15] MSA-level elasticity measures are only available for 60% of the loans in the sample. Loans with no elasticity measure most frequently reside in rural geographies with high housing supply elasticity. As such, in the spirit of being conservative, we assign the loans in rural geographies with no available elasticity measure the highest elasticity score calculated by Saiz.[16] The loan-level elasticity data are then

[13] Mian and Sufi (2009) also rely on this argument when disentangling a housing price expectations hypothesis from a risk premium hypothesis.

[14] Saiz (2010) further argues, "Coastal cities can hardly expand toward the sea. Major coastal land reclamation projects in areas with an extremely constrained supply of land are very costly and rare. Other geographic constraints to development include abundance of other water bodies such as lakes and rivers (as in New Orleans), heavy slopes and mountainous areas (as in Los Angeles), and wetlands (as in the Miami's Everglades)."

[15] Saiz creates housing supply elasticity measures at the primary-MSA level. Primary-MSA's are nested within MSA's. In circumstances where a loan resides within a primary-MSA that does not have an elasticity score, but is in an MSA that contains a seperate primary-MSA that does have an elasticity score, we assign the loan the average primary-MSA score within a given MSA.

[16] We perform two robustness checks to test how sensitive our analysis is to this approach. First, we assign all missing loans the average elasticity measure. Second, we treat all missing loans as missing, and exclude them in calculating loan-weighted deal-level elasticity scores. Assigning missing observations the average elasticity score does reduce the statistical significance of some of the results. Overall, we find that our results are not qualitatively impacted too substantially using either of the alternative methods.

aggregated to the deal level using loan sizes as weights.

Deal-level housing elasticity measures allow us to estimate the relationship between deal-level collateral concentration and expected rates of house price appreciation. We control for other important collateral characteristics such as deal-level FICO scores, loan-to-value ratios, the percentage of loans with adjustable rates, percent owner occupied, percent second mortgages, percent no documentation or low documentation, and the percentage of loans that are refinances. Estimates of OLS regressions are reported in Table 7. Columns (1) and (2) report results omitting year fixed effects while Columns (3) and (4) report results including year fixed effects. Columns (1) and (3) employ the Herfindahl measure of concentration while Columns (2) and (4) employ the % California measure. Each of the estimates indicates the existence of a strong negative association between housing supply elasticity and deal-level concentration. Concentrated deals are built with mortgage collateral originated in areas with low housing supply elasticity (higher expected price appreciation). The results are of substantial magnitude economically. Using the Herfindahl measure of concentration, a one standard deviation decrease in deal-level housing supply elasticity is associated with a one-half standard deviation increase in deal-level concentration.[17] We interpret these results as being consistent with the important impact that collateral values have in the expected loss of a mortgage pool. Results are similar using the "% loans in California" measure of concentration.

4.4 The Industrial Organization of the Mortgage Market

The results of table 7 are clear in demonstrating that expected collateral values play an important role in explaining where deal collateral is concentrated, when it is indeed concentrated. Further evidence (Table 6) indicates that geographic clustering appears to be a demand-driven phenomenom. However, neither of the results conclusively identifies

[17] The strong statistical relationship between housing supply elasticity and collateral concentration may prompt the concern that the variables are by construction, mechanically, negatively correlated. The relationship is not mechanical. A positive correlation, rather than the observed negative correlation, would exist if deals constructed pools of collateral concentrated in areas of high housing supply elasticity.

causes of cross-sectional differences in concentration. It is likely that many causal explanations of cross-sectional variation exist, two of which we discuss. First, deal arrangers could simply be "taking a view" on the expected path of house prices. The "taking a view" hypothesis is difficult to rule out. Even if views on the housing market differed across underwriters, the optimal expression of such views in a portfolio of arranged deals is unclear ex ante. Are strong views on the direction of the housing market best expressed in every deal, or in a few select deals?

A second hypothesis, one which lends itself more readily to empirical tests, points to factors in the industrial organization of the mortgage origination-securitization market. An established affiliation between a deal arranger and mortgage originator could influence the nature of the collateral employed in a particular deal. Many of the deal arrangers in our sample, particularly commercial banks, have affiliated mortgage origination arms. Securitizing loans originated through an affiliate reduces transactions costs and costs associated with asymmetric information. A consequence, however, of obtaining loans through an affiliate is a reduced ability to actively select the geographic composition of the collateral. Though deal arrangers are free to sell mortgage loans originated through an affiliate and replace them with loans originated through a non-affiliate, doing so increases transactions costs, thereby eliminating a primary benefit of the affiliate relationship. Our hypothesis is that securitizing loans originated solely by an affiliate likely results in more highly concentrated pools, given that originating affiliates are more likely to be concentrated in a given geography, on average. Ultimately however, the impact of originator affiliation on deal-level concentration is an empirical question.[18]

We test whether deal arrangers who securitize mortgage pools originated by affiliates have more highly concentrated collateral. We classify deals as being originated by affiliates

[18] The decision to pool a set of loans originated solely by an affiliate is made at the deal origination level, and should reflect the influence of secondary market demand. Loans originated by a primary market originator that is a subsidiary of, or is affiliated with, a deal underwriter do not always end up as collateral in securitized deals underwritten by their affiliated underwriter. The data support this view. As an example, Chase Mortgage, a subsidiary of JPMorgan Chase, originated loans that appeared in 48 different deals that were not underwritten by JPMorgan Chase.

in the following way. In circumstances where a securitized pool was originated by a single originator, Bloomberg reports the name of the originator. Using a time-series of Dun & Bradstreet publications titled "Who Owns Whom," we determine whether the primary originator of the collateral, as listed by Bloomberg, is a subsidiary or affiliate of the bank responsible for arranging and underwriting the securitization. We then create an indicator variable equal to one for deals with collateral originated by an affiliate. The size of our estimation sample, 915 deals, is dictated by the number of deals for which Bloomberg reports the originator of mortgage collateral and for which we have LoanPerformance collateral-attribute data.

Results are tabulated in Table 8. Column (1) reports the results of an OLS regression with deal-level Herfindahl as the dependent variable and our affiliated indicator variable as the independent variable of interest. We include the deal-level housing supply elasticity variable, the standard set of controls, fixed effects, and clustering of standard errors as in previous tests. The estimated coefficient on the "originator affiliated with underwriter" indicator variable reported in Column (1) is positive, and statistically significant at the 10% level. Collateral originated through an affiliate is 1.8 percentage points more concentrated than non-affiliate collateral (the standard deviation in collateral concentration is 7.8 percentage points). The affiliate indicator variable is positive but not statistically significant using the California measure of concentration.

We test whether affiliated collateral is systematically concentrated in areas of high expected house price appreciation by creating an affiliate-housing supply elasticity interaction term. The affiliate-housing supply elasticity interaction term is highly collinear with the stand-alone affiliate indicator when housing supply elasticity is specified as a continuos variable. To reduce the correlation between the two variables, we construct an above-median housing supply elasticity indicator and interact the above-median indicator with the affiliate indicator. We estimate an OLS regression including the interaction term and the standard set of controls and report results in Column (3) of Table 8. The significant, negative coefficient on the above-median, affiliate interaction term indicates that when affiliated deals are concentrated, they are substantially less concentrated in areas with high housing supply elasticity (more concentrated in areas with low housing supply elasticity - high expected house price appreciation). After controlling for the interaction effect, the positive and significant coefficient on the stand-alone affiliated indicator suggests that underwriters that pool mortgages originated by an affiliate put together deals that are more geographically concentrated, on average. The same pattern of results holds when specifying concentration using the "% in California" measure.

The results indicate that cross-sectional differences in concentration can be explained by differences in methods of obtaining collateral. We stress however that the underwriter/affiliate result cannot be interpreted as conclusive evidence of any one particular hypothesis. The nature of the relationship between a deal underwriter and affiliated mortgage originator is itself not random. The creation of an affiliate relationship could be the mechanism by which deal underwriters "take a view" on expected house price appreciation, or it could be a play to increase efficiencies in obtaining mortgage collateral for securitization deals. Thus, while our tests are able to determine that the industrial organization of the mortgage market has power in explaining cross-sectional variation, we are unable to conclusively identify the economics which might drive an affiliate/underwriting relationship.[19]

5 Is Collateral Correlation Priced?

Tables 2 and 3 present compelling evidence that collateral concentration influences deal-level default rates. The evidence prompts the question, "were investors adequately compensated for this risk at deal origination?" While some argue that the rise in deal default

[19] It is worth noting that we have considered other tests which evaluate the influence of the industrial organization in the origination/securitization market. For example, we investigate whether obtaining collateral through multiple originators has explanatory power. We also consider differences in types of deal underwriting banks (investment banks vs commercial banks). While these additional tests are interesting, they do not distinguish an industrial organization hypothesis from a "taking a view" hypothesis more effectively than the tests we formally describe.

rates is simply the result of a bad draw from the housing market, the stylized model discussed in Section 2 suggests that unconditionally, collateral correlation is harmful to senior tranches. As such, when analyzing the cross-section of deals, senior bonds with more highly correlated collateral had higher ex ante probabilities of default. In this section we analyze whether collateral correlation had any influence on the pricing of subprime bonds.

In order to the answer the pricing question carefully, we require that the bond-transaction sample meet very specific criteria. First, we keep only bonds that traded at par at the time of deal origination or within 12 months of deal origination. Holding bond prices constant allows us to focus on differences in promised coupon payments. Second, we limit the sample to include floating rate bonds only in order to eliminate the influence of changes in interest rates. This ensures that the coupon spread over LIBOR is an adequate measure of the compensation provided investors. Observing transactions at or near origination also ensures that realizations of pre-payments or defaults or other unobserved issues related to deal seasoning have not impacted perceived bond value, though constraining the sample to bonds priced at par should also adequately address the issue.

In a regression framework, we test whether investors in bonds with highly concentrated collateral received a higher coupon spread over LIBOR. In addition to the standard set of deal-level controls included in the default regressions, we include controls that are specific to a bond that might influence coupon spreads. These controls include the amount of subordination and over-collateralization afforded each bond. We attempt to control for differences in expected duration across bonds using the weighted average maturity of the collateral and the percentage of the deal's collateral with prepayment penalties. We include credit rating fixed effects as well as month-specific fixed-effects to control for any time-varying factors that are unobserved.

Table 9 reports the results. The dependent variable is the coupon spread over LIBOR. The sample reported in Columns (1) and (3) include bonds priced at par at origination or within 12 months of origination. Columns (2) and (4) include only bonds priced at par at origination. The estimates indicate that collateral concentration had very little impact on cross-sectional differences in bond coupon spreads. Though the estimated sign is consistent with higher concentration deals paying higher spreads, neither measure of concentration in either specification is statistically different from zero. Very few of the collateral-specific attributes explain differences in spreads, with the exception of FICO scores. Deals with one-standard deviation higher FICO scores are associated with 4 basis point lower spreads. Bond-specific credit ratings, captured with fixed effects, have substantial explanatory power and are responsible for the large R-squared's. The regression results do not provide convincing evidence that investors were compensated for correlation risk at the time of bond origination.

5.1 Detecting the Pricing of Correlation Using Portfolio Sorts

Failure to detect the influence of concentration on bond pricing could be on account of two possibilities. One, correlation risk truly was not priced at the time of bond origination, or two, lack of power in our tests. In an attempt to rule out lack of power, we analyze differences in coupon spreads using two-way portfolio sorts.

Given the results on FICO scores reported in Table 9, we choose to sort on FICO scores and our Herfindahl measure of concentration. We constrain the sample to include bonds priced at par at or within 12 months of origination. We further constrain the sample to include only AAA-rated bonds. Sorts are performed separately on at-par transactions occurring in 2005 and 2006. The 2005 sample has 181 observations that meet the criteria. The 2006 sample has 261 observations. Table 10, panel A reports the results arising from a simple one-way FICO sort. Deal-level FICO scores are first sorted into quartiles. We report the average coupon spreads paid to investors on bonds purchased at par for the highest quartile and the lowest quartile. The difference of 6.2 basis points between the high and low FICO quartiles, with an associated t-statistic of 3.57 provides some evidence that sorts have power to pick up differences in pricing on account of deal attributes.

Panel B of Table 10 reports results arising from two-way conditional sorts. We first sort deal-level FICO scores into quartiles. Within each FICO quartile we sort Herfindahl measures into quartiles. We then report the coupon spread for deals that fall into each of the 16 buckets (with an average of 11 and 23 observations per bucket for the 2005 and 2006 samples, respectively). Within each FICO quartile, differences in coupon spreads for the high concentration minus low concentration portfolios are negative on average, though none are statistically different from zero, suggesting that bonds in the high concentration portfolios did not compensate investors for concentration risk. The low FICO portfolios have 6 basis point higher spreads on average than high FICO buckets, and three of the four high minus low FICO differences are statistically different from zero. This result suggests two things. First, consistent with the regression results, investors were compensated for low-FICO risk. Second, the portfolio sorts do display some power in detecting differences in pricing based on collateral attributes.

Panels C and D of Table 10 repeat the one and two-way portfolio sorts using AAA-rated bonds from 2006. The results are less conclusive than 2005 tests. The high minus low correlation difference within the lowest FICO quartile does reveal a statistically significant concentration premium of 5 basis points, though the average across all quartiles is slightly negative, and not significant. Surprisingly, low FICO portfolios did not pay a premium over high FICO portfolios in the 2006 sample.

Table 11 performs similar one and two-way sorts, conditioning on the Herfindahl measure of concentration first, instead of FICO scores. Panels A and B of Table 11 employ the AAA-rated 2005 sample, with panels C and D reporting results from the AAA-rated 2006 sample. Conditioning first on concentration does not change our inference. In 2005, FICO scores pay some premium on average, with two of the four bins statistically different from zero. High concentration portfolios do not pay a premium over low concentration portfolios in the 2005 sample. The results in the 2006 sample are even less conclusive.

Regression results and two-way conditional sorts do not provide evidence that investors in subprime bonds were compensated ex ante for correlation risk. Lack of pricing evidence could be the result of poor power. However, the results on coupon discounts for low-FICO deals provide some confidence that the regression tests and portfolio sorts have sufficient power to detect differences in pricing on account of difference in collateral attributes. We conclude that it is unlikely that investors in subprime bonds at the time of deal origination were compensated for correlation in deal collateral. The regression results provide evidence that bond credit ratings were most responsible for explaining cross-sectional differences in coupon spreads. Aside from small differences in coupons based on FICO scores, investors apparently priced bonds based on credit ratings, at least at deal origination. This result was reported by Coval, Jurek, and Stafford, 2009.

6 Conclusion

Prioritized cash flow rules in structured finance create securities which fail to pay off in the worst states of the world. We document that the exposure of structured finance securities to economic risk can be influenced substantially by the nature of collateral. In particular, correlated collateral can exacerbate the impact of economic catastrophes.

Our results also contribute to our understanding of the economics of securitization. To date, the literature has focused on how securitization altered lenders' screening incentives at the point of loan origination. Indeed, the originate-to-distribute model resulted in the origination of poor-quality credit (Purnanandam (2010)). In this paper we demonstrate that once originated, the manner in which loans were pooled had a substantial impact on deal-level default rates. In a cross-section of deals, pools with geographically concentrated collateral defaulted at substantially higher rates during the financial crisis. In fact, only FICO scores and LTV ratios have a larger impact on deal-level defaults. We feel this result is important given the belief that the explosion in mortgage default rates represents a bad draw from a distribution of potential outcomes in the housing market. The structure of deals themselves contributed to the explosion in defaults at the deal-level.

In explaining why collateral is concentrated, we find that secondary market demand is at play as opposed to a primary market supply explanation. Perhaps not surprisingly, given the contractual features of subrprime mortgages, when concentrated, mortgage collateral is concentrated in areas with high expected rates of house price appreciation (as proxied by housing supply elasticities). Finally, we show that collateral obtained through an affiliated, single-name originator results in more concentrated deals.

We are unable to find pricing effects in OLS regression tests as well as two-way conditional portfolio sorts. We fail to find any evidence indicating that investors in bonds with more highly concentrated collateral received higher coupon payments. Lack of evidence could be the result of a lack of power in our tests, although our tests are able to discern pricing differences on account of differences in deal-level FICO scores.

Ultimately, this paper demonstrates that the manner in which a securitization deal is put together can have a substantial impact on that deal's performance. Many important questions remained unanswered. Did deal underwriters actively seek collateral along other observable dimensions in a manner which influenced default outcomes? Did securitization activity itself, particularly the demand for mortgage collateral from concentrated geographies, influence the realized housing market outcome and contribute to the financial crisis? How carefully did market analysts, particularly rating agencies, account for the influence of securitization activity itself on the future path of housing prices? Finally, we have considered the issue of correlation from the standpoint of senior bond holders. Correlated collateral *reduces* risk for the most subordinated bonds in structured finance. At what point in the capital structure of a structured finance deal does correlation turn from being harmful to helpful? Calculating the total welfare implications of collateral correlation requires a very specific identification of what bonds in a structured finance deal suffer from correlation and what bonds thrive under it. We leave these questions to future research.

References

[1] Barclays Capital. 2007. "Correlation Modelling – from Vanilla to Exotic", Structured Credit Strategist, April 19, 2007.

[2] Coval, J., Jurek, J., Stafford, E., 2009. "Economic Catastrophe Bonds," *American Economic Review* 99, no. 3 (June 2009).

[3] Coval, J., Jurek, J., Stafford, E., 2009. "The Economics of Structured Finance," *Journal of Economic Perspectives* 23, no. 1 (winter 2009): 3-25.

[4] Cowan, A., Cowan C., 2004. "Default Correlation: An Empirical Investigation of a Subprime Lender" *Journal of Banking and Finance*, 28, pg. 753-771.

[5] Dell'Ariccia, G., Igan, D., Laeven, L., 2008. "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market," IMF Working Paper.

[6] Demyanyk, Y.,Van Hemert, O., 2008. "Understanding the Subprime Mortgage Crisis," *Review of Financial Studies.*

[7] DeMarzo, P. 2005. "The Pooling and Tranching of Securities: A Model of Informed Intermediation," *Review of Financial Studies*, 18, 1-35.

[8] Deng, Y., Quigley, J., Van Order, R., 2000. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options," *Econometrica*, 68, 275-307.

[9] Diamond, D. Rajan, R. 2009. "The Credit Crisis: Conjectures about Causes and Remedies", University of Chicago GSB Working Paper.

[10] Gaur, V, Seshadri, S., Subrahmanyam, M., 2003. "Market Incompleteness and Super Value Additivity: Implications for Securitization," mimeo, New York University.

[11] Gerardi, Lehnert, Sherlund, and Willen 2009. "Making Sense of the Subprime Crisis." Brookings Panel on Economic Activity, forthcoming.

[12] Gerardi, K., Shapiro, A., and Willen, P., 2007. "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures," Federal Reserve Bank of Boston Working Paper No. 07-15.

[13] Glaeser, E., Gyourko, J., and Saiz, A., 2008. "Housing Supply and Housing Bubbles," Harvard University Working Paper.

[14] Gorton, G. 2010, "The Panic of 2007," Yale University and NBER Working Paper.

[15] Gorton, G., Metrick, A. 2010. "Securitized Banking and the Run on the Repo," Yale University Working Paper.

[16] Ho, G., and Pennington-Cross, A., 2006. "The Termination of Subprime Hybrid and Fixed Rate Mortgages," Federal Reserve Bank of St. Louis Working Paper, 2006-042A.

[17] Keys, B., Mukherjee, T., Seru, A., Vig, V., 2010a. "Securitization and Screening: Evidence From Subprime Loans 2001-2006," *Quarterly Journal of Economics,* 125(1).

[18] Keys, B., Seru, A., Vig, V., 2010. "Lender Screening and the Role of Securitization: Evidence from Prime and Subprime Mortgage Markets, forthcoming *Review of Financial Studies.*

[19] Krishnamurthy, A. 2008, "The Financial Meltdown: Data and Diagnoses," Working Paper.

[20] Krugman, P. 2007, "Just Say AAA", New York Times Op. Ed., July 2, 2007.

[21] Lando, D. and Stenbo Nielsen, M. 2010. Correlation in Corporate Defaults: Contagion or Conditional Independence?. Unpublished Working Paper.

[22] Longstaff, F., Arvind R., 2008. "An Empirical Analysis of the Pricing of Collateralized Debt Obligations," *Journal of Finance*, 43(2) pg. 529-563.

[23] Longstaff, F., Rajan., 2008 Correlation and CDOs. working paper.

[24] Markowitz, H, 1952, Portfolio Selection, *Journal of Finance*, 7: 77-91.

[25] Mayer, C. Pence, K., 2008. "Subprime Mortgages: What, Where and to Whom?" NBER Working Paper #14083.

[26] Mayer, C. Pence, K. and Sherlund, S. 2009. "The Rise in Mortgage Defaults." *Journal of Economic Perspectives*, 23, 27-50.

[27] Merton, R. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance*, Vol. 29, No. 2.

[28] Mian, A. and Sufi, A., 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the 2007 Mortgage Default Crisis," *Quarterly Journal of Economics*, forthcoming.

[29] Moody's Investor Services, 2004. "Introduction to Moody's Structure Finance Analysis and Modelling." Presentation given by Frederic Drevon, May 13.

[30] Nadauld, T., and Sherlund, S. 2009. "The Impact of Secondary Mortgage Market Activity on the Expansion of Subprime Credit", Working Paper.

[31] Pagano, M. and Volpin, P. 2008. "Securitization, Transparency, and Liquidity," Working Paper.

[32] Purnanandam, A., 2011. "Originate-to-Distribute Model and the Subprime Mortgage Crisis," *Review of Financial Studies*, 24(6), 1881-1915.

[33] Rajan, U., Seru. A., and Vig, V. 2010. "The Failure of Models That Predict Failure: Distance, Incentives, and Defaults," Working Paper.

[34] Siaz, A., 2009. "On Local Housing Supply Elasticity," Wharton Working Paper, 2008.

[35] Sherlund, S., 2008, "The Past, Present, and Future of Subprime Mortgages," *Finance and Economic Discussion Series 2008-63,* Federal Reserve Board.

[36] Shiller, R., 2007. "Understanding Recent Trends in House Prices and Home Ownership," Cowles Foundation Discussion Paper No. 1630.

[37] Standard and Poor's, 2007. "Principles-Based Rating Methodology for Global Structured Finance Securities." Standard & Poor's Ratings Direct Research, N.Y.

7 Appendix

7.1 Appendix A.1 Sample Construction

The RMBS deal-level data relies on the intersection of two datasets provided by LoanPerformance and ABSNet. LoanPerformance, a subsidiary of First American Trust, reports borrower attributes and loan information for about 75% of all subprime mortgage loans sold into securitization deals.[20] The database contains detailed information on subprime loans at the time of origination and tracks the performance of individual loans through time, a feature we exploit in computing changes in deal-level default rates through time. Most relevant for our purpose is loan-level information on the original loan balance, FICO score, loan-to-value ratio, debt-to-income ratio, and loan type. Our house price measures employ the LoanPerformance repeat-sales house price index, which is available at the zip code level. If a ZIP-code index is not available, we use MSA and state-level repeat sales indexes. Also crucial to our analysis is the location of the borrower, which is reported at the ZIP-code level.

Deal level summary data come from ABSNet, a subsidiary of Standard and Poor's. ABSNet reports important, deal-level summary information for the universe of subprime RMBS. Important summary information includes the origination date, origination amount, deal underwriters, original credit ratings, original principal amounts of each tranche, tranche CUSIPs (each tranche, or bond, has a unique CUSIP), and credit enhancements. Figure 2 provides an example of a deal structure from our sample. The deal, originated by Goldman Sachs in 2006 has total principal of $714.2 million, carved into 14 bonds and an equity tranche. The tranche-weighted coupon spread over LIBOR for the deal was .258%. All of this information is reported in a deal summary by ABSNet.

In analyzing the relationship between default rates and collateral correlation, our primary unit of analysis is at the deal level. We take the following steps to identify and

[20] The coverage of LoanPerformance varies by year, but it is more complete in the later years of our sample.

aggregate individual residential subprime loan data to the deal level. First, we obtain the deal summary for residential mortgage-backed securitization deals originated between 1998 and 2007 from ABSNet. Because ABSNet does not classify the residential securitization deals as being subprime, we rely on the classification of subprime loans provided by LoanPerformance. No unique identifier exists between the deal summary data from ABSNet and the LoanPerformance database, so we match by hand using deal names. The total number of subprime deals included in our sample is dictated by the number of subprime deals in the LoanPerformance database that can be matched to the universe of ABSNet, which totals 1,254 subprime deals.[21] We double check that our matching process correctly matched the LoanPerformance and ABSNet data by examining a sub-sample of deal names and deal summaries from Bloomberg. The median securitization deal in our sample has 5,219 mortgage loans serving as collateral. In our deal-level analysis, we control for attributes of the loan collateral by computing deal-level summary measures. We aggregate the loan attributes to the deal level by summing over the loan-weighted average attribute of each loan in a deal. Weights are determined by the size of each individual loan relative to the size of the total deal. The loan level data for attribute i of loan k in deal j is aggregated to the deal level as follows: $LoanAttribute_{i,k,j} = Attribute_{i,k} * \left[\frac{principal_k}{\sum principal_k}\right]$.

The final deal-level data set includes 1,254 securitization deals and summary measures of the over 6 million loans that serve as collateral in the deals.[22]

Our analysis in Section 4 requires data on the identity of loan originators. Bloomberg reports the specific name of the loan originator for a given securitization deal if the bulk of the loans in a single deal were originated by the same loan originator. For deals with multiple originators, Bloomberg reports the originator as "multiple." We are able to obtain information on loan originators for 915 of the 1,254 deals in our final sample. In deals

[21] In addressing potential concerns about whether our final sample of deals is systematically biased in any way, we conclude that our sample likely under-represents deal activity that occurred early in our sample period on account of less complete coverage of subprime activity by LoanPerformance.

[22] Our analysis of deal-level credit ratings relies on a sample of 1,225 deals. This is the total number of deals we were able to match that have the full set of information on tranche credit ratings, subordination, and other relevant deal-level data required in the estimation.

where the collateral arise from a single originator (645 deals), we determine whether the originator is affiliated with the deal underwriter in any way using Dun and Bradstreet's "Who Owns Whom."

Our analysis of correlation and pricing takes place at the bond-level.[23] Bond-level transaction data are compiled and provided by Thomson Reuters EMaxx services. The observations represent open-market transactions primarily of insurance funds, but include transactions of government state pension funds and mutual funds. Insurance funds file quarterly reports detailing their transactions and asset holdings with the National Association of Insurance Commissioners (NAIC). Thomson Financial data analysts gather data on the transactions of government state pension funds, and mutual fund data are gathered through Thomson's relationship with Lipper. The bond price data includes the bond's credit rating at the time of the transaction, the transaction date, and price. We match the bond-level transaction data with deal-level securitization data by CUSIP. We restrict our sample to only include transactions on floating rate bonds purchased at par either at deal origination, or at par within 12 months of deal origination. The at par and at-or-within 12 months of origination sample includes 1,234 unique bond-transactions, while the at-origination only sample includes 745 unique bond-transactions. Although we cannot use the full sample of transaction prices in our analysis, in Figure 5 we provide a plot of the prices paid on subprime bonds in the open market from 2000-2009. The decline in prices paid for subprime bonds through the heart of the financial crisis is visually stunning.

7.2 Appendix A.2 Correlation Model

In this appendix we present a simple model to demonstrate how correlation and default relate in a simple securitization framework. Our model is a simplified version of Coval, Jurek, and Stafford's (2009a), hereafter CJS (2009a), model, in that this model will make specific assumptions regarding the correlation between assets while for the most part their

[23] A subprime securitization deal produces about 17 bonds, on average, but the bonds share the same collateral. In the pricing analysis we focus on specific bonds arising from a deal.

model leaves correlation unspecified.[24]

Following CJS (2009a) we assume that a deal is constructed using N underlying assets; subprime mortgages in our case. These N assets are pooled together and, for simplicity, have an aggregate notional amount of \$1. The deal's capital structure is defined as a standard series of tranches which we formalize later. To introduce correlation, again following CJS (2009a), we adopt a conditional independence approach, common to both the academic literature and industry practice (Lando and Nielsen (2010) and Barclays Capital (2007)). We introduce the conditional independence assumption in determining the value of each individual mortgage. Specifically, the mortgage value for security i (X_i) is assumed to follow:

$$X_i = \sqrt{\rho}F + \sqrt{1-\rho}\varepsilon_i, \tag{1}$$

where F represents a factor influencing the prices of the underlying mortgages in the pool. For example, in this paper's analysis of subprime mortgages as assets, F can be thought of as a local macroeconomic factor that would influence all of the underlying mortgage values of a deal constructed using a pool of mortgages in any geographic location. For simplicity, we assume that F follows a $N(0,1)$ distribution.[25] As seen from equation (1), the value of each asset, X_i, is also assumed to have an asset specific component represented by ε_i, which also follows a $N(0,1)$ distribution. The asset-specific component ε_i is assumed to be independent of all other asset specific components (i.e. $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$) and independent of the factor F (i.e. $\text{cov}(\varepsilon_i, F) = 0$).

The final term in equation (1) is ρ, which determines the correlation between the mortgage values in the model. In particular, ρ helps to determine the relative influence

[24] Coval, Jurek, and Stafford (2009a) do not explicitly study collateral correlation. A refinement in their model, one they briefly explore, does allow for the possiblity of correlation in the underlying collateral through collateral's dependence on a common factor. This model is also related to the mixed binomial model found in Lando (2004).

[25] This assumption describing mortgage values simplifies the default assessement of deal tranches, but this assumptions has weaknesses, primarily in that this specification allows mortgage values to be negative (potentially infinitely negative).

between the systematic component of variation (F) and the asset-specific component (ε_i). Similar to a standard correlation coefficient, we bound ρ to lie between 0 and 1. It is straightforward to show, from equation (1), that the covariance between any two assets i and j is ρ. Allowing ρ to take on a range of values will allow us to consider a variety of dependence cases. For example, a structured deal that pools mortgages across the whole geography of the US will have less dependence on a Florida-specific factor and consequently in this case would be modelled with a low ρ, while a deal that pools only Florida loans will have a strong common dependence on F and consequently would be modelled with a large ρ. Our assumptions regarding F, ε_i, and ρ also lead to the asset value X_i following a $N(0,1)$ distribution which simplifies the process of determining default probabilities on various tranches of a deal.

We follow Merton (1974) in determining the default of any given asset. In particular, an asset will be deemed in default if its value drops below some level D. The probability of default for a given asset, denoted as p, can be shown as:

$$
\begin{aligned}
p &= E\left[P\left(X_i < D\right)\right] && (2)\\
& E_F\left[P\left(X_i < D \mid F\right)\right] && (3)
\end{aligned}
$$

by the law of iterated expectations, and where $E_F[\cdot]$ corresponds to integrating over the distribution of factor F. By first conditioning on F, the conditional probability values of default can be drawn from a normal distribution regardless of the influence of F.

From equation (1), we know that the conditional probability of default, p_F:

$$
\begin{aligned}
p_F &= P\left(X_i < D \mid F\right) \\
&= P\left(\sqrt{\rho}F + \sqrt{1-\rho}\varepsilon_i < D \mid F\right) && (4)\\
&= P\left(\varepsilon_i < \frac{D-\sqrt{\rho}F}{\sqrt{1-\rho}} \mid F\right) \\
&= N\left(\frac{D-\sqrt{\rho}F}{\sqrt{1-\rho}}\right), && (5)
\end{aligned}
$$

where $N(.)$ denotes the Normal distribution function.

Similar to CJS (2009a), we adopt a standard capital structure by defining a sequence of fractions $(\lambda_1, \lambda_2, \lambda_3, ...\lambda_k)$, such that $(\lambda_1 < \lambda_2 < \lambda_3 < ... < \lambda_k)$ which represent the attachment points of each respective tranche. Each attachment point, λ_i, corresponds to the percent of underlying asset defaults that would trigger a default on that particular tranche. For example, tranche 1 corresponds to attachment point λ_1 and would be considered in default if $n_1 = \lambda_1 * N$ assets are in default.[26] Given the relationship between the attachment points, the higher the subscript, the more senior the tranche (the more mortgages that must default to make that corresponding tranche default). To calculate the probability of a tranche default, we need to calculate the probability of the number of assets corresponding to a tranche defaulting. Consequently, for the λ_1 tranche, the probability of n_1 assets defaulting in the pool of N assets can be calculated using the binomial distribution function. We denote this probability as $p_F^{\lambda_1}$, which can be calculated as follows:

$$p_F^{\lambda_1} = \begin{pmatrix} N \\ n_1 \end{pmatrix} p_F \left(1 - p_F\right)^{N-n_1}, \tag{6}$$

where p_F is the conditional (on F) default probability from equation (5). To calculate the unconditional probability of default, p^{λ_1}, we only need to integrate over the density of F,

$$p^{\lambda_1} = P\left(X_i < D\right) = \int p_F^{\lambda_1} g(F) dF, \tag{7}$$

with $g(F)$ denoting the probability density of F which in our case is assumed to be a normal density. Equation (7) demonstrates that the default probabilities are readily obtainable because, conditional on the factor, all asset values are independent normal variables.[27] This simple model allows us to determine default probabilities for various

[26] Our model ignores some of the common subtleties in the bylaws (cash flow rules, performance triggers, and prepayment rates) of structured finance deals, which would also play an important role in determining the sequence of defaults.

[27] Other approaches to modelling correlation between assets exist (i.e., contagion) and may provide a

points in the capital structure while varying the correlation between the underlying mortgages. Though the model is relatively simple, the impact of correlation across tranche types is subtle as can be seen in Figure 1 which plots the expected default across the spectrum of tranches (λ) and for a range of correlations in the underlying mortgages. Our interest is on the senior tranches of a securitization deal, and as can be seen in Figure 1, for senior tranches the relationship between correlation and expected default is positive.

richer relationship between asset values than this model.

Figure 1. The Relationship Between Collateral Correlation, Tranche Seniority, and Default Probability.

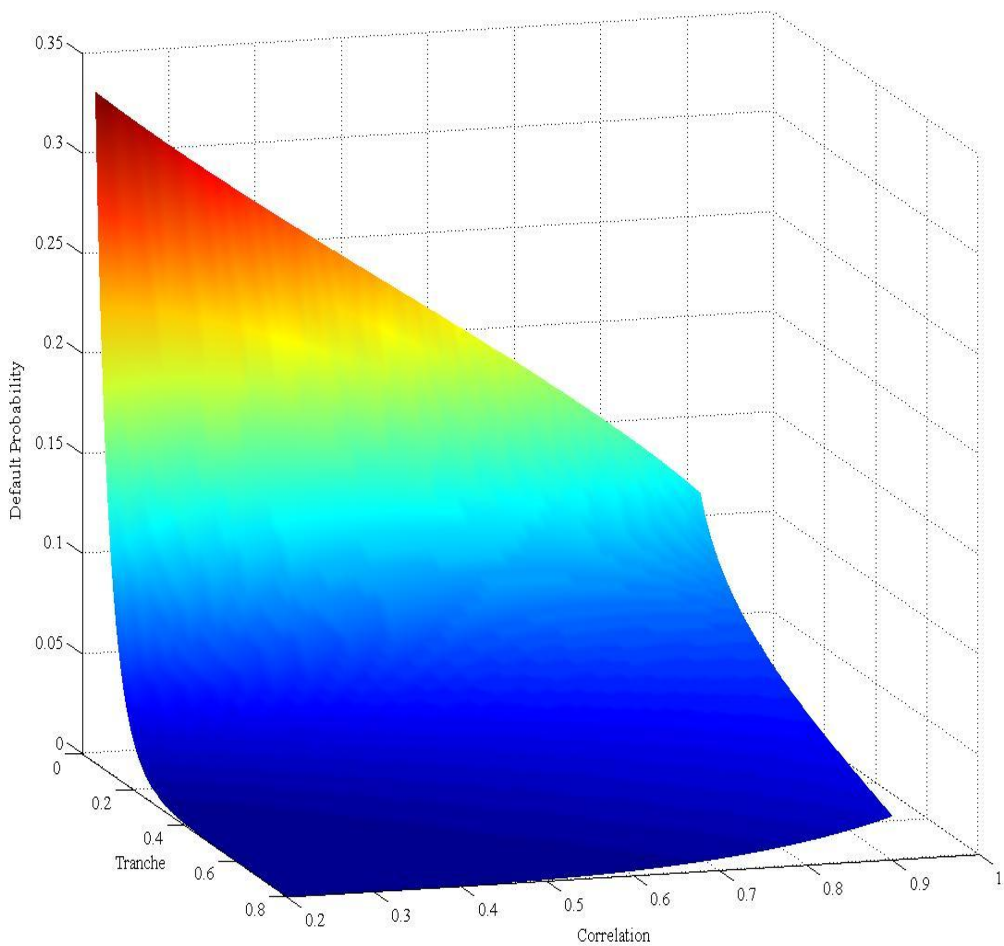
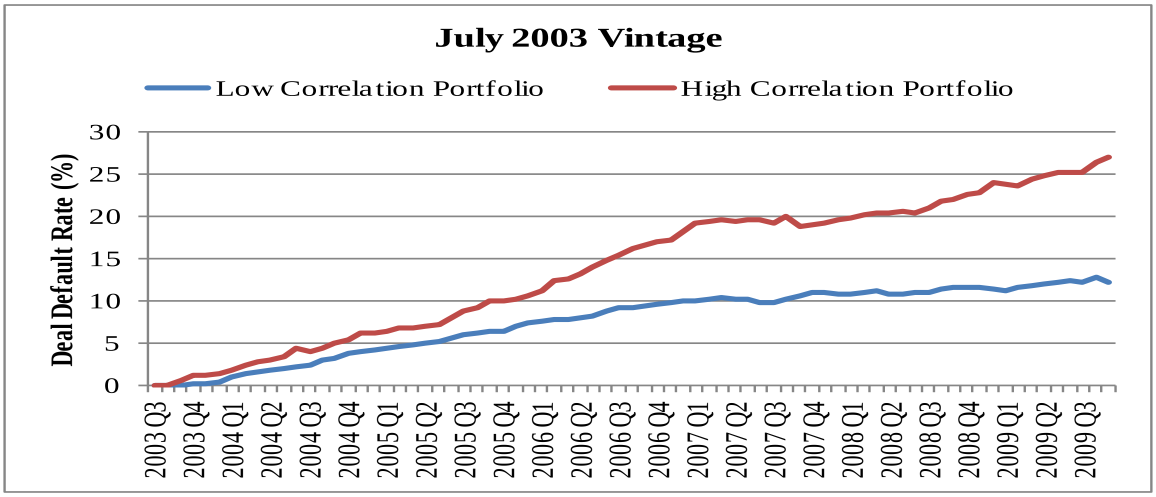


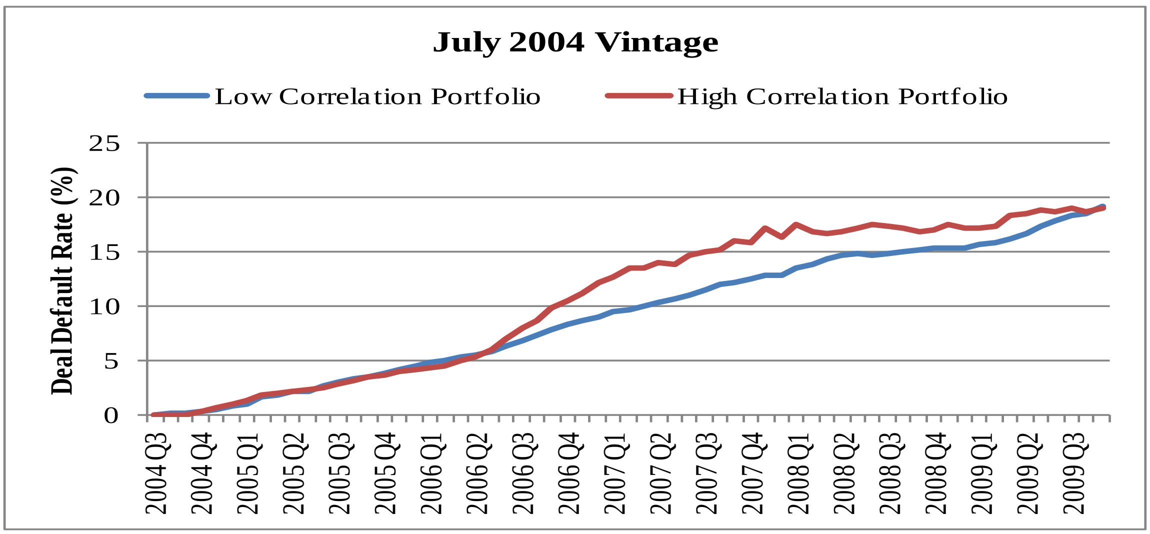
Figure 2. The Structure of a Subprime Securitization Deal.

This table presents the structural details of a subprime securitization deal in our sample. The deal, named GSAMP Trust 2006-NC1 was arranged by Goldman Sachs and was issued in February 2006. The total deal principal is $714.2 million, with $8.93 million serving as overcollateralization. One-month LIBOR in February 2006 was 4.58. The tranche-weighted spread over LIBOR is .258%.

Class Name	Original Balance (000's)	Percent of Total Principal	Original Rating (S&P)	Coupon Type	First Coupon Rate	Spread Over 1-Mth LIBOR	Spread over Reference Treasury
A-1	$310,299	43.4%	AAA	Floating	4.675%	0.095%	
A-2	$224,955	31.5%	AAA	Floating	4.785%	0.205%	
A-3	$42,565	6.0%	AAA	Floating	4.895%	0.315%	
M-1	$23,213	3.3%	AA+	Floating	4.965%	0.385%	
M-2	$21,784	3.0%	AA+	Floating	4.985%	0.405%	
M-3	$12,857	1.8%	AA	Floating	5.005%	0.425%	
M-4	$11,070	1.5%	AA	Floating	5.105%	0.525%	
M-5	$10,714	1.5%	AA-	Floating	5.125%	0.545%	
M-6	$9,642	1.3%	A+	Fixed	6.000%		1.430%
B-1	$9,285	1.3%	A	Floating	5.205%	0.625%	
B-2	$10,000	1.4%	A-	Floating	5.745%	1.165%	
B-3	$6,428	0.9%	BBB+	Fixed	6.000%		1.530%
B-4	$5,357	0.8%	BBB	Fixed	6.000%		1.530%
B-5	$7,142	1.0%	BBB-	Fixed	6.000%		1.530%
CE	$8,928	1.3%	N.R.				

Figure. 3. Vintage Default Rates by High (75th%tile) and Low (25th%tile) Collateral Correlation.





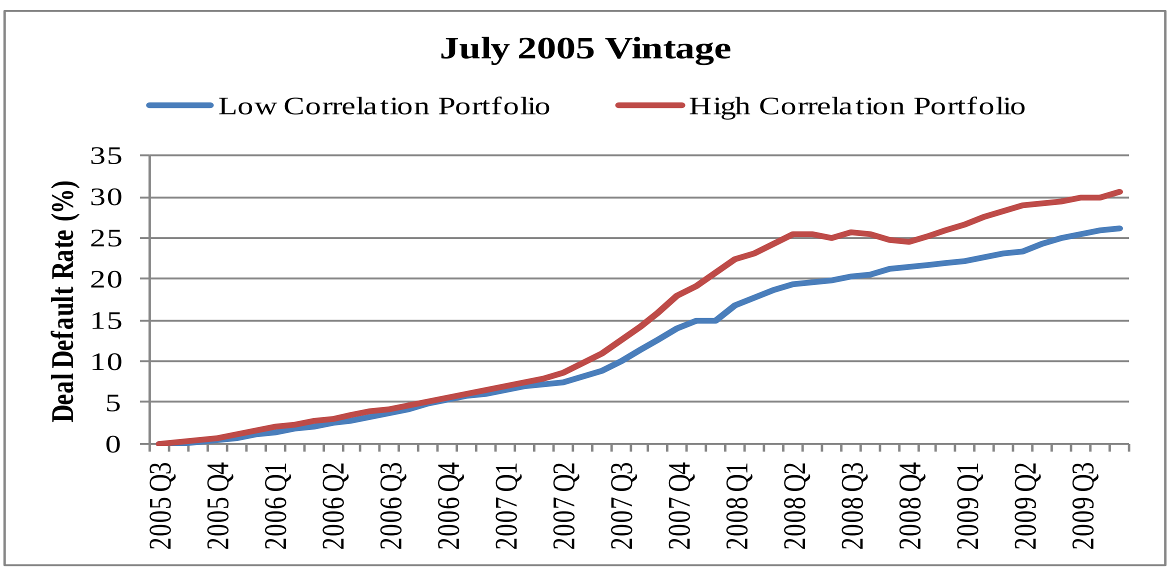


Figure 4. Plotting the Aggregate Difference in Deal-Level Default Rates for Highly Concentrated Deals vs Low Concentrated Deals.

The solid line represents the within-vintage difference in default rates for highly-concentrated deals compared to low-concentrated deals. Differences in vintage-default rates are calculated as follows. Within each month of our sample, beginning in January 2003, we sort vintages with at least 8 deals into quartiles of collateral concentration. Within each vintage we calculate the difference between average defaults rates for deals in the highest concentration portfolio and subtract average default rates for deals in the lowest concentrated portfolio. As of January 2003, there were only 7 vintages that met the 8 deal criteria. Thus, the data point in January 2003 represents differences in default rates between high and low concentrated deals for 7 vintages worth of deals as of January 2003. As of January 2007, 55 vintages are in the sample. Differences in default rates on account of seasoning are accounted for by comparing differences *within* vintage. We construct confidence intervals using the variance in differences in default rates between high and low concentrated deals.

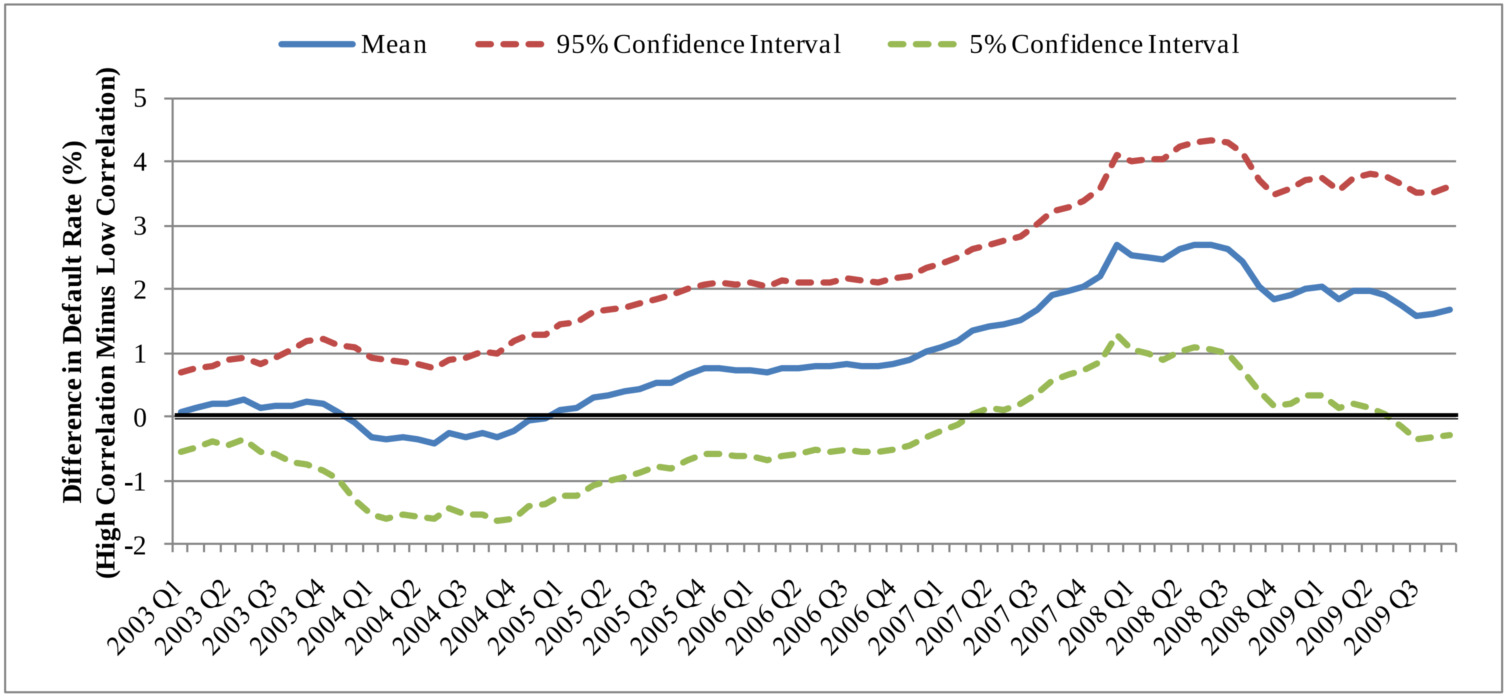


Figure 5. Transaction Prices on Subprime Bonds Through Time.

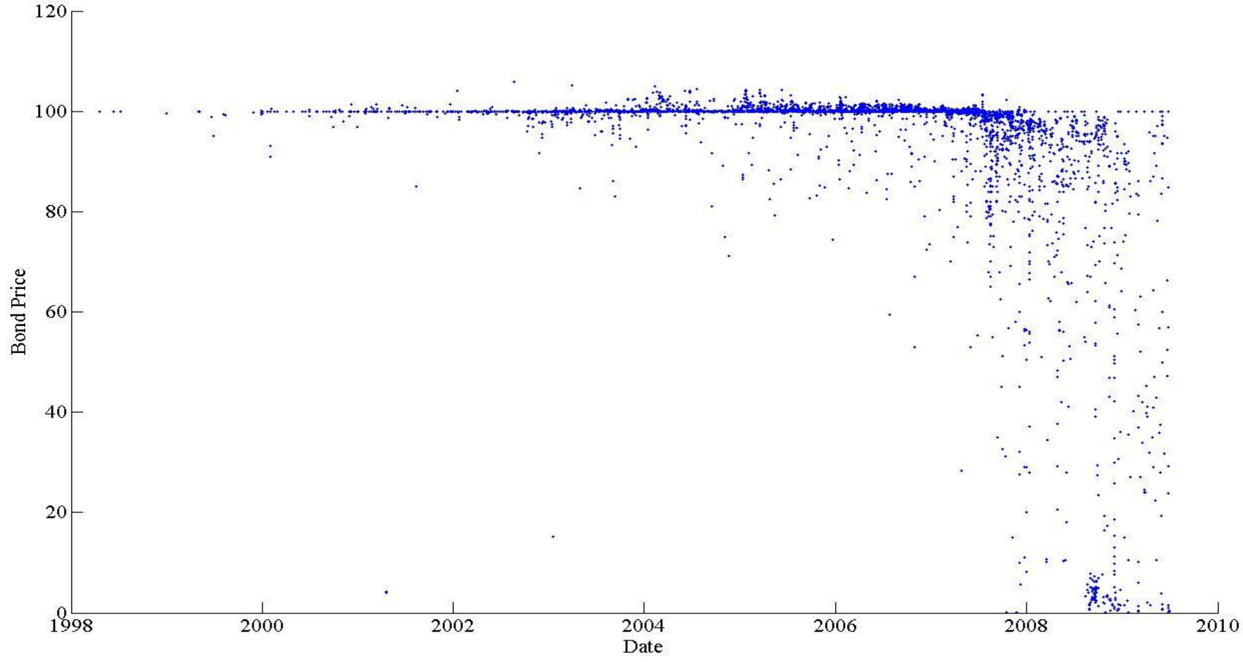


Table 1. Summary Statistics on Deal-Level Collateral.

Panel A tabulates summary statistics on the collateral attributes of subprime mortgage-backed securitization deals through time. Loan-level attributes are aggregated to the deal-level using loan size as weights. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Collateral in California* calculates the number of loans (equal weighted) in each deal originated in the state of California. *12-mth, 24-mth, and 36-mth default rates* refer to deal-level default rates, calculated as the total number of defaults in a deal divided by the total loans outstanding 12-, 24-, and 36-months after deal origination. *Deal Housing Supply Elasticity* is calculated as follows. Each individual loan is assigned a housing supply elasticity measure, as calculated by Saiz (2009), for the MSA in which it resides. Loan-level elasticity measures are aggregated to the deal level using loan sizes as weights. We discuss the handling of loans with missing elasticity measures in the text. *Deal H.P.A.* measures loan-weighted, deal-level rates of house price appreciation in the year prior to deal close. Individual loans are matched with ZIP, MSA, or state-level house price indexes depending on index availability for the geographic area in which the loan was originated. *Deal FICO* measures loan-weighted, deal-level FICO scores. *Deal CLTV* measures loan-weighted, deal-level combined loan-to-value ratios at the time of loan origination. *Deal % A.R.M.* calculates the loan-weighted percent of loans in the deal with adjustable-rate features.

Panel A

			Herfindahl		% Collateral in California		12-mth Default Rate		24-mth Default Rate		36-mth Default Rate	
Year	N		Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
1998	3		0.093	0.066	0.221	0.120	5.23	4.17	11.04	5.95	17.01	9.40
1999	11		0.083	0.052	0.185	0.093	6.20	5.44	9.69	7.48	13.49	9.45
2000	15		0.094	0.046	0.154	0.122	6.74	5.53	13.28	9.84	18.34	13.44
2001	31		0.101	0.038	0.215	0.108	8.98	5.56	16.70	8.25	24.50	11.18
2002	68		0.137	0.070	0.299	0.125	4.99	3.59	11.61	6.43	17.24	8.74
2003	135		0.151	0.083	0.314	0.135	3.98	2.98	8.61	5.05	13.31	7.15
2004	233		0.165	0.093	0.335	0.136	3.77	2.08	9.02	3.77	14.93	6.33
2005	302		0.154	0.086	0.311	0.133	4.71	2.39	12.37	5.19	21.98	8.91
2006	321		0.139	0.064	0.289	0.111	8.77	4.89	20.08	9.28	29.27	13.58
2007	135		0.118	0.049	0.252	0.096	12.77	6.40	24.08	11.29	--	--
Total	1254	Avg.	0.144	0.078	0.297	0.127	6.574	5.008	14.516	9.045	20.387	11.323

Table 1 continued…

Panel B												
			Deal Housing Supply Elasticity		Deal H.P.A.t-1		Deal FICO		Deal CLTV		Deal % A.R.M.	
Year	N		Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
1998	3		1.49	0.26	5.14	0.94	612.88	10.13	71.88	8.64	0.252	0.279
1999	11		1.50	0.19	6.46	1.29	598.20	20.45	77.09	4.00	0.505	0.195
2000	15		1.51	0.16	7.69	1.27	591.29	31.38	77.57	4.24	0.492	0.287
2001	31		1.48	0.11	9.55	1.21	597.77	19.24	78.52	2.55	0.652	0.287
2002	68		1.43	0.14	8.93	1.23	613.57	24.41	79.76	4.84	0.717	0.184
2003	135		1.41	0.15	10.84	2.05	622.89	28.75	80.26	8.48	0.676	0.170
2004	233		1.40	0.14	13.84	2.39	622.46	16.60	83.47	4.54	0.765	0.142
2005	302		1.42	0.13	16.64	2.27	628.73	20.17	85.20	5.64	0.841	0.129
2006	321		1.43	0.13	14.56	2.95	626.56	19.27	86.30	4.86	0.846	0.099
2007	135		1.45	0.12	6.40	3.15	624.30	19.93	85.01	5.58	0.754	0.187
Total	1254	Avg.	11.202	6.296	13.044	4.169	623.563	21.888	83.957	6.058	0.781	0.169

Table 2. Deal-Level Default Rates Sampled *Prior* to January 2007.

This table estimates the relationship between deal-level default rates and deal-level attributes using an OLS regression. The dependent variables, *12-mth, 24-mth, 36-mth, and 48-mth default rates* refer to deal-level default rates, calculated as the total number of defaults in a deal divided by the total loans outstanding ,12, 24, 36, and 48 months after deal origination. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Collateral in California* calculates the number of loans (equal weighted) in each deal originated in the state of California. *Deal H.P.A.* measures loan-weighted, deal-level rates of house price appreciation in the year prior to deal close. Individual loans are matched with ZIP, MSA, or state-level house price indexes depending on index availability for the geographic area in which the loan was originated. *Cumulative H.P.A.* measures cumulative, deal-level rates of house price appreciation from the time of deal origination to the month in which the default rate is being measured. Deal-level measures of collateral quality are aggregated to the deal level using loan size as weights. Collateral quality measures include *FICO* scores, *combined loan-to-value ratios, % loans with adjustable-rate features, % loans owner occupied, % loans second mortgages, % low doc/no doc, % refinancing.* Our sample includes deals originated between 1998 and 2007. The sample is a panel which calculates deal-level default rates each month beginning in the month of deal inception. The sample used in this table only includes deal default rates measured *prior* to January 2007. We include month fixed effects and cluster standard errors by month. The symbols ***, ** and * indicate significance at the 1, 5, and 10% levels, respectively.

Sample: 12, 24, 36, and 48-month deal-level default rates measured ***prior*** to Jan. 2007								
	Dependent Variable: Deal-Level Default Rates							
	Months Seasoning From Deal Origination				Months Seasoning From Deal Origination			
	12-mth	24-mth	36-mth	48-mth	12-mth	24-mth	36-mth	48-mth
Herfindahl (Geographic Concentration)	0.519	-1.130	1.120	0.981				
	(0.484)	(-0.406)	(0.159)	(0.0739)				
% Loans in California					2.450**	3.568	8.918**	14.850**
					(2.544)	(1.372)	(2.062)	(2.043)
Deal-Level Measures of Collateral Quality:								
Deal-Level H.P.A.$_{t-1}$	-0.122***	-0.142	-0.356	1.056	-0.216***	-0.337**	-0.883**	0.260
	(-2.715)	(-1.381)	(-0.857)	(1.452)	(-4.070)	(-2.539)	(-2.259)	(0.356)
Deal-Level Cumulative H.P.A.	0.012***	-0.001	-0.007	-0.012	0.012***	-0.002	-0.010	-0.017
	(3.109)	(-0.155)	(-0.545)	(-0.358)	(3.036)	(-0.285)	(-0.755)	(-0.527)
Deal-Level FICO Score	-0.059***	-0.099***	-0.119***	-0.117***	-0.060***	-0.100***	-0.119***	-0.119***
	(-10.68)	(-9.136)	(-6.972)	(-4.094)	(-10.86)	(-9.179)	(-6.959)	(-4.066)
Deal-Level Combined Loan-to-Value Ratio	0.035	0.051	0.043	0.013	0.034	0.049	0.037	0.007
	(1.484)	(1.046)	(0.547)	(0.125)	(1.476)	(1.035)	(0.489)	(0.0695)
Deal-Level % of Loans with A.R.M.'s	-0.873**	0.030	3.135**	3.504*	-1.022**	-0.299	2.536*	2.384
	(-2.198)	(0.0409)	(2.450)	(1.776)	(-2.479)	(-0.388)	(1.855)	(1.083)
Deal-Level % Owner Occupied	-1.863	1.865	7.328	4.927	-2.493	1.018	5.154	2.483
	(-0.669)	(0.543)	(1.410)	(0.969)	(-0.863)	(0.284)	(0.958)	(0.564)
Deal-Level % Second Mortgages	-2.550***	-4.369***	-6.837***	-11.241**	-2.599***	-4.539***	-7.349***	-11.885**
	(-3.907)	(-3.447)	(-2.912)	(-2.237)	(-3.996)	(-3.625)	(-3.244)	(-2.427)
Deal-Level % Low Doc., No Doc.	2.802***	5.414***	9.619***	12.603**	2.740***	5.355***	9.629***	11.647*
	(3.718)	(3.741)	(3.107)	(2.016)	(3.629)	(3.785)	(3.312)	(1.956)
Deal-Level % Refinance	-2.536***	-5.538***	-6.505***	-3.614	-2.337***	-5.078***	-5.602***	-2.395
	(-5.474)	(-5.782)	(-3.305)	(-1.603)	(-5.037)	(-5.390)	(-3.018)	(-1.114)
Constant	47.420***	79.776***	103.006***	69.632***	49.111***	82.744***	109.902***	76.406***
	(10.49)	(10.17)	(7.860)	(3.622)	(10.66)	(10.05)	(8.132)	(3.961)
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Std. Error Clustered by Month	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1279	840	503	274	1279	840	503	274
Adjusted R-squared	0.477	0.554	0.594	0.608	0.481	0.556	0.601	0.622

Table 3. Deal-Level Default Rates Sampled *After* January 2007.

This table estimates the relationship between deal-level default rates and deal-level attributes using an OLS regression. The dependent variables, *12-mth, 24-mth, and 36-mth default rates* refer to deal-level default rates, calculated as the total number of defaults in a deal divided by the total loans outstanding ,12, 24, and 36 months after deal origination. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Collateral in California* calculates the number of loans (equal weighted) in each deal originated in the state of California. *Deal H.P.A.* measures loan-weighted, deal-level rates of house price appreciation in the year prior to deal close. Individual loans are matched with ZIP, MSA, or state-level house price indexes depending on index availability for the geographic area in which the loan was originated. *Cumulative H.P.A.* measures cumulative, deal-level rates of house price appreciation from the time of deal origination to the month in which the default rate is being measured. Deal-level measures of collateral quality are aggregated to the deal level using loan size as weights. Collateral quality measures include *FICO* scores, *combined loan-to-value ratios, % loans with adjustable-rate features, % loans owner occupied, % loans second mortgages, % low doc/no doc, % refinancing.* Our sample includes deals originated between 1998 and 2007. The sample is a panel which calculates deal-level default rates each month beginning in the month of deal inception. The sample used in this table only includes deal-level default rates measured *after* January 2007. We include month fixed effects and cluster standard errors by month. The symbols ***, ** and * indicate significance at the 1, 5, and 10% levels, respectively.

Sample: 12, 24, and 36-month deal-level default rates measured ***after*** to Jan. 2007						
	Dependent Variable: Deal-Level Default Rates					
	Months Seasoning From Deal Origination			Months Seasoning From Deal Origination		
	12-mth	24-mth	36-mth	12-mth	24-mth	36-mth
Herfindahl (Geographic Concentration)	12.732***	15.504***	11.499***			
	(4.692)	(7.234)	(3.331)			
% Loans in California				8.800***	10.845***	5.840***
				(7.009)	(9.328)	(2.764)
Deal-Level Measures of Collateral Quality:						
Deal-Level H.P.A.$_{t-1}$	0.005	0.123	0.154	-0.030	0.054	0.186
	(0.0549)	(1.200)	(1.113)	(-0.284)	(0.548)	(1.174)
Deal-Level Cumulative H.P.A.	0.006	-0.008	-0.041**	0.008	-0.006	-0.040**
	(0.858)	(-0.709)	(-2.708)	(1.041)	(-0.522)	(-2.697)
Deal-Level FICO Score	-0.111***	-0.141***	-0.135***	-0.111***	-0.139***	-0.132***
	(-10.05)	(-16.61)	(-12.06)	(-10.02)	(-16.00)	(-11.96)
Deal-Level Combined Loan-to-Value Ratio	0.218***	0.310***	0.390***	0.231***	0.324***	0.401***
	(5.557)	(6.865)	(10.96)	(6.367)	(7.052)	(10.81)
Deal-Level % of Loans with A.R.M.'s	1.042*	2.951***	4.880***	1.238**	3.238***	5.113***
	(1.899)	(4.326)	(6.136)	(2.442)	(4.107)	(5.960)
Deal-Level % Owner Occupied	-14.349***	-8.932***	3.939	-15.615***	-10.338***	3.872
	(-5.742)	(-3.313)	(1.513)	(-6.205)	(-3.704)	(1.441)
Deal-Level % Second Mortgages	-7.085***	-14.506***	-19.575***	-7.015***	-14.471***	-19.579***
	(-7.604)	(-9.059)	(-11.02)	(-8.076)	(-8.875)	(-10.80)
Deal-Level % Low Doc., No Doc.	3.730***	5.178***	5.005**	3.150***	4.453***	4.823**
	(6.925)	(4.055)	(2.480)	(4.465)	(3.202)	(2.384)
Deal-Level % Refinance	-1.797**	-1.639*	-1.586	-1.563	-1.212	-1.356
	(-2.083)	(-1.791)	(-1.279)	(-1.713)	(-1.234)	(-1.069)
Constant	68.793***	74.920***	58.050***	68.797***	73.825***	54.669***
	(7.447)	(10.80)	(6.464)	(7.294)	(10.26)	(5.882)
Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Std. Error Clustered by Month	Yes	Yes	Yes	Yes	Yes	Yes
Observations	710	1138	1151	710	1138	1151
Adjusted R-squared	0.691	0.740	0.718	0.698	0.743	0.717

Table 4. Credit Ratings and Collateral Concentration.

This table reports the results of a mutinomial logit estimate of the determinants of credit rating outcomes. The sample consists of 870 unique bonds from 870 unique deals originated between the years 2000-2007. The bonds represent the senior-most tranche in each deal, and were rated AAA at deal origination. Credit rating outcomes fall into one of three categories. "*No ratings change*" applies to bonds that have maintained at least a A-rating throughout their life. "*Ratings downgrade"* applies to bonds that have been downgraded to a BBB, BB, B, CCC, CC, or D credit rating. "*Not currently rated"* applies to bonds that do not have a current credit rating because the bond has prepaid. The multinomial logit assigns bonds in the no ratings change category as the baseline. Thus, estimated coefficients are relative to the baseline of no ratings change. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Loans in California* calculates the number of loans (equal weighted) in each deal originated in the state of California. *Deal H.P.A.* measures loan-weighted, deal-level rates of house price appreciation in the year prior to deal close. Individual loans are matched with ZIP, MSA, or state-level house price indexes depending on index availability for the geographic area in which the loan was originated. Deal-level measures of collateral quality are aggregated to the deal level using loan size as weights. Collateral quality measures include *FICO* scores, *combined loan-to-value ratios, % loans with adjustable-rate features, % loans owner occupied, % loans second mortgages, % low doc/no doc, % refinancing.* Standard errors are clustered by year. The symbols ***, ** and * indicate significance at the 1, 5, and 10% levels, respectively.

	Multinomial Logit: Credit Rating Outcomes			Multinomial Logit: Credit Rating Outcomes		
VARIABLES	Baseline = No Ratings Change (Still AAA)	Ratings Downgrade	Not Currently Rated (Prepayment)	Baseline = No Ratings Change (Still AAA)	Ratings Downgrade	Not Currently Rated (Prepayment)
	(0)	(1)	(2)	(3)	(4)	(5)
Herfindahl (Geographic Concentration)		19.669***	8.257***			
		(3.989)	(2.958)			
% Loans in California					11.766***	4.634***
					(12.22)	(3.077)
Deal-Level Measures of Collateral Quality:						
Deal-Level H.P.A.$_{t-1}$		-0.452***	-0.206**		-0.453***	-0.203**
		(-4.192)	(-2.131)		(-4.993)	(-2.384)
Deal-Level FICO Score		0.004	-0.003		0.007	-0.000
		(0.366)	(-0.483)		(0.590)	(-0.107)
Deal-Level Combined Loan-to-Value Ratio		-0.018	0.051		-0.011	0.050
		(-0.568)	(1.353)		(-0.390)	(1.398)
Deal-Level % of Loans with A.R.M.'s		1.236	0.731		1.233	0.580
		(0.722)	(0.666)		(0.607)	(0.531)
Deal-Level % Owner Occupied		-23.831***	-14.413***		-21.630***	-14.078***
		(-6.400)	(-4.739)		(-4.520)	(-4.343)
Deal-Level % Second Mortgages		3.012**	4.044***		3.457***	4.199***
		(2.313)	(3.950)		(2.669)	(4.773)
Deal-Level % Low Doc., No Doc.		3.298***	0.836		3.150***	0.811
		(3.038)	(0.775)		(2.978)	(0.690)
Deal-Level % Refinance		-1.907	1.744		-1.551	1.755
		(-0.725)	(1.216)		(-0.604)	(1.198)
Constant		23.822***	9.363		18.806***	7.565
		(3.086)	(1.319)		(2.870)	(1.256)
Year Fixed Effects		Yes	Yes		Yes	Yes
Std. Errors Clustered by Year		Yes	Yes		Yes	Yes
Observations		870	870		870	870
Psuedo R-squared		0.379	0.379		0.379	0.379

Table 5. Documenting Cross-Sectional Variation in Collateral Concentration.

This table presents summary statistics on the collateral attributes of subprime mortgage-backed securitization deals through time. Deal data is calculated by aggregating loan-level data to the deal level. Details of this process are included in the text and data appendix. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Collateral in California* calculates the number of loans (equal weighted) in each deal originated in the state of California.

Panel A. Deal Level Herfindahl

Year	N	10th %tile	25th%tile	Median	Mean	75th %tile	90th %tile
1998	3	0.050	0.050	0.060	0.093	0.169	0.169
1999	11	0.051	0.053	0.061	0.083	0.100	0.129
2000	15	0.043	0.058	0.084	0.094	0.136	0.164
2001	31	0.052	0.073	0.096	0.101	0.124	0.147
2002	68	0.063	0.082	0.132	0.137	0.171	0.237
2003	135	0.061	0.086	0.144	0.151	0.196	0.258
2004	233	0.071	0.100	0.146	0.165	0.205	0.276
2005	302	0.063	0.089	0.133	0.154	0.189	0.286
2006	321	0.064	0.092	0.128	0.139	0.174	0.221
2007	135	0.062	0.082	0.111	0.118	0.145	0.178
Full Sample	1254	0.061	0.088	0.128	0.144	0.178	0.248

Panel B. Deal Level % of Deals in California

Year	N	10th %tile	25th%tile	Median	Mean	75th %tile	90th %tile
1998	3	0.149	0.149	0.156	0.221	0.359	0.359
1999	11	0.093	0.127	0.174	0.185	0.243	0.299
2000	15	0.001	0.030	0.176	0.154	0.227	0.321
2001	31	0.057	0.149	0.234	0.215	0.296	0.327
2002	68	0.113	0.201	0.313	0.299	0.388	0.456
2003	135	0.139	0.223	0.324	0.314	0.414	0.479
2004	233	0.158	0.239	0.331	0.335	0.418	0.499
2005	302	0.142	0.216	0.296	0.311	0.393	0.499
2006	321	0.143	0.216	0.279	0.289	0.370	0.426
2007	135	0.127	0.186	0.263	0.252	0.318	0.375
Full Sample	1254	0.136	0.210	0.290	0.297	0.382	0.470

Table 6. Explaining Deal-Level Concentration: Primary Market Supply or Secondary Market Demand?

This table documents measures of geographic concentration in the mortgage origination market, total securitization market, and at the deal-level. *Panel A* reports equal-weighted Herfindahl measures of geographic concentration. *Panel B* reports equal-weighted % Loans in California measures of geographic concentration. The column titled "*Primary Market Originations"* reports primary market mortgage origination measures of concentration, by year using HMDA data. The column titled *"All Securitized Loans"* reports total securitization market measures of concentration, by year using LoanPerformance data. The columns titled *"Deal-Level Herfindahl",* and *"Deal-Level % California"* report deal-level measures of concentration. We calculate the difference between deal-level measures of concentration and origination market levels of concentration and report T-statistics associated with the difference. The final two columns tabulate the total percentage of deals in a given year that are more geographically concentrated then the origination market level of concentration and the total securitization market level of concentration, in each year.

Panel A. Equal-Weighted Herfindahl Measures of Concentration

Year	Primary Market Originations	All Securitized Loans	Deal-Level Herfindahl	Deal-Level Herfindahl minus Primary Market Originations	T-statistic	Percent of Deals with Higher Deal Concentration than Origination Concentration	Percent of Deals with Higher Deal Concentration than Total Securitized Concentration
2004	0.0571	0.0934	0.1036	0.0465	(10.68)	78.31%	42.47%
2005	0.0671	0.0894	0.0912	0.0241	(7.09)	57.46%	34.23%
2006	0.0665	0.0901	0.0816	0.0151	(7.75)	56.31%	27.34%

Panel B. % Loans in California Measures of Concentration

	Primary Market Originations	All Securitized Loans	Deal-Level % California	Deal Level % California minus Primary Market % in California	T-statistic	Percent of Deals with Higher Deal % in California than Originated % in California	Deals with Higher Deal % in California than Total Securitized % in California
2004	0.1596	0.2587	0.3167	0.1571	(19.54)	84.55%	64.14%
2005	0.1869	0.2419	0.2984	0.1115	(17.01)	77.95%	65.23%
2006	0.1778	0.2361	0.2767	0.0989	(19.84)	83.74%	63.82%

Table 7. Why is Deal Collateral Concentrated? Expected Rates of House Price Appreciation.

This table presents the results of an OLS regression of deal-level measures of collateral concentration on deal-level collateral attributes. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Collateral in California* calculates the number of loans (equal weighted) in each deal originated in the state of California. *Deal Housing Supply Elasticity* is calculated as follows. Each individual loan is assigned a housing supply elasticity measure, as calculated by Saiz (2009), for the MSA in which it resides. Loan-level elasticity measures are aggregated to the deal level using loan sizes as weights. We discuss the handling of loans with missing elasticity measures in the text. Deal-level measures of collateral quality are aggregated to the deal level using loan size as weights. Collateral quality measures include *FICO* scores, *combined loan-to-value ratios, % loans with adjustable-rate features, % loans owner occupied, % loans second mortgages, % low doc/no doc, % refinancing.* All deal attributes are measured at the time of deal creation. The sample includes 1,232 deals originated between the years 1998 and 2007. The estimation controls for year fixed effects in Columns (3) and (4). We cluster standard errors by underwriter and year. The symbols ***, ** and * indicate significance at the 1, 5, and 10% levels, respectively.

Variables	Herfindahl	% Loans in California	Herfindahl	% Loans in California
	(1)	(2)	(3)	(4)
Deal-Level Housing Supply Elasticity	-0.404***	-0.675***	-0.381***	-0.632***
	(10.27)	(12.06)	(9.92)	(11.84)
Deal-Level Measures of Collateral Quality:				
Deal-Level FICO Score	0.0004	0.001*	0.0004	0.001*
	(1.57)	(1.94)	(1.38)	(1.68)
Deal-Level Combined Loan-to-Value Ratio	-0.0002	-0.001	0.0001	-0.000
	(0.26)	(0.73)	(0.13)	(0.32)
Deal-Level % of Loans with A.R.M.'s	0.00001	0.062*	0.004	0.070**
	(0.00)	(1.84)	(0.19)	(2.18)
Deal-Level % Owner Occupied	0.219***	0.321**	0.202***	0.278*
	(5.34)	(2.17)	(4.71)	(1.94)
Deal-Level % Second Mortgages	0.026	-0.017	0.043	0.016
	(0.83)	(0.36)	(1.34)	(0.34)
Deal-Level % Low Doc., No Doc.	0.009	0.023	0.034	0.065**
	(0.33)	(0.83)	(1.20)	(2.36)
Deal-Level % Refinance	-0.071***	-0.109***	-0.076**	-0.122***
	(2.60)	(2.83)	(2.29)	(2.80)
Constant	0.302	0.672***	0.255	0.627***
	(1.50)	(2.67)	(1.20)	(2.78)
Year Fixed Effects	No	No	Yes	Yes
Std. Errors Clustered by Underwriter	Yes	Yes	Yes	Yes
Std. Errors Clustered by Year	Yes	Yes	Yes	Yes
Observations	1237	1237	1237	1237
Adjusted R-squared	0.640	0.683	0.661	0.706

Table 8. Why is Deal Collateral Concentrated? Affiliated Mortgage Originators.

This table presents the results of an OLS regression of deal-level measures of collateral concentration on deal-level collateral attributes. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Collateral in California* calculates the number of loans (equal weighted) in each deal originated in the state of California. *Deal Housing Supply Elasticity* is calculated as follows. Each individual loan is assigned a housing supply elasticity measure, as calculated by Saiz (2009), for the MSA in which it resides. Loan-level elasticity measures are aggregated to the deal level using loan sizes as weights. We discuss the handling of loans with missing elasticity measures in the text. *Above-Median Housing Elasticity Indicator* is an indicator variable equal to one for deals with an above-median level of deal-level housing supply elasticity, and zero otherwise. *Originator Affiliated with Underwriter* is an indicator variable equal to one if the deal arrangers obtained mortgage collateral from a single-name mortgage originator that is affiliated with the deal underwriter. We identify deal underwriters as being affiliated with mortgage originators using Dun & Bradstreet's "Who Owns Whom," published annually. Deal-level measures of collateral quality are aggregated to the deal level using loan size as weights. Collateral quality measures include *FICO* scores, *combined loan-to-value ratios, % loans with adjustable-rate features, % loans owner occupied, % loans second mortgages, % low doc/no doc, % refinancing.* All deal attributes are measured at the time of deal creation. The estimation controls for year fixed effects. We cluster standard errors by underwriter and year. The symbols ***, ** and * indicate significance at the 1, 5, and 10% levels, respectively.

Table 8 continued….

Variables	Herfindahl	Herfindahl	Herfindahl	% in California	% in California	% in California
	Dependent Variables					
	(1)	(2)	(3)	(4)	(5)	(6)
Originator Affiliated with Underwriter Indicator* Above-Median Housing Elasticity Indicator			-0.029*			-0.053*
			(1.76)			(1.66)
Originator Affiliated with Underwriter Indicator	0.018*	0.019	0.034*	0.014	0.016	0.044**
	(1.68)	(1.34)	(1.71)	(0.61)	(0.53)	(2.34)
Above-Median Housing Elasticity Indicator		-0.073***	-0.069***		-0.119***	-0.112***
		(8.11)	(6.92)		(8.15)	(6.00)
Deal-Level Housing Supply Elasticity	-0.398***			-0.645***		
	(10.27)			(10.92)		
Deal-Level Measures of Collateral Quality:						
Deal-Level FICO Score	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	(3.22)	(3.06)	(3.02)	(5.27)	(5.15)	(4.84)
Deal-Level Combined Loan-to-Value Ratio	-0.001**	-0.002***	-0.002***	-0.002**	-0.003**	-0.003**
	(2.33)	(3.15)	(3.17)	(2.01)	(2.45)	(2.48)
Deal-Level % of Loans with A.R.M.'s	0.038**	0.062***	0.062***	0.104***	0.143***	0.142***
	(2.37)	(6.84)	(6.60)	(3.61)	(8.84)	(9.15)
Deal-Level % Owner Occupied	0.233***	0.245***	0.258***	0.239**	0.260***	0.283***
	(6.85)	(6.85)	(7.26)	(2.43)	(4.69)	(10.19)
Deal-Level % Second Mortgages	0.073*	0.060	0.061	0.066	0.045	0.046
	(1.73)	(1.23)	(1.26)	(1.05)	(0.66)	(0.71)
Deal-Level % Low Doc., No Doc.	0.013	0.103***	0.106***	0.030	0.175***	0.180***
	(0.47)	(3.90)	(4.10)	(1.10)	(6.55)	(7.02)
Deal-Level % Refinance	-0.073**	-0.088***	-0.091***	-0.100**	-0.124**	-0.131**
	(2.32)	(2.92)	(2.81)	(1.98)	(2.58)	(2.51)
Constant	0.198	-0.302	-0.300	0.451*	-0.359*	-0.356**
	(1.04)	(1.40)	(1.42)	(1.89)	(1.93)	(2.05)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Std. Errors Clustered by Underwriter	Yes	Yes	Yes	Yes	Yes	Yes
Std. Errors Clustered by Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	915	915	915	915	915	915
Adjusted R-squared	0.685	0.564	0.567	0.709	0.585	0.588

Table 9. Pricing of Correlation Risk: Spread over LIBOR on Bonds Sold at Par.

This table explains variations in the spread over LIBOR being paid to bondholders on a sample of subprime bonds purchased at par. Columns (1) and (3) estimate spreads on bonds purchased at par within 12 months of the deal origination. Columns (2) and (4) estimate spreads on bonds purchased at par at the time of deal origination. The variable *Herfindahl* is a Herfindahl index, calculated as the sum of squared weights, where weights are calculated as the percent of each deal's principal originated in a given state. *% Collateral in California* calculates the number of loans (equal weighted) in each deal originated in the state of California. *Original Subordination* calculates the percentage of deal principal that is junior to the bond included in the sample. *Over Collateralization* is a measure of deal-level credit support. *Weighted-average Maturity* is an estimate of the expected maturity of the bond as calculated by Bloomberg. *% with Prepayment Penalty* calculates the percentage of loans in the deal collateral with prepayment penalties, a variable which should influence bond maturity. *Deal H.P.A.* measures loan-weighted, deal-level rates of house price appreciation in the year prior to deal close. Individual loans are matched with ZIP, MSA, or state-level house price indexes depending on index availability for the geographic area in which the loan was originated. Deal-level measures of collateral quality are aggregated to the deal level using loan size as weights. Collateral quality measures include *FICO* scores, *combined loan-to-value ratios, % loans with adjustable-rate features, % loans owner occupied, % loans second mortgages, % low doc/no doc, % refinancing.* All deal attributes are measured at the time of deal creation. We include month and credit-rating fixed-effects. Standard errors are clustered by month of transaction. The symbols ***, ** and * indicate significance at the 1, 5, and 10% levels, respectively.

Table 9 continued…

	Dependent Variable: Bond-Specific Spread Over Libor at Origination			
	(1)	(2)	(3)	(4)
Herfindahl (Geographic Concentration)	0.188	0.073		
	(1.023)	(0.914)		
% Loans in California			0.096	0.033
			(0.875)	(0.637)
Original Subordination	-0.050	-0.040**	-0.050	-0.039**
	(-1.275)	(-2.493)	(-1.273)	(-2.463)
Over-Collateralization	-0.009	-0.000	-0.008	-0.000
	(-1.568)	(-0.031)	(-1.521)	(-0.034)
Weighted-Average Maturity	0.002**	-0.000	0.002**	-0.000
	(2.571)	(-0.001)	(2.551)	(-0.030)
% with Prepayment Penalty	-0.132	-0.086	-0.135	-0.087
	(-1.312)	(-1.582)	(-1.320)	(-1.573)
Deal-Level Measures of Collateral Quality:				
Deal-Level H.P.A.$_{t-1}$	-0.001	-0.002	-0.000	-0.002
	(-0.184)	(-0.575)	(-0.044)	(-0.517)
Deal-Level FICO Score	-0.001***	-0.001***	-0.001***	-0.001***
	(-4.681)	(-4.439)	(-4.376)	(-4.296)
Deal-Level Combined Loan-to-Value Ratio	0.003	0.002	0.003	0.002
	(0.896)	(0.995)	(0.921)	(0.990)
Deal-Level % of Loans with A.R.M.'s	0.164**	0.084	0.163**	0.081
	(2.174)	(0.859)	(2.134)	(0.829)
Deal-Level % Owner Occupied	0.350	0.030	0.361	0.037
	(1.213)	(0.154)	(1.260)	(0.192)
Deal-Level % Second Mortgages	-0.043	-0.161**	-0.042	-0.160**
	(-0.449)	(-2.102)	(-0.444)	(-2.106)
Deal-Level % Low Doc., No Doc.	0.022	-0.044	0.022	-0.045
	(0.323)	(-0.876)	(0.328)	(-0.903)
Deal-Level % Refinance	0.037	-0.058	0.043	-0.055
	(0.365)	(-0.852)	(0.426)	(-0.807)
Constant	-0.172	0.719***	-0.187	0.714***
	(-0.435)	(3.583)	(-0.470)	(3.596)
Month Fixed Effects	Yes	Yes	Yes	Yes
Credit Rating Fixed Effects	Yes	Yes	Yes	Yes
Purchased at Par Within 12 Months of Deal Origination	Yes	No	Yes	No
Purchased at Par at Time of Deal Origination	No	Yes	No	Yes
Standard Errors Clustered by Month of Transaction	Yes	Yes	Yes	Yes
Observations	1234	745	1234	745
Adjusted R-squared	0.784	0.864	0.784	0.864

Table 10. Pricing of Correlation Risk: Conditional Portfolio Sorts.

These tables examine coupon spreads based on portfolio sorts in years 2005 and 2006, separately. One-way FICO sorts evaluate differences in coupon spreads between the highest quartile deal-level FICO scores and lowest quartile deal-level FICO scores. Two-way sorts first separate deals into quartiles by FICO score, then into quartiles by the measure of geographic concentration within each FICO quartile. We test whether differences between high and low FICO and concentration buckets are significant, and report associated t-statistics. Deal-level geographic concentration is measured using the *Herfindahl* index discussed in the text and summary statistics in Table 1.

Panel A

One-Way Sort: All AAA-rated Bonds Issued in 2005

Coupon Spread on High FICO Portfolio (>75th%tile)			Coupon Spread on Low FICO Portfolio (< 25th%tile)			Low-High	T-Stat.
N	Spread over LIBOR	Std. Dev.	N	Spread over LIBOR	Std. Dev.		
44	0.211	0.071	44	0.272	0.089	0.062	(3.57)

Panel B

Two-Way Conditional Sort: All AAA-rated Bonds Issued in 2005

		Herphindahl (Geographic Concentration)					
		< 25%tile	25th-50th%tile	50th-75th%tile	> 75th%tile	High Con. - Low Con.	T-Stat.
F	< 25%tile	0.310	0.282	0.308	0.206	-0.104	(-3.18)
I	25th-50th%tile	0.262	0.253	0.299	0.267	0.006	(0.13)
C	50th-75th%tile	0.219	0.181	0.242	0.254	0.035	(0.69)
O	> 75th%tile	0.237	0.230	0.223	0.170	-0.067	(-2.02)
	Low Fico - High FICO	0.073	0.053	0.085	0.036		
	T-Stat.	(1.76)	(1.38)	(2.39)	(1.67)		

Table 10 Continued...

Panel C

One-Way Sort: All AAA-rated Bonds Issued in 2006

Coupon Spread on High FICO Portfolio (>75th%tile)			Coupon Spread on Low FICO Portfolio (< 25th%tile)			Low-High	T-Stat.
N	Spread over LIBOR	Std. Dev.	N	Spread over LIBOR	Std. Dev.		
61	0.181	0.102	67	0.174	0.090	-0.025	(-1.46)

Panel D

Two-Way Conditional Sort: All AAA-rated Bonds Issued in 2006

		Herphindahl (Geographic Concentration)					
		< 25%tile	25th-50th%tile	50th-75th%tile	> 75th%tile	High Con. - Low Con.	T-Stat.
F	< 25%tile	0.161	0.139	0.107	0.213	0.052	(1.89)
I	25th-50th%tile	0.146	0.208	0.183	0.180	0.035	(1.31)
C	50th-75th%tile	0.180	0.181	0.113	0.041	-0.139	(-1.09)
O	> 75th%tile	0.198	0.152	0.151	0.221	0.023	(0.57)
	Low Fico - High FICO	-0.037	-0.014	-0.044	-0.008		
	T-Stat.	(-1.11)	(-0.49)	(-1.24)	(-0.23)		

Table 11. Pricing of Correlation Risk: Conditional Portfolio Sorts.

These tables examine coupon spreads based on portfolio sorts in years 2005 and 2006, separately. One-way concentration sorts evaluate differences in coupon spreads between the highest quartile concentrated-deals and lowest quartile concentrated-deals. Two-way sorts first separate deals into quartiles by geographic concentration, then into quartiles by average deal-level FICO scores within each concentration quartile. We test whether differences between high and low concentration deals and FICO buckets are significant, and report associated t-statistics. Deal-level geographic concentration is measured using the *Herfindahl* index discussed in the text and summary statistics in Table 1.

Panel A

One-Way Sort: All AAA-rated Bonds Issued in 2005

Coupon Spread on High Con. Portfolio (>75th%tile)			Coupon Spread on Low Con. Portfolio (< 25th%tile)			High - Low	T-Stat.
N	Spread over LIBOR	Std. Dev.	N	Spread over LIBOR	Std. Dev.		
45	0.234	0.082	46	0.255	0.106	-0.020	(-1.03)

Panel B

Two-Way Conditional Sort: All AAA-rated Bonds Issued in 2005

	Loan-Weighted Deal-Level FICO					
Concentration	< 25%tile	25th-50th%tile	50th-75th%tile	> 75th%tile	Low Fico - High Fico	T-Stat.
< 25%tile	0.310	0.249	0.233	0.222	0.088	(2.25)
25th-50th%tile	0.245	0.320	0.289	0.157	0.088	(1.62)
50th-75th%tile	0.210	0.288	0.290	0.276	-0.066	(-1.85)
> 75th%tile	0.273	0.220	0.203	0.239	0.034	(0.84)
High Con. - Low Con.	-0.038	-0.029	-0.031	0.017		
T-Stat.	(-0.94)	(-0.74)	(-0.90)	(0.44)		

Table 11 Continued….

Panel C

One-Way Sort: All AAA-rated Bonds Issued in 2006

Coupon Spread on High Con. Portfolio (>75th%tile)			Coupon Spread on Low Con. Portfolio (< 25th%tile)			High - Low	T-Stat.
N	Spread over LIBOR	Std. Dev.	N	Spread over LIBOR	Std. Dev.		
65	0.183	0.089	67	0.174	0.090	0.009	(0.58)

Panel D

Two-Way Conditional Sort: All AAA-rated Bonds Issued in 2006

	Loan-Weighted Deal-Level FICO					
Concentration	< 25%tile	25th-50th%tile	50th-75th%tile	> 75th%tile	Low Fico - High Fico	T-Stat.
< 25%tile	0.161	0.154	0.157	0.215	-0.054	(-1.71)
25th-50th%tile	0.130	0.122	0.195	0.171	-0.040	(-1.76)
50th-75th%tile	0.193	0.192	0.123	0.064	0.129	(1.27)
> 75th%tile	0.182	0.211	0.168	0.172	0.009	(0.31)
High Con. - Low Con.	0.021	0.057	0.012	-0.042		
T-Stat.	(0.75)	(1.93)	(0.37)	(-1.27)		