Information Frictions in Securitization Markets: Unsophisticated Investors or Opaque Assets?

David Echeverry *

July 5, 2018

Abstract

Bond prices are informative about subsequent downgrades, but the level of informativeness is affected by two frictions. I use a measure of deal opacity in collateralized mortgage obligation markets to show that the predominant friction is (in)completeness of documentation on the underlying loans. The agency friction between sophisticated junior investors and unsophisticated senior ones is secondary. In particular, prices of junior tranches appear no more informative than those of AAA tranches among "low-doc" deals, and the latter no less informative for "full-doc" deals. The results suggest that documentation transparency can be an effective complement to skin in the game requirements.

JEL classification: G21, G24

Keywords: credit rating agencies, mortgages, mortgage-backed securities, Gaussian copula, credit risk.

*Email:david.echeverry@berkeley.edu. I am grateful to Martijn Cremers, Nicolae Gârleanu, Amir Kermani, Ilénin Kondo, Haoyang Liu, Christopher Palmer, James Vickery and Nancy Wallace for their valuable feedback. I also thank seminar participants at Universidad de los Andes, Universidad Javeriana and University of Notre Dame for their helpful comments and suggestions. I thank Paulo Issler, Sanket Korgaonkar and Haoyang Liu for their assistance with source data.

Information frictions in private label mortgage-backed securities (MBS) are cited among the root causes of the financial crisis of 2007. In spite of the collapse of private label securitization and subsequent stagnation,[1] private investment remains vital to the mortgage market. Finkelstein, Strzodka, and Vickery (2018) document that Fannie Mae and Freddie Mac (the government-sponsored entities, or GSEs) have reduced Federal government exposure to credit risk on close to $1.8tn mortgages by transfering a (growing) share of it to the private sector. But in spite of their importance in securitization markets, much remains to be understood about the information frictions affecting private investors in securitization markets. This paper aims to bridge this gap.

I say that bond prices are informative to the extent that they predict subsequent bond downgrades, controlling for agency rating at the time of transaction. Early prices of collateralized mortgage obligations[2] (CMOs) are informative, as argued by Ashcraft, Goldsmith-Pinkham, Hull, and Vickery (2011). But there are information frictions which affect the level of informativeness. According to Ashcraft and Schuermann (2008), two frictions take place between the investor and the originator of the securities. The first one is investor (un)sophistication, whereby junior investors are better informed than senior ones (Boot and Thakor, 1993). According to Gorton and Pennacchi (1990), the latter seek information-insensitive tranches (i.e. the AAA-rated ones) while the former are better suited to handle the information-sensitive ones (i.e. the junior tranches).[3] This gives rise to a principal-agent problem. Adelino (2009) argues that early AAA prices are not informative, unlike junior tranche prices, giving evidence of this information differential between the -unsophisticated-senior investor and the -sophisticated- junior investor.

A less-explored friction relates to lack of due diligence on the quality of the underlying assets, making these assets opaque. The incompleteness of information gives rise to adverse selection (Skreta and Veldkamp, 2009). The main contribution of this paper is to provide evidence that, of the two frictions distinguished by Ashcraft and Schuermann (2008), asset opacity takes precedence over investor sophistication. I measure deal opacity as an index of documentation completeness, linking the loans underlying each deal. I find that prices from junior tranches are no more informative than those of AAA tranches among deals backed by low documentation loans, where the value of the asset is opaque. Conversely, AAA prices are no less informative than junior ones within "full-doc deals, which are not opaque. Among deals with intermediate levels of documentation, I find evidence in line with Adelino (2009), namely that junior bond prices appear more informative than AAA ones. The evidence agrees this principal-agent problem indeed takes place, but it is mediated by the extent of asset opacity affecting the deal.

The results suggest that regulation interventions focusing on the agency problem, such as risk retention in the form of skin in the game, can be complemented by market transparency initiatives -achieving better documentation on the underlying loans-. To the extent that incomplete information is easier to tackle than differential sophistication, such transparency initiatives can

[1] The market for non-conforming loans is far from its 2007 peak origination of over $1tn despite the recent upwards trend (in 2017, $4.1bn of securities backed by so-called nonprime loans were issued -according to Inside Mortgage Finance- with issuance in Q1 2018 being roughly twice that of Q1 2017).

[2] A collateralized mortgage obligation is a tranche of a mortgage-backed security, usually backed by private label loans. As discussed by Finkelstein et al. (2018), the use of a tranching structure as a credit risk transfer mechanism has also been introduced to agency securities since 2013.

[3] The efficiency of this arrangement is discussed by Dang, Gorton, and Holmström (2013). In particular, when information is costly this helps the market liquidity (Gorton and Ordonez, 2013).

be an effective instrument to help price informativeness in private label securitization markets. In particular, risk retention requirements could be tied to the extent to which documentation is complete on the underlying loans.

As explained in IOSCO (2008) the key step in the rating process of a structured product is to determine the amount of subordination that will ensure a given rating, in particular a Standard & Poor's AAA. This makes the subordination structure an essential aspect of the bondholder's risk assessment. In fact, between price, coupon and subordination, the latter is the most sensitive to asset opacity. Whereas the level of informativeness of bond price does not vary much as a function of documentation completeness (a fall in price is uniformly predictive of a downgrade, controlling for rating), subordination is predictive of downgrades among "high-doc" deals but not among "low-doc" ones. Moreover, the amount of AAA issuance is decreasing in documentation completeness (controlling for deal average probability of default). Prior evidence of rating inflation in CDOs includes Griffin and Tang (2012) who speak of subjective ratings. An, Deng, Nichols, and Sanders (2015) use the number of tranches as a proxy for complexity and argue that more complex CMBS structures see lower subordination levels. Then in RMBS markets, Benmelech and Dlugosz (2010) link rating inflation to rating shopping, but not to asset opacity. My result is in line with the theoretical predictions of Skreta and Veldkamp (2009) that ratings are more likely to be inflated when asset quality is opaque (or "complex" to use their term).

I construct a summary measure that takes into account price, coupon, probability of default, probability of prepayment and subordination by computing the implied default correlation of a given tranche.[4] Because a central premise of securitization is diversification through pooling, default correlation is essential to the value of the security. Thus prices of structured products that are subject to default risk reflect investors' beliefs about default correlation. The higher it is, the more volatility there will be in the portfolio cashflows, which benefits junior bondholders at the expense of senior ones (Duffie and Gârleanu, 2001). I use a single factor Gaussian copula (Li, 2000), which Hull and White (2006) call "the standard market model for valuing collateralized debt obligations and similar instruments".[5] I estimate the probability of default (PD) and loss given default (LGD) from loan performance data (following common practice in CDO pricing models that PD and LGD on the underlying asset are given by the underlying loans) and default correlations are implied from the market price.

Heterogeneity in the information content of implied correlations means investors disagree about the value of this parameter. By taking default probabilities as fixed and estimating default correlations, the implicit assumption in the Gaussian copula approach is that the main source of disagreement among investors in a given deal is the default correlation. The literature has examined the role of

[4]Default correlations can be computed ex post from default experience instead of inferring them from bond prices. See Cowan and Cowan (2004); de Servigny and Renault (2002); Geidosh (2014); Gordy (2000); Nagpal and Bahar (2001). Though default-based measures are not directly comparable to ours (Frye, 2008), one study based on default experience worth noting here is Griffin and Nickerson (2016). They infer rating agency beliefs about corporate default correlations by studying collateralized loan obligation (CLO). Their results suggest such beliefs were revised upwards after the crisis, but not sufficiently so when benchmarked against a performance-based estimator accounting for unobserved frailty in the default generating process (Duffie, Eckner, Horel, and Saita, 2009). The results suggest that agency ratings adapted more slowly to the crisis than market prices.

[5]See Brunne (2006); D'Amato and Gyntelberg (2005); Duffie and Singleton (2012); Elizalde (2005); Hull and White (2004, 2006, 2008); McGinty, Beinstein, Ahluwalia, and Watts (2004); Tzani and Polychronakos (2008).

disagreement about other risk attributes such as the probability of a crisis (Simsek, 2013) or the prepayment speed (Carlin, Longstaff, and Matoba, 2014; Diep, Eisfeldt, and Richardson, 2016). The prominence of Gaussian copulas in the CDO literature suggests that the primary source of disagreement across bonds in such a structure is the default correlation.

Coval, Jurek, and Stafford (2009a) use a Gaussian copula model[6] to show that security prices are sensitive to underlying default correlations, and that this sensitivity compounds along the structured finance chain. As (Cordell, Huang, and Williams, 2012) show (see Figure N.1) the underlying collateral of cash CDOs is predominantly mezzanine tranches of CMOs, which in turn are composed of mortgages. This means that CDOs are very sensitive to loan default correlation, much like the CDO^2 in Coval et al. (2009a).[7] This highlights the importance of CMO default correlations to structured finance markets, while leaving the question open as to which of the investors, the junior or the senior ones, may be miscalculating them. The evidence we provide suggests senior investors were less informed, but only for intermediate levels of opacity.

This paper argues that the main information friction is deal opacity stemming from low documentation loans. I use loan documentation completeness indicators to construct an deal level index of opacity. A number of papers have used a similar measure to study the effects of opacity in mortgage markets. JEC (2007) documents a relative decline in the number of full documentation subprime loans in the running to the crisis. Keys, Mukherjee, Seru, and Vig (2010) argue that the "low-doc" loans underperformed (in terms of defaults) relative to otherwise similar but better documented loans. This underperformance of low-doc loans is confirmed by the results of Kau, Keenan, Lyubimov, and Slawson (2011). Moreover,Ashcraft, Goldsmith-Pinkham, and Vickery (2010) use a loan-level measure of documentation completeness (similar to the one I use) to document the underperformance of "low-doc" deals. The emphasis in this paper is on investor information in bond markets rather than collateral performance. Finally, Adelino, Gerardi, and Hartman-Glaser (2016) argue that investors in the secondary market for loans deal with opacity by *skimming* the underlying loans; they look at the time to sale of loans in the secondary market, while I consider the channel of bond prices.

The focus on early originations leaves out developments that took place over the boom. A recent literature suggests that deal opacity had an increasingly important role in the running to the crisis. Using six measures of deal complexity built from the prospectuses of subprime securities issued between 2002 and 2007, Ghent, Torous, and Valkanov (2016) offer evidence of growing obfuscation between the issuer and the senior investor, so that the latter didn't price in the higher risks due to security complexity. They argue that complex deals facilitated the collusion between the issuer and the junior investor (Demiroglu and James, 2012) to divert cash flows from senior securities to junior ones. My finding that pre-boom originations across the spectrum (prime, alt-A and

[6] Using their parameters I replicate their results (see Figure 3.1).

[7] Gorton (2009) argues that the information destruction in structured products was caused by their layered structure. Because of this, CMO prices are the closest reflection of the market view on default correlations. We provide a measure of default correlations directly from RMBS prices, which contributes to prior estimates from the CDO pricing literature. Among those, Duffie and Gârleanu (2001) and Duffie and Singleton (2012) discuss the pricing of cash CDOs. Otherwise, the literature has mostly focused on synthetic CDOs and tranches of credit default swap baskets (**?**Andreoli, Ballestra, and Pacelli, 2016; **?**; Brunne, 2006; Buzková and Teplý, 2012; Coval, Jurek, and Stafford, 2009b; Elizalde, 2005; D'Amato and Gyntelberg, 2005; Hull and White, 2004, 2006; **?**; **?**; Stanton and Wallace, 2011).

subprime) see both senior and junior investors are equally affected by loan opacity suggests that collusion became a problem over the boom.

The paper proceeds as follows. Section 1 presents the data. Section 2 presents my empirical strategy based on price, coupon and subordination. Section 3 lays out the copula model from which I infer default correlations and copula model estimates. Section 4 replicates the results from Section 2 using implied correlation as the independent variable. Section 5 concludes.

1 Data

CMOs are traded over the counter, but proprietary datasets collect transaction information. I use data from Thomson Reuters, which records bond prices from January 2004 onwards. I obtain series of prices for CMOs originated before and up to June 2005, i.e. prior to the pre-crisis mortgage boom. ABSNet collects monthly information about private label securitization deals, providing snapshots of all tranches inside a given deal between the time of origination and the end of 2016. Starting July 2009, ABSNet started recording bond prices over time, which allows me to cross-check prices across sources by matching on bond CUSIP, year and month (keeping the nearest transaction to the rating observation date[8]). I check the consistency between the ABSNet price and the mid price in Thomson Reuters. I find a median absolute difference of \$0.06 and a 99th percentile of \$1.51, the difference being consistent with small time differences in the date of observation across sources.

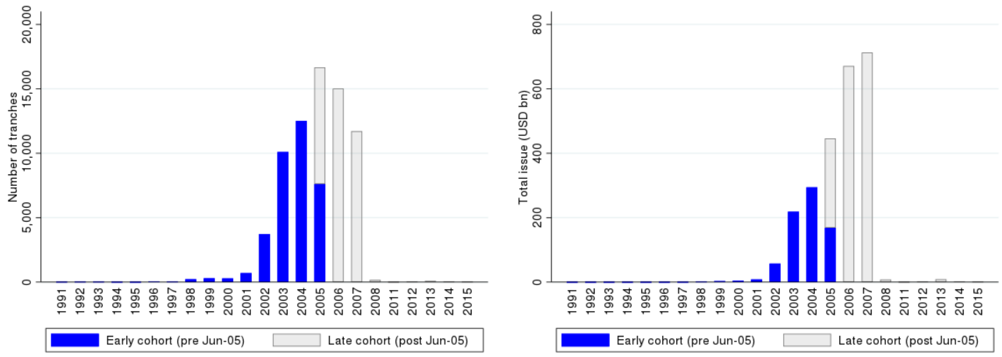
For each month, ABSNet records rating, subordination, bond maturity and coupon for each tranche. I collect all the snapshots available from each deal in their website. Tranches are organized in a matrix format by increasing subordination levels, which determine the default cushion for a given tranche. From there I derive the detachment point for each tranche, and thus the waterfall of losses for the given deal.[9]

From the early cohorts (i.e. those originated before June 2005) I observe 35,692 tranches (about 14 tranches per deal on average) for a total \$1,854.8bn of originated securities. See Table 48. In comparison, Adelino (2009) includes boom-time data to obtain 67,412 securities from JP Morgan's MBS database, a total issue of \$4,204.8bn. I follow his data cleaning procedures such as removing Interest Only, Principal Only, Inverse Floater and Fixed to Variable bonds from the sample. Alt-A and subprime deals are the largest classes (see Table 46). Though the size of these asset classes mostly built up in the running to the crisis (Gorton, 2009), my estimation sample is also composed mostly of suprime and Alt-A bonds.

Most of the bond issuance in my sample are rated AAA at origination (see Figure N.2). As Figure N.3 shows, the bonds were mostly priced at par, or even slight premium, at the moment of origination, which we observe for the tranches originated in 2004 and 2005. This applies in particular to BBB bonds, which Deng, Gabriel, and Sanders (2011) link to demand pressures from

[8]The average distance in days is 1.83, the median is 0 and the 99th percentile 53 days

[9]Some deals have more than one structure inside, each structure giving rise to its own subordination waterfall. I source each structure separately, and treat different structures as if they were different deals.



(a) Number of tranches

(b) Total issue

Figure 1.1: Number of tranches and amount issued by vintage year for private label collateralized mortgage obligations. Source: ABSNet bond data. The counts in our estimation sample (early vintages, prior to June 2005) are recorded in blue, while the numbers for late vintage tranches are illustrated in light grey.

the surge of CDO markets. Within two months of issue, prices drop. Bonds then remain priced at a discount over subsequent trades. As Figure 1.2 shows, discounts are higher in the running to the crisis for AAA bonds, and within AAA they are higher for prime and Alt-A bonds. Over 2007 prices fall, but BBB bonds see a sharp fall compared to the relatively mild fluctuation in AAA prices. In comparison, AAA and BBB bond coupons have a similar pattern over time as shown by Figure 1.3. Aside from the wider fluctuations for BBB subprime and second lien bonds compared to the corresponding AAA ones, the difference over time across seniorities is less over prices than over coupons.

ABSNet provides the Standard & Poor's (S&P) rating, which is the main ordinal variable we use to capture the cash flow sequence among the bonds in a given deal. When the security has no S&P rating we use the one issued by Fitch, which uses the same grading scale. Figure 1.4 shows the average subordination percentage by rating at origination. Tranching becomes steeper as the rating increases, and Second Lien/Subprime deals in general require more subordination at each rating grade. The average tranching structure lines up in general with the one Cordell et al. (2012) obtain from Intex data (see Table 47 for a comparison), apart from relatively thicker AAA tranches in our sample.[10]

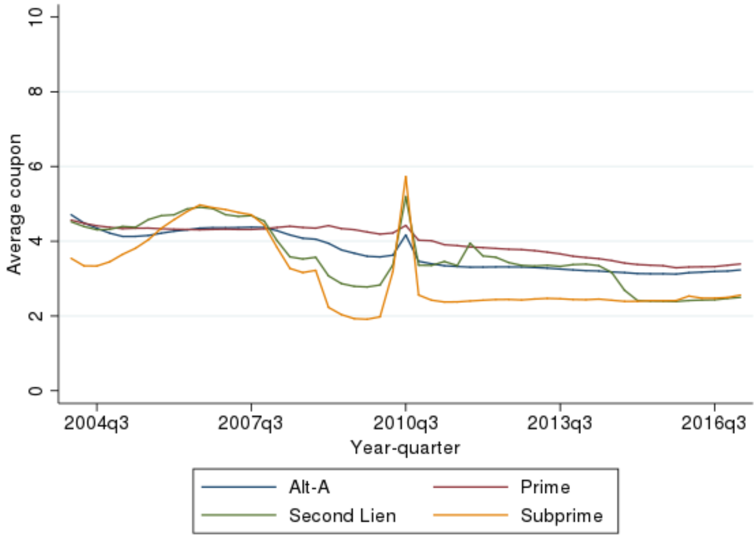
Changes in subordination percentage take place over the cycle, though mostly for subprime deals, reflecting the effect of defaults and prepayments. This is shown in Figure N.4, which depicts the point-in-time difference in average subordination between AAA and BBB tranches. While the difference remains close to constant for Alt-A and prime deals, the difference rises for subprime deals in the running to the crisis, with a slight downward trend over time afterwards. In summary, among the tranche-level variables we use for the pricing model, i.e. price, coupon and subordination structure, the first two show exhibit more cyclical variation than the latter.

[10]Rule 144A of the Securities Act of 1933 allows private companies to sell unregistered securities to qualified institutional buyers. Intex contains data on 144A deals, which are not in our sample.

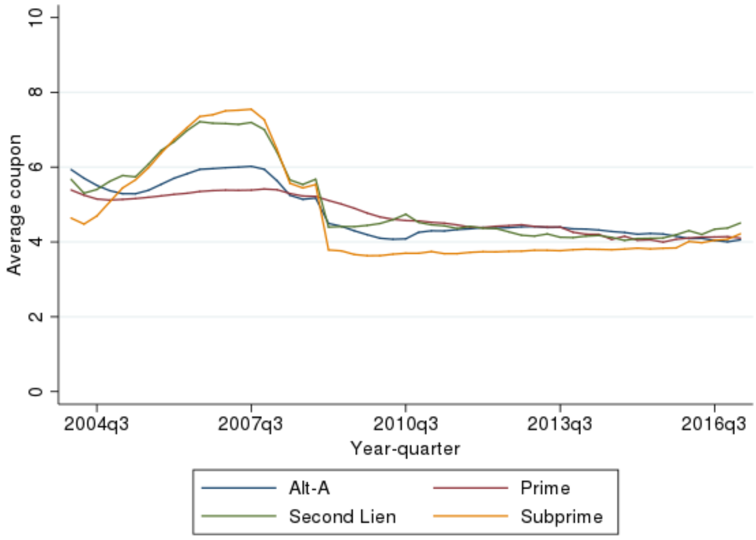
(a) Tranches rated AAA at origination

(b) Tranches rated BBB at origination

Figure 1.2: Average price by initial rating. Source: Thomson Reuters. For all the prices observed within a given month we use the closest to month end. The figure presents average price over trading time (for early vintages, prior to June 2005) controlling for initial rating.



(a) Tranches rated AAA at origination



(b) Tranches rated BBB at origination

Figure 1.3: Average coupon by initial rating. Source: ABSNet bond data.The figure presents average coupon rate over trading time (for early vintages, prior to June 2005) controlling for initial rating.

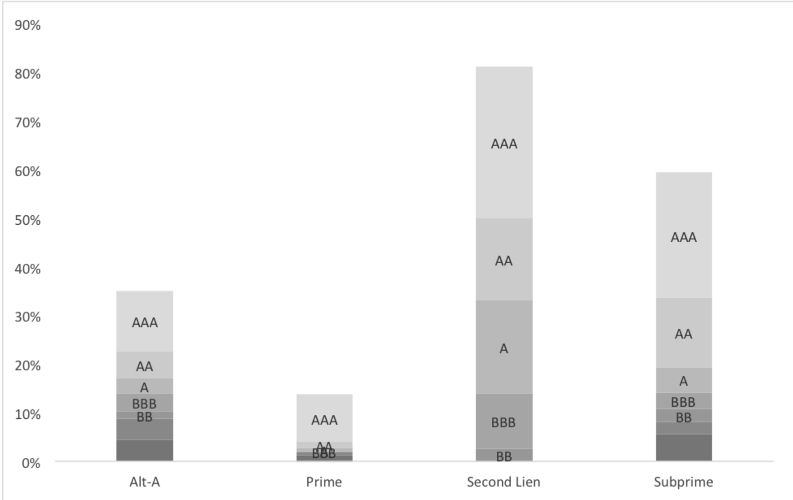
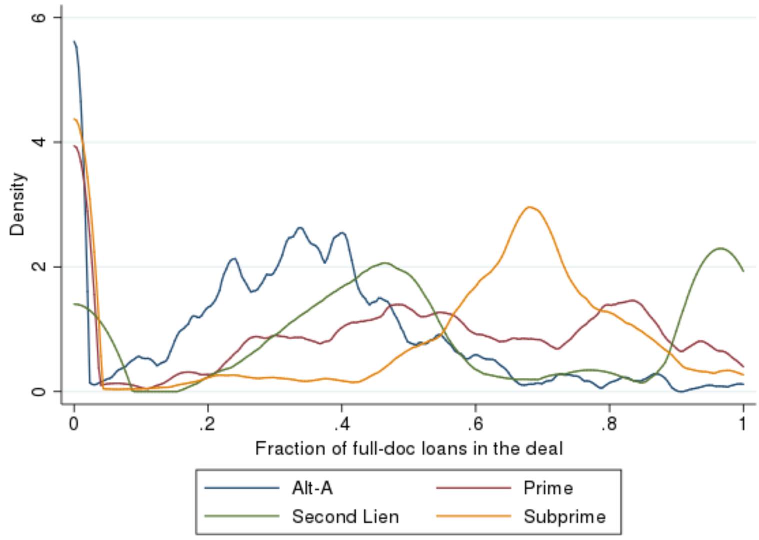


Figure 1.4: Deal structure. Source: ABSNet bond data. For our sample of early vintage deals, we look at the difference in subordination between tranches with consecutive S&P ratings. We then average the outcome by rating and asset type, aggregating at coarse grade level (see mapping in Table 38). This average difference is represented here, stacked by asset type.

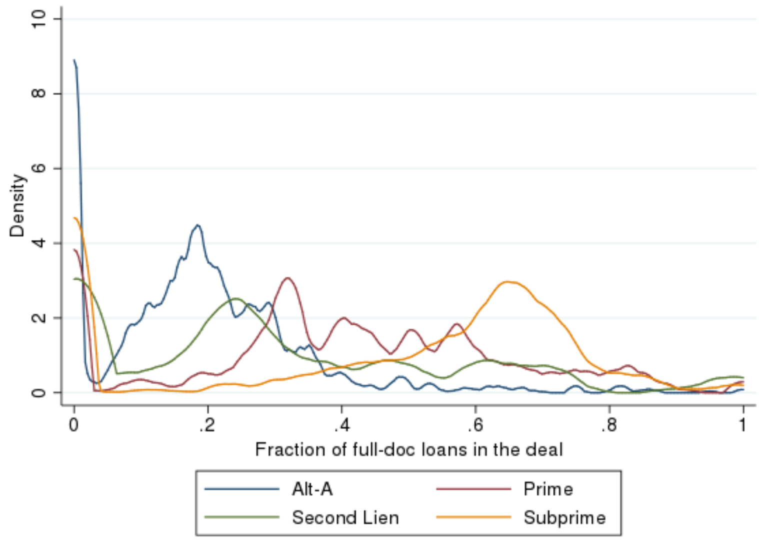
I use origination and monthly performance data on the underlying loans by ABSNet. Loans are linked to their respective deals. We start with a sample of 6,453,799 loans of which 2,944,014 are originated by 2005. We have loan and borrower characteristics such as FICO score, owner occupancy, original loan amount and original LTV, which we will use in Section 3.1 to estimate default and prepayment hazard models.

The loan data also provides a documentation completeness indicator for each loan. This is categorized as full, limited, alternative or no documentation. Loans with full documentation provide verification of income as well as assets. Loans with limited documentation provide no information about borrower income but do provide some about their assets. "No-documentation" loans provide no information about income or assets. Figure 1.5 shows a distribution of the share (at the deal level) of loans with full documentation in our sample of vintages prior to June 2005. It suggests subprime loans were relatively better documented than Alt-A deals, with densities peaking around 0.7 and 0.35 approximately. Prime deals show a higher dispersion in terms of documentation completeness. In comparison, density plots on post-June 2005 issues suggest that documentation completeness deteriorated more among Alt-A, second lien and prime deals relative to subprime ones in the running to the crisis.

Including cases of partial and alternative documentation, I assign a documentation score to each loan (no documentation=0; partial=0.1; alternative=0.3; full=1). In comparison Keys et al. (2010) use percentage of completeness, which is equivalent to excluding the intermediate values from my score. I average documentation scores into a deal level opacity score. Figure 1.6 presents the averages by asset type and vintage year. Note that Alt-A markets can only be characterized by low documentation levels -relative to other types- from year 2000 onwards. The downward slope in Figure 1.6 reflects the decline in lending standards in the running to the crisis observed on subprime loans by Dell'Ariccia, Igan, and Laeven (2012) and Keys et al. (2010).



(a) Originated before June 2005



(b) Originated after June 2005

Figure 1.5: Kernel density plot of the distribution of full-documentation loans by deal asset type. For each deal we obtain the percentage of fully documented loans associated to it. The figure represents a kernel density plot of the distribution of deals along this measure. A separate plot on vintages later than June 2005 is provided for comparison.

Figure 1.6: Average documentation index by vintage year. Source: ABSNet loan data. We assign a documentation score to each loan (no documentation=0; partial=0.1; alternative=0.3; full=1). Then for a given deal we compute the average documentation index, and present the averages by asset type and vintage year.

Other data include dynamic covariates such as CBSA level home price indices from FHFA and interest rate data; I use the difference between the loan original interest rate from ABSNet and the original ten year Treasury rate from FRED. Using Treasury rates I also compute the coupon gap (the difference between the ten year rate at origination and the current ten year rate). From Bloomberg I extract bond contractual maturities and the weighted average life.

2 Empirical strategy

I assess the information content of the different bond attributes considered so far (price, coupon and subordination) by estimating regressions of the form

$$downgrade_{i,2009} = f(\alpha + \beta X_{i0} + \eta_{rating_{i0}} + \varepsilon_i) \tag{1}$$

where X_{i0} is a vector of bond attributes at origination such as price, subordination and coupon, controlling for deal vintage and tranche rating at origination.

Table 24 presents regression results for specification (1). A higher bond price is predictive of a lower probability of downgrade, and a higher percentage subordination has the same effect. Both are significant predictors of downgrades. A higher coupon significantly predicts lower downgrades, though this only holds for below-AAA bonds. Now we split the sample by value of the opacity index derived in Section 1, using four buckets of size 0.25. Table 25 shows that the effect most clearly driven by documentation quality is that of subordination percentage: the corresponding regression coefficient decreases monotonically from insignificant, for the lowest documentation indices, to negative and significant for the highest ones.

	downgrade		
	(1)	(2)	(3)
	All	AAA only	Non-AAA only
Price	-0.0187***	-0.0457***	-0.00932***
	(0.00151)	(0.00299)	(0.00149)
Coupon	-0.123***	-0.0365	-0.184***
	(0.0178)	(0.0245)	(0.0240)
Subordination	-3.130***	-3.944***	-3.978***
	(0.268)	(0.565)	(0.310)
Observations	26,242	14,034	12,206
Rating at first transaction	Y	N	Y
Vintage year	Y	Y	Y

Standard errors in parentheses

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

Table 1: Regression results from running logit regression 1 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include price, subordination, coupon and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Column (1) includes all issues; columns (2) and (3) split the sample between bonds rated AAA at origination and the rest, respectively. Errors are clustered at deal level.

Comparing the subsample of AAA bonds and the rest, which we do in Table 26, I find evidence of this monotonicity of the regression coefficient on subordination percentage for both AAA bonds and the rest. So while the effect of price is always negative and significant and that of coupon depends on whether the bond is AAA at origination, the effect of subordination depends on the quality of documentation on the underlying loans as measured by our opacity index. In order to weigh the relative contribution of these different components I will price the bonds using a Gaussian copula model. The outcome of the pricing model, namely the implied correlation, works as a summary statistic of the variables considered so far.

3 Implied correlation: a summary measure

I use the asymptotic single risk factor model implemented by the IRB approach in Basel II. Credit risk in this basic framework has two components, one systematic and the other idiosyncratic, so that correlation is captured by codependence on the realization of the systematic factor (Crouhy, Galai, and Mark, 2000). I use a Large Homogeneous Gaussian Copula (LHGC) model (Brunne, 2006; D'Amato and Gyntelberg, 2005; Duffie and Singleton, 2012; Elizalde, 2005; McGinty et al., 2004; Tzani and Polychronakos, 2008).[11] In the LHGC setup two assumptions apply: all loans in a given pool have the same (known) probability of default PD, and all have the same recovery rate RR. The homogeneity allows us to abstract from individual loan sizes, which we normalize to one. Consider a pool of N mortgages. Default times $\tau = \tau_1, \ldots, \tau_N$ are correlated random

[11] Following Li (2000) the Gaussian copula offered a conceptually simple framework for pricing structured securities, which made its use widespread. The model was also used for risk management, which Jarrow (2011) shows is inappropriate. The inapproapriate use of copulas is blamed for a surge in investor overconfidence and eventually set the stage for the financial crisis in 2007. See Felix Salmon, *Recipe for Disaster: the Formula that Killed Wall Street* (`https://www.wired.com/2009/02/wp-quant/`)

	(1) [0, 0.25)	(2) [0.25, 0.5)	(3) [0.5, 0.75)	(4) [0.75, 1]
	Downgrade indicator			
Price	-0.0159***	-0.0200***	-0.0110***	-0.0169***
	(0.00606)	(0.00333)	(0.00267)	(0.00354)
Coupon	-0.142**	-0.0380	-0.117***	-0.0780*
	(0.0640)	(0.0304)	(0.0441)	(0.0466)
Subordination	0.00163	-1.857***	-4.016***	-5.722***
	(0.864)	(0.657)	(0.489)	(0.943)
Observations	2,489	5,513	7,073	5,049
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 2: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include price, subordination, coupon and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

variables. Correlation is captured by the loading on one -exogenous- systematic factor S, which in this setting follows a standard normal distribution. I use a one-factor model, where the individual default probability is given by

$$p(s,T) := Pr(\tau \leq T|S = s) = \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho}s}{\sqrt{1-\rho}}\right). \tag{2}$$

In equation 2, PD is the unconditional default probability. Defaults are independent conditional on the realization of the systematic factor S, i.e.

$$Pr(\tau_1 \leq t, \ldots, \tau_N \leq t|S = s) = \prod_{k=1}^{N} Pr(\tau_k \leq t|S = s)$$

which simplifies computations.

Total losses from the pool accumulate over time to $l(t) = \frac{1}{N}\sum_{k=1}^{N}(1 - RR)1_{(\tau_k \leq t)}$. The losses are distributed along the tranches from the deal. A given tranche's position in the waterfall is characterized by its lower and upper attachment points a and b where $0 \leq a < b \leq 1$. Its notional is a proportion $b - a$ of the total pool notional N. The losses borne by this tranche are given by

$$l_{[a,b]}(t) = \frac{[l(t) - a]^+ - [l(t) - b]^-}{b - a}.$$

This exposure to risk affects the expected payoff of the CMO tranche. Using the recovery rate, equation (2) yields the following estimate of expected losses within the $[a, b]$ tranche by payment

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator - AAA only			
Price	-0.0352***	-0.0360***	-0.0347***	-0.0539***
	(0.00900)	(0.00529)	(0.00632)	(0.0127)
Coupon	0.0508***	0.0546	0.0919	0.118*
	(0.0161)	(0.0451)	(0.0575)	(0.0625)
Subordination	-0.0174	-2.774**	-2.014	-9.907***
	(1.622)	(1.229)	(1.881)	(3.612)
Observations	1,325	3,073	3,272	2,926
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y
	Downgrade indicator - not AAA			
Price	-0.0163**	-0.0129***	-0.00786***	-0.0113***
	(0.00714)	(0.00371)	(0.00250)	(0.00358)
Coupon	-0.367***	-0.167***	-0.201***	-0.156***
	(0.102)	(0.0475)	(0.0529)	(0.0603)
Subordination	-0.309	-2.648***	-4.501***	-4.193***
	(1.881)	(0.880)	(0.538)	(0.784)
Observations	1,038	2,248	3,757	2,111
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 3: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include price, subordination, coupon and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

date T_i:

$$E[l_{[a,b]}(T_i)] = \frac{1}{b-a}\int_{-\infty}^{\infty} \frac{e^{-s^2/2}}{\sqrt{2\pi}} \left([(1-RR)p(s,T_i)-a]^+ - [(1-RR)p(s,T_i)-b]^+\right) ds \quad (3)$$

Duffie and Gârleanu (2001) and Coval et al. (2009a) look at the sensitivity of expected recovery to default correlation. Figure 3.1 replicates the exercise in Coval et al. (2009a) by plotting expected recovery for each value of ρ, normalized by the value corresponding to $\rho = 20\%$.

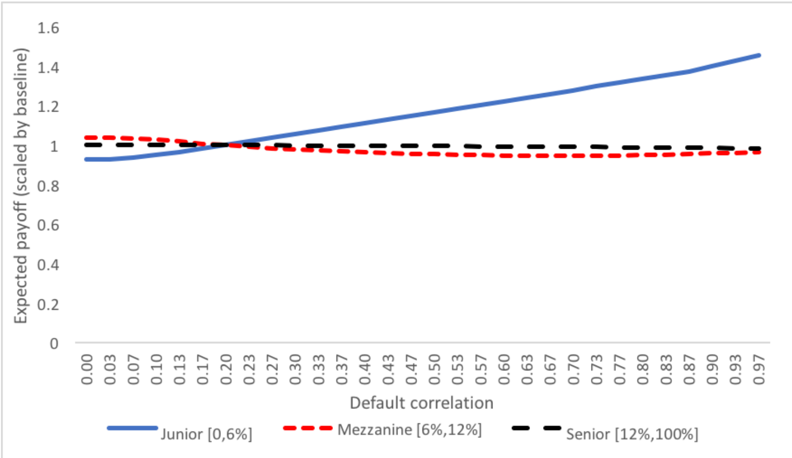


Figure 3.1: Sensitivity of a simulated CMO structure to default correlations. We plot the expected payoff within a given tranche, for each value of the underlying correlation ρ (parameters are PD=5% and LGD=50% as in Coval et al. (2009a)). The results are normalized by baseline estimate, based on the same parameters and a correlation $\rho = 20\%$. No prepayments are incorporated (i.e. SMM=0%) for comparability of outcomes.

Using payment dates $0 < T_1 < \cdots < T_m = T$ (where T is the maturity of the security), write the pricing equation of the security

$$\frac{V_{[a,b]}}{N(b-a)} = c\sum_{i=1}^{m} B(0,T_i)\Delta(T_{i-1},T_i)(1-l_{[a,b]}(T_i)). \quad (4)$$

Formula (4) equates current price to the sum (in expectation) of two terms: the discounted cash-flows from coupon payments and the residual value (after accounting for defaults) of principal outstanding. Here $B(t_1,t_2)$ discounts a payoff at t_2 to t_1, c denotes the tranche coupon and $\Delta(T_{i-1},T_i)$ is the time difference between two payment dates (for mortgage bonds we use $\Delta(T_{i-1},T_i) \equiv 1/12$). The pricing equation is then $pN(b-a) = E[V_{[a,b]}]$. Writing $e_i^{[a,b]} = E[1-l_{[a,b]}(T_i)]$ the following holds at origination:[12]

$$p_0 = c\sum_{i=1}^{m} B(0,T_i)\Delta(T_{i-1},T_i)e_i^{[a,b]} \quad (5)$$

[12] Note that formula (6) implies that default occurs immediately after the following period payment.

The pool is exposed to prepayment risk.[13] As prepayments happen, the coupon rate is applied to the balance outstanding, while the prepaid amount is allocated across tranches according to the order specified in the prospectus. In the absence of data about the order of the cashflows for each deal, I make the simplifying assumption that prepayments are uniformly distributed across tranches.[14] I obtain

$$p_t = \sum_{i=t+1}^{m} B(t,T_i) e_i^{[a,b]} \prod_{k=t+1}^{i-1} (1 - SMM_k) \left(\underbrace{c\Delta(T_{i-1},T_i)(1 - SMM_i)}_{\text{coupon payment}} + \underbrace{SMM_i}_{\text{prepaid principal}} \right) \tag{6}$$

where SMM_k is the single month mortality rate at time k, and is given by the prepayment speed model. Given the unconditional default probability PD, the recovery rate RR and prepayment rate SMM_k, pricing equation (6) pins down a value of ρ, the market estimate of default correlation for the given pool of loans. Note that expression (2) is only defined for $\rho \in [0,1)$ and thus the existence of a solution to equation (6) is not guaranteed for an arbitrary choice of p and c. So instead of solving the equation, we solve

$$min_{\rho\in[0,1)} \left| p_t - \sum_{i=t+1}^{m} B(t,T_i) e_i^{[a,b]} \prod_{k=t+1}^{i-1} (1 - SMM_k) \left(c\Delta(T_{i-1},T_i)(1 - SMM_i) + SMM_i \right) \right| \tag{7}$$

Expected losses are monotonically increasing in default correlation ρ for the senior tranche, and monotonically decreasing for the junior tranche (see Figure 3.1). The mezzanine tranche behaves like a senior tranche for low correlations and like a junior tranche for high ones (Ashcraft and Schuermann, 2008; Duffie, 2008).[15] This gives the market estimate of default correlations which I now compute on our panel of security prices.

3.1 Model parameters: default and prepayment

The present analysis is focused on expected losses (EL). Equation 3 uses the identity $EL = PD \times LGD$, where PD and LGD respectively denote the probability of default and the loss given default. Both factors must be based on the same definition of default. Since recoveries in our data are based on liquidated values, I use of liquidation as the default event.

Figure N.5 shows an increase in cumulative liquidation rates in the running to the crisis, though the trend is only upward sloping from 2005 vintages onward. Alt-A default rates were roughly half those of subprime deals until early 2005, when both rates soared in the running to the crisis. By 2008, securitization issuance have collapsed. One difference is that while the 90+ delinquency rate

[13]The Standard Prepayment Model of The Bond Market Association specifies a prepayment percentage for each month in the life of the underlying mortgages, expressed on an annualized basis. In early versions of the model I use the common assumption that prepayment speed is given by 150% PSA (see Figure N.8). The final version has a model for both PD and prepayment speed.

[14]Duffie and Singleton (2012) discuss two prioritization schemes (uniform and fast). Both imply prepayment cash flows are sequential over seniorities. I do not have deal-level information about the allocation of cash flows.

[15]For those cases two minima could arise in principle (as would also be the case if solving for equation (6) instead of (7)).

they report remains lower for Alt-A deals, I find that their cumulative liquidation rate, initially similar to that of prime deals, caught up with that of subprime ones in the running to the crisis.

From loss event data I compute LGDs at the deal level (see Figure N.12 for a count of observations by vintage and asset type). Figure N.6 shows that LGD was nearly monotonically increasing from 1990 onwards (except for a peak in 1996) in the running to 2007, so that investors may have been adjusting their expectations of LGD over the cycle. However, for LGDs to be computed the full post-workout must be observed, which usually takes a substantial observation time after default. Recent advances in modeling LGDs with incomplete workouts (see Rapisarda and Echeverry (2013)) are far from the norm, especially in the running to the crisis. I apply the common assumption of constant LGD, using the long run (weighted) average on our sample of 60% that is also typically assumed in the literature (Altman, 2006; Brunne, 2006; Coval et al., 2009b; Hull and White, 2004, 2008).

Investors' beliefs about default rates are elicited with a regression model establishing the likelihood of default as a function of loan covariates and estimated on default history. I use a proportional hazard model on a prepayment indicator to assess investors' beliefs about prepayment speeds. The model is estimated as a separable hazard model, treating observations representing default as censored as in Palmer (2015) and Liu (2016). Default and prepayment are termination reasons happening at a random time τ^{term}, whose intensity (for termination cause $term \in \{default, prepayment\}$) is given by equation (8).

$$\lambda_i^{term}(t) = \lim_{\epsilon \to 0} \frac{Pr_i(t-\epsilon < \tau^{term} \leq t \mid t-\epsilon < \tau^{term},\ X)}{\epsilon}. \tag{8}$$

Here i denotes loan, and t denotes time after origination. The density function in equation 8 is modeled as

$$\frac{\lambda_i^{term}(t)}{\lambda_0^{term}(t)} = \exp(X_{it}'\beta^{term}) \tag{9}$$

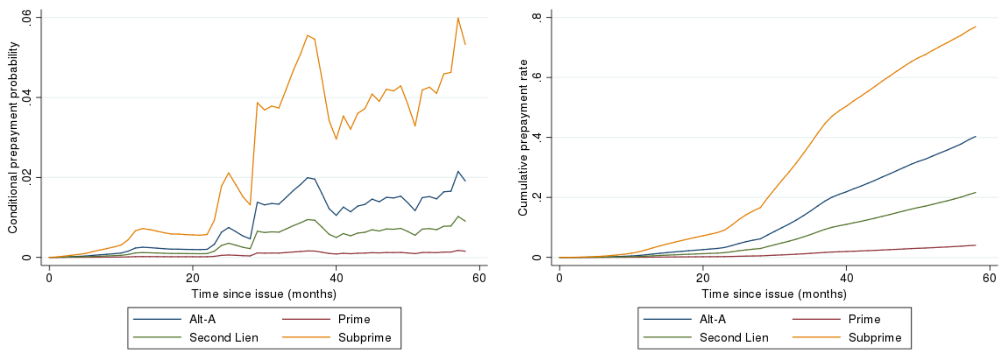
where $\lambda_0^{term}(t)$is the baseline hazard function that depends only on the time since origination t. Covariates in X_{it} include loan attributes (loan amount, coupon gap relative to 10 year constant maturity Treasury, LTV, prepayment penalty indicator), agent characteristics (FICO score, owner occupancy) and variables at the CBSA level such as home price appreciation and unemployment rate. The exponential model specified in equation 8 has a continuous time specification. To estimate it on discrete time data I accumulate the intensity process λ over time intervals per equation (10).

$$Pr_i(t < \tau^{term} \mid t-1 < \tau^{term}) = \exp\left(-\int_{t-1}^{t} \lambda_i^{term}(u)du\right) \tag{10}$$

This leads to the complementary log-log specification in equation (11):

$$Pr_i(t < \tau^{term} \mid t-1 < \tau^{term}) = \exp(-\exp(X_{it}'\beta^{term})\lambda_0^{term}(t)) \tag{11}$$

I estimate specification (11), with month since origination fixed effects to obtain the hazard functions over the first 60 months of the loan. I document the results in Table 27 and plot the resulting prepayment rates on Figure 3.2. I find that adjustable rate mortgages are both more likely to default and to prepay than fixed rate types. Subprime loans are the asset type most likely to default. In terms of prepayment hazard, there is no significant difference across asset types other than prime loans being less subject to prepayment than other types.



(a) Prepayment hazard

(b) Cumulative prepayment rate

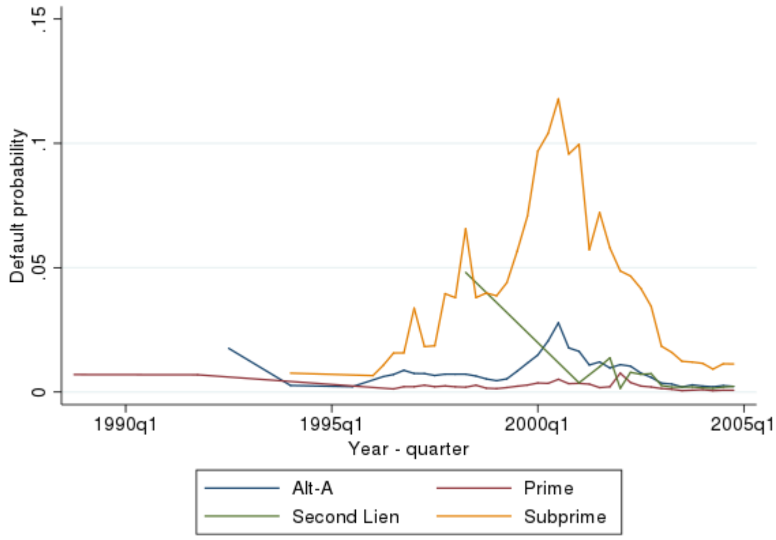
Figure 3.2: Marginal and cumulative prepayment rates implied from the model (11), as summarized in Table 28. Using loan covariates at origination, prepayment hazard rates are computed at the loan level. Averages are computed by asset type and month after origination, and plotted here.

I now compare the results results from Table 27 with the ones obtained by Liu (2016), who uses the same model to estimate default and prepayment hazard rates on loans backed by the government-sponsored entities (Fannie Mae and Freddie Mac).[16] On one hand, we find the same sign for the effect of FICO score, the difference between the original loan interest rate and the original 10 year rate and the unemployment rate. Moreover, in terms of default hazard we find similar effects of LTV and home price appreciation.

On the other hand we find a few differences, mostly about the link between home prices and prepayment rates. Liu (2016) finds that home price appreciation increases prepayment hazard while we find the opposite. Similarly, he finds that higher LTV reduces prepayment hazard while we find no clear link. As discussed by Gorton (2009), while the prepayment option is always

[16] Adding late originations (up to 2007) we find a number of similarities. The main difference that arises is that now subprime loans can be seen to be prepaying significantly more than other types, and significantly more than early vintages. This suggests that the link between subprime origination and home prices through prepayments was specific to the pre-crisis boom rather than a constitutive characteristic of subprime loans from their inception. Macroeconomic factors such as home price appreciation and unemployment exhibit a similar effect on defaults and prepayments when adding late vintages. Instead, for coupon gap there is a change compared to the early sample. The coupon gap, i.e. the change in 10 year rates between origination and present, reflects stronger incentives to refinance. The expectation is that this leads to a higher probability of prepayment and a lower probability of default, which we see once we add late cohorts but not in the early sample.

valuable for prime, 30-year fixed rate mortgages (i.e. if house prices rise borrowers build up equity), for subprime loans lenders hold an implicit option to benefit from house price changes. Table 27 shows prepayment penalties, this being the way in which the lender exercises its option, are a strong deterrent against this termination type.



(a) Probability of default

Figure 3.3: Probability of default implied from the complementary log-log model, estimates of which are in Table 28. Using loan covariates at origination, default probabilities are computed at the loan level. Averages are computed by asset type and month after origination, and plotted here.

The break-even probabilities of a crisis computed by Beltran, Cordell, and Thomas (2017) from CDO prices show a decrease from early cohorts (pre 2006 per their definition) to late ones, which suggests a relatively high risk premium was charged in early cohorts. Though there are no studies on risk premia in mortgage markets, I can benchmark against the corporate market. Berndt, Douglas, Duffie, Ferguson, and Schranz (2005) derive actual and risk-neutral probabilities from CDS market quotes. They find that the corresponding coverage factors (ratio of risk neutral probability to real probability) oscillate between 1.5 and 3.5 over time, between 2002 and 2003. I use a coverage ratio of 3.[17]

Using the model in Table 28 I predict prepayment hazards and default probabilities at the loan level, and average them at the deal level. Both the default probability and the hazard rate are estimated deal by deal (in Section K I use a constant PD and prepayment speed, as a robustness check). As for the prepayment hazard, we will use the full schedule in order to estimate the average prepayment speed for the given deal over the first 60 months. As Figure 3.2 illustrates, subprime loans have the highest prepayment rates, followed by Alt-A loans. They also have the highest default probabilities, as shown in Figure 3.3. We use the model-implied PDs from Table 28 (see Figure 3.3) and include them as controls in our regressions.

[17]Heynderickx, Cariboni, Schoutens, and Smits (2016) quantify coverage factors from CDS quotes of European corporates and find that they range between 1.27 for Caa (Moody's) ratings to 13.51 for Aaa ones on pre-crisis data. Like Heynderickx et al. (2016), Denzler, Dacorogna, Müller, and McNeil (2006) argue that risk spreads exhibit a scaling law, whereby risk premia are decreasing in the probability of default. The results in Table 39 imply coverage ratios between 2.03 for subprime deals and 3.27 for Alt-A ones, in line with the literature.

We source contractual maturity from Bloomberg, which for most bonds is close to 30 years. These figures are high compared with realized maturity (defined as the first observation where the tranche balance is zero) the difference being 16.27 years on average on a sample of 5,507 tranches. Figure N.7 also suggests that bonds do not live that long on average. Adelino (2009) uses weighted average life (WAL) instead of contract maturity, which is closer to the realized maturity. We also source WAL for a sample of our loans where we could find it, but found that WALs are low compared to realized maturities in the data (the average difference is 6.77 years on a sample of 16,894 tranches, see Figure N.10 for a further breakdown of the difference). We will use contractual maturity, relying on the prepayment speed model to achieve an accurate reduction of tranche balance over time.

The model in Table 28 incorporates all observations. In reality, agents' expectations about default evolve over time, especially as the business cycle unfolds. As an example take home prices, which fluctuate over the cycle. As Table 9 shows, home price appreciation is the variable whose effect on defaults changes the most over the cycle. In particular, the negative relationship between price appreciation and defaults documented in Table 28 is an average between the positive effect recorded in the early years of the sample (up to 2002) and the negative effect in subsequent years. I expect this to have a modest impact on the pricing model, given that over the times of the prices we are interested in (mostly 2004 and 2005) the coefficients in Table 9 tend to be close to those in Table 28.

Loan performance data gives a basis for consensus about probability of default, loss given default and prepayment speed. Instead, as discussed in the introduction, default correlation is the parameter market participants are more likely to disagree about[18]. Seeing these disagreements as the starting point for differential information, we will use the pricing model from Section 2 to generate a summary statistic that signals future downgrades, and study how asset opacity drives the informativeness of the signal.

3.2 Implied default correlations from CMO data

For a given bond we compute its implied correlation ρ using the coupon rate c, market price p, attachment point and detachment point $a \geq 0$ and $a < b \leq 1$. The probability of default and prepayment speed are estimated per Section 3.1. The recovery rate is $RR = 60\%$. We use the discount rate $r = 4.27\%$, the average 10-year constant maturity treasury (annual) rate between 1995 and 2015. The numerical computations of loss probability are evaluated using a trapezoidal rule, which Brunne (2006) deems faster than Gauss-Legendre and Gauss-Hermite methods. Figure N.14 provides a summary of observations.

The distribution of individual outcomes is bimodal (see Figure N.13). Tzani and Polychronakos (2008) find that in CDS markets model correlations would often have had to exceed 100% in order to price supersenior tranches, which is suggested by Figure K.1. The extreme values suggest there is a role for market incompleteness as in Andreoli et al. (2016) and Stanton and Wallace (2011).

[18] "Currently, the weakest link in the risk measurement and pricing of CDOs is the modeling of default correlation." Duffie (2008)

Figure N.14 also shows evidence of a correlation smile in prices both before and after the crisis.[19]

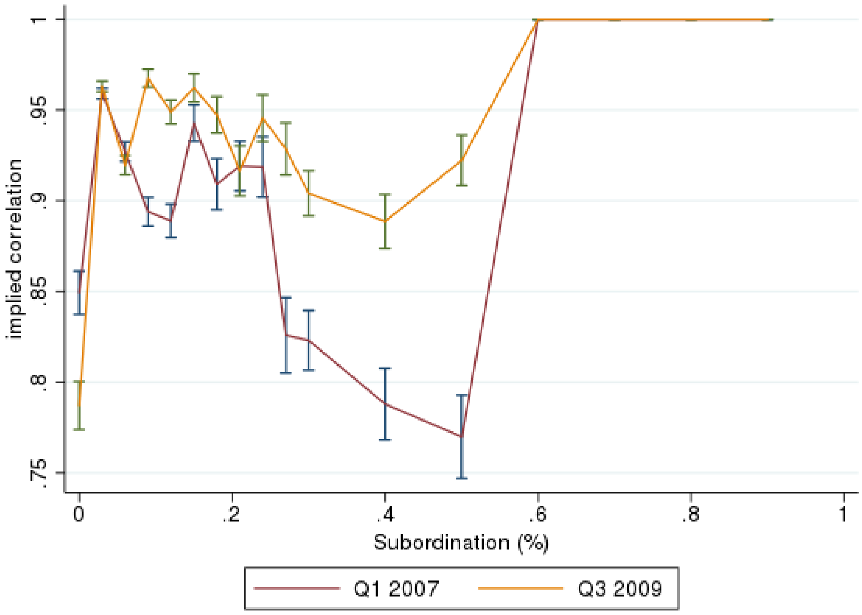


Figure 3.4: Average correlation plotted against tranche subordination percentage, on two given dates. We use the sample of early vintage bonds (originated prior to June 2005). Subordinations are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.

Using a one factor Gaussian copula model, Buzková and Teplỳ (2012) analyze prices of the 5-year, North American investment grade CDX (V3) index between September 2007 and February 2009. They report that for synthetic CDOs, implied correlations show a large increase, from 0.15 to 0.55 on average over that time period. In comparison, we observe a significant increase over the same period, though of smaller magnitude (from 0.89 to 0.93). Breaking the change by asset type we see an increase for Alt-A tranches (from 0.81 to 0.97, significant at 99%) and for subprime deals (from 0.85 to 0.89, significant at 99%) and no change for prime ones (0.93). The upward adjustment was thus the largest for Alt-A issues (see Figure N.15). In terms of seniorities, the difference observed by Buzková and Teplỳ (2012) over the crisis is mainly driven by mezzanine tranches (7%-10% and 10%-15%). Figure N.14 also suggests the increase in correlations is larger among intermediate seniorities.

[19]The correlation smile is an artifact from the compound correlation method (O'Kane and Livesey, 2004). A method that is used to derive increasing correlations is the base correlation, which is computed as follows: let the attachment points in the full waterfall be given by $(b_1, \dots, b_n)$, where $b_n = 1$. First, solve 6 for the tranche $[0, b_k]$, $k = 1 \dots n$. This gives an estimate of $e_i^{[0,b_k]}$. Using the identity

$$(b-a)e_i^{[a,b]} = be_i^{[0,b]} - ae_i^{[0,a]},$$

the expected losses in tranche $[a, b]$ can be sequentially computed along the waterfall: once the $[b_{k-1}, b_k]$ tranche has been priced, the following one can be priced using

$$(b_{k+1} - b_k)e_i^{[b_k,b_{k+1}]} = b_{k+1}e_i^{[0,b_{k+1}]} - b_k e_i^{[0,b_k]}.$$

Base correlations price all tranches in a deal simultaneously, and thus do not use base correlations because we are pricing tranches that trade separately over time.

We now consider the trend over time (see Figure 3.5). Ratings were mostly stagnant ahead of the crisis, especially for AAA tranches, in comparison with default correlations. BBB tranches even see an improvement in ratings ahead of the crisis while correlations are increasing (except for subprime deals, which see both downwards and upwards changes). The sharpness of rating downgrades suggests this is a concern for BBB tranches. A potential explanation is given by Griffin and Tang (2012), who argue that AAA ratings were inflated in CDO securities, with optimistic ratings applied to a large share of bonds issued. Because CDOs are mainly composed of CMO tranches, a potential channel for rating inflation in AAA CDO tranches is rating inflation in the underlying BBB tranches *after* origination. This gives a possible channel for ratings inflation over the cycle other than that of boom time originations.

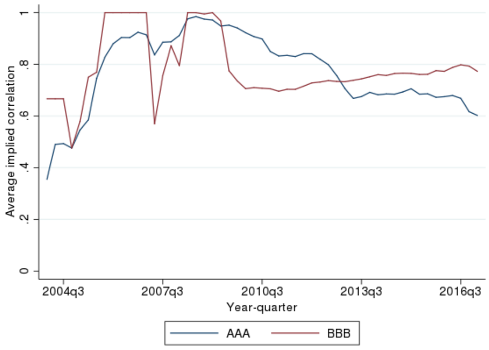
4 The information content of implied correlations

This section will focus on whether correlations implied from early prices are informative of subsequent downgrades. Using this data I first replicate the findings by Ashcraft et al. (2011) that market prices contain information about bond performance which is not captured by the agency ratings. Then we replicate the result in Adelino (2009) that the information content is a priori less significant for AAA tranches than for non-AAA tranches. We start with a logit specification similar to that in Adelino (2009), where the dependent variable is whether bond i was downgraded by December 2009. More specifically,

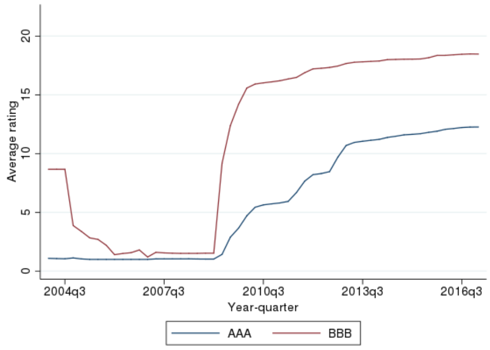
$$downgrade_{i,2009} = f(\alpha + \beta\rho_{i0} + \eta_{rating_{i0}} + \gamma X_{i0} + \varepsilon_i). \tag{12}$$

The independent variable of interest is the implied correlation at the first transaction, ρ_{i0}. High correlations are detrimental to senior bondholders but beneficial to subordinate ones (Duffie and Gârleanu, 2001). In line with this we expect that (except for bonds with zero subordination percentage, which we do not often observe) a higher implied correlation should predict a more likely downgrade. We control for rating at origination using dummy indicators and for vintage year. Also we cluster standard errors in all tests at the deal level, to control for the fact that several classes in the same deal are often (down)graded at the same time.

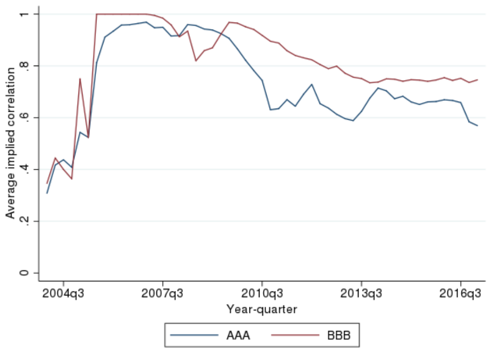
The results in Table 29 replicate the findings by Ashcraft et al. (2011) that ratings at origination are not statistically sufficient for implied correlations (in their case, coupon premium) in predicting subsequent bond downgrades. Their proxy for the bond price is the coupon premium to treasury, the hypothesis being that higher premium is reflective of more risk and thus of more downgrades. The implied correlation gives a similar intuition. I find a positive, significant coefficient, so that higher implied correlation increases the likelihood of downgrades. Table 29 breaks down this result between bonds initially rated AAA and the rest. While the coefficient for correlation at first transaction remains significant for grades below AAA, implied correlations seem to have no predictive power in terms of bond downgrades, similar to the findings in Adelino (2009).



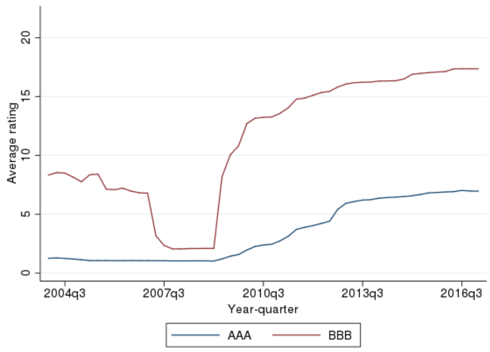
(a) Implied correlation - Alt-A



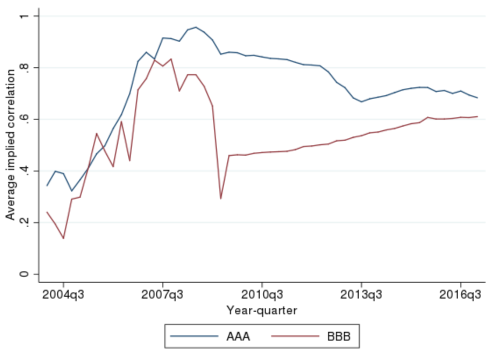
(b) Rating - Alt-A



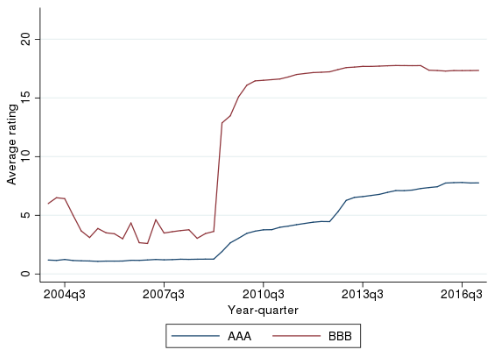
(c) Implied correlation - prime



(d) Rating - prime



(e) Implied correlation - subprime



(f) Rating - subprime

Figure 3.5: Performance of early vintage tranches: average implied correlation and average rating for bonds originated before June 2005. For a given tranche we compute the implied correlation, at each point in time. The average is taken by transaction period, by coarse rating at origination (AAA=1,..., BBB=4,..., D=8).

	downgrade		
	(1)	(2)	(3)
	All	AAA only	Non-AAA only
Correlation at first transaction	0.414***	0.299	0.268***
	(0.0629)	(0.201)	(0.0644)
Model-implied PD	2.294**	4.308	1.503
	(0.922)	(3.648)	(1.023)
Observations	28,991	16,618	12,371
Rating at first transaction	Y	N	Y
Vintage year	Y	Y	Y
Asset type	Y	Y	Y

Standard errors in parentheses

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

Table 4: Regression results from running logit regression 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default, as estimated in subsection 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Column (1) includes all issues; columns (2) and (3) split the sample between bonds rated AAA at origination and the rest, respectively. Errors are clustered at deal level.

I use my opacity index to break down the sample by increments of 0.25, and present the results in Table 30. I find a ranking along the index similar to the one discussed in Section 2, whereby the coefficient on implied correlations is monotonically increasing in the value of the opacity index, from insignificant at 10% for tranches below 0.25 to positive and significant at 1% for tranches above 0.75.

Breaking down the results in Table 30 between AAA tranches and others. The results, shown in Table 31, show a similar pattern across the two rating categories. For tranches where the documentation index is above 0.5 implied correlation is predictive of bond downgrades. Seen together, the results suggest that uninformed investors are not so much those in AAA tranches as those subject to "low-doc" deals.

As a robustness check, I run the same set of regressions as before, using the deal level average correlation (clustering errors at deal level) instead of the tranche implied correlation. I control for initial prices, coupons and subordinations. The results, shown in Table 11, suggest that correlation loses its predictive power when averaged across the deal. The average at rating level, instead, retains some predictive power about subsequent downgrades. Breaking the results down by opacity index in Table 12, I find the same monotonicity in predictiveness of implied correlations, though the coefficient becomes significant only for the highest values of the documentation index. However, once I break down the results between AAA and sub-AAA tranches in Table 13, only AAA tranche implied correlations are predictive (still, only in the highest documentation index values). In all these tables, the monotonicity property observed before is best represented by the subordination percentage.

Low-doc assets should in principle require a form of compensation: all else constant, a sophisticated investor requires more subordination when the underlying assets are opaque. Instead, Skreta and

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator			
Correlation at first transaction	0.243	0.605***	0.476***	0.569***
	(0.250)	(0.200)	(0.102)	(0.135)
Model-implied PD	0.381	13.60	4.331	4.225*
	(1.675)	(10.51)	(3.000)	(2.521)
Observations	2,723	6,285	7,808	5,565
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 5: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default as estimated in subsection 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the value of the documentation index corresponding to the given deal. Errors are clustered at deal level.

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator - AAA only			
Correlation at first transaction	1.018	0.430	1.647***	0.842***
	(0.703)	(0.599)	(0.627)	(0.321)
Model-implied PD	47.95	-13.93	12.91***	3.301
	(48.60)	(45.34)	(4.648)	(2.970)
Observations	1,529	3,765	3,975	3,429
Rating at first transaction	N	N	N	N
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y
	Downgrade indicator - not AAA			
Correlation at first transaction	0.0485	0.370**	0.314***	0.353**
	(0.283)	(0.155)	(0.109)	(0.158)
Model-implied PD	-2.323	26.14**	1.906	5.083
	(2.661)	(10.85)	(2.607)	(3.180)
Observations	1,045	2,289	3,787	2,124
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 6: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default as estimated in subsection 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the value of the documentation index corresponding to the given deal. Errors are clustered at deal level.

Veldkamp (2009) predict that rating inflation is worse when assessing the true value of the asset is difficult (making ratings noisier and more varied). For their result to hold, investors must be unable to infer the rating selection bias. Similarly in our case, investors who are unaware of the deficiency in documentation are more likely to be subjected to inflated ratings. Table 32 provides evidence that AAA share at origination is decreasing in our opacity index (controlling for the model-implied probability of default). This suggests that unsophisticated investors select into low-doc deals, where rating inflation is more likely to occur.

5 Summary and discussion

This paper assesses the relative importance of two key information frictions that take place between the investor and the securitizer. Though there is a role for what the literature calls investor unsophistication, proxyed by a AAA rating at origination, asset opacity is the predominant friction. I capture this using a deal-level index of documentation completeness. I observe less of a differential in information content across seniorities than across low-doc assets and "full-doc" ones. The latter exhibit better information content across the rating spectrum. In particular, AAA implied correlations are no less predictive than the rest when the bond comes from a deal with a high standard of documentation. The evidence suggests that more opaque deals tend to issue a higher proportion of AAA bonds, controlling for risk attributes of the deal, consistent with ratings inflation. Accordingly, the results suggest that errors in computing default correlations in the running to the crisis were not a problem of AAA investors per se, but rather a problem of "low-doc" investors.

Implied correlations are large in subprime deals compared to other asset classes, which reflects a design feature of subprime loans that made them jointly dependent on house prices. The literature has historically attributed default clustering to joint dependence on a systematic shock (Bisias, Flood, Lo, and Valavanis, 2012; Chan-Lau, Espinosa, Giesecke, and Sole, 2009; Bullard, Neely, Wheelock, et al., 2009; Khandani, Lo, and Merton, 2013). I have followed this approach, using a Gaussian copula. Recent literature distinguishes two additional sources of default clustering: unobserved frailty (Duffie et al., 2009; Kau, Keenan, and Li, 2011; Griffin and Nickerson, 2016) and contagion (see appendix M).[20] In particular Azizpour, Giesecke, and Schwenkler (2016); Gupta (2016) and Sirignano, Sadhwani, and Giesecke (2016) suggest the contagion channel is important. In light of this literature, this paper is the first of several steps to understand which sources of default clustering are priced in mortgage markets.

[20]For a review of recent literature on contagion see Bai, Collin-Dufresne, Goldstein, and Helwege (2015).

A Default and prepayment models

	with data up to 2004		with data up to 2007	
	(1)	(2)	(3)	(4)
	Default	Prepayment	Default	Prepayment
log(FICO)	-1.468***	1.408***	-2.076***	0.305**
	-0.157	-0.155	-0.199	-0.12
owner occupied	0.039	-0.024	-0.098*	0.024
	-0.05	-0.02	-0.054	-0.02
original r - original 10 year rate	0.475***	0.249***	0.252***	0.066***
	-0.01	-0.017	-0.011	-0.006
log(original amount)	0.421***	0.257***	0.143***	0.02
	-0.043	-0.031	-0.041	-0.026
log(original LTV)	0.439***	-0.007	0.183***	0.069***
	-0.043	-0.036	-0.033	-0.02
prepayment penalty	-1.866***	-1.034***	-0.914***	-0.950***
	-0.08	-0.073	-0.031	-0.025
adjustable rate mortgage	0.655***	0.493***	0.367***	0.467***
	-0.062	-0.047	-0.038	-0.015
log(Cumulative HPA)	-8.398***	-7.780***	-6.482***	-2.474***
	-1.041	-0.963	-0.652	-0.41
coupon gap	0.400***	0.120*	-0.255***	-0.144**
	-0.05	-0.062	-0.04	-0.06
unemployment	0.330***	0.320***	0.201***	0.319***
	-0.072	-0.075	-0.068	-0.075
Asset type: Prime	-1.008***	-0.147***	-1.130***	-0.603***
	-0.078	-0.027	-0.078	-0.033
Asset type: Second Lien	-0.580***	0.124	0.843***	0.385***
	-0.142	-0.079	-0.064	-0.028
Asset type: Subprime	0.504***	-0.021	1.113***	0.201***
	-0.053	-0.05	-0.037	-0.02
CBSA FE	Y	Y	Y	Y
Month since origination FE	Y	Y	Y	Y
Observations	68,634,789	76,206,672	121,236,208	126,625,633

Standard errors in parentheses

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

Table 7: This table shows estimates using the maximum likelihood estimation of the complementary log-log specification in (11), using a nonparametric baseline hazard, on the loan level data available from ABSNet for private label loans (purchases only). The model treats competing risks independently, indicating 1 for failure and 0 for censoring. Each coefficient is the effect of the corresponding variable on the log hazard rate for either the default or prepayment of a mortgage. The conditional hazard is captured by performance month dummies, where performance is tracked over the first 60 months of the sample. The sample is truncated at December 2004 for columns (1) and (2), and at June 2007 for columns (3) and (4). Errors are clustered at CBSA level.

B Data cleaning

B.1 Bond data

We start with 16,397,826 panel observations, corresponding to 127,963 tranches. I remove data entry errors such as subordination percentages larger than one. In those cases all observations for

	(1) default	(2) prepayment
log(FICO)	-2.481***	0.448***
	-0.064	-0.018
owner occupied	0.025*	0.372***
	-0.014	-0.005
original r - original 10 year rate	0.429***	-0.011***
	-0.004	-0.001
log(original amount)	0.137***	0.324***
	-0.01	-0.003
log(original LTV)	0.572***	0.183***
	-0.012	-0.005
adjustable rate mortgage	0.487***	0.579***
	-0.016	-0.004
log(Cumulative HPA)	-1.826***	-1.581***
	-0.051	-0.011
coupon gap	0.848***	-0.261***
	-0.007	-0.002
unemployment	0.080***	0.001
	-0.004	-0.001
Asset type: Prime	-0.808***	-2.719***
	-0.044	-0.014
Asset type: Second Lien	-0.794***	0.298***
	-0.038	-0.011
Asset type: Subprime	0.402***	1.079***
	-0.025	-0.005
CBSA FE	N	N
Month since origination FE	N	Y
Observations	2,630,290	76,374,400

Standard errors in parentheses

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

Table 8: This table shows estimates using the maximum likelihood estimation of a complementary log-log specification, using a hazard specification for prepayments and a dummy indicator for default, on the loan level data available from ABSNet for private label loans (purchases only). The hazard model treats default risk as censored. Each coefficient is the effect of the corresponding variable on the log hazard rate for prepayment or the log probability of default of a mortgage. The conditional hazard is captured by performance month dummies, where performance is tracked over the first 60 months of the sample. The sample is truncated at December 2004.

	default indicator (by the end of the given year)									
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
log(FICO)	-2.485***	-3.599***	-4.816***	-2.583***	-3.141***	-3.682***	-4.561***	-4.160***	-3.029***	-2.059***
	(0.494)	(0.257)	(0.151)	(0.101)	(0.071)	(0.052)	(0.038)	(0.027)	(0.018)	(0.014)
owner occupied	-0.318**	0.037	0.263***	0.133***	-0.214***	-0.329***	-0.333***	-0.247***	-0.097***	-0.148***
	(0.139)	(0.069)	(0.041)	(0.024)	(0.016)	(0.011)	(0.008)	(0.006)	(0.004)	(0.003)
original r - original	-0.052	0.277***	0.199***	0.431***	0.459***	0.343***	0.164***	0.178***	0.158***	0.102***
10 year rate	(0.038)	(0.018)	(0.011)	(0.006)	(0.004)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
log(original amount)	-0.053	-0.038	-0.235***	0.125***	0.150***	-0.026***	-0.297***	-0.126***	-0.029***	-0.013***
	(0.084)	(0.043)	(0.025)	(0.016)	(0.011)	(0.007)	(0.005)	(0.003)	(0.002)	(0.002)
log(original LTV)	0.828***	0.698***	0.585***	0.772***	0.682***	0.548***	0.445***	0.178***	0.124***	0.078***
	(0.266)	(0.099)	(0.030)	(0.019)	(0.014)	(0.010)	(0.007)	(0.004)	(0.003)	(0.002)
adjustable rate	-0.707***	0.145**	0.305***	0.335***	0.261***	0.269***	0.291***	-0.045***	-0.130***	0.001
mortgage	(0.104)	(0.058)	(0.035)	(0.023)	(0.016)	(0.011)	(0.008)	(0.006)	(0.004)	(0.003)
log(cumulative HPA)	1.921***	2.981***	4.548***	-3.303***	-1.878***	-0.877***	0.412***	-1.998***	-5.796***	-4.319***
	(0.676)	(0.248)	(0.122)	(0.103)	(0.054)	(0.030)	(0.018)	(0.017)	(0.011)	(0.007)
coupon gap	-1.930***	0.216***	-0.591***	1.234***	0.998***	0.832***	0.170***	-1.057***	-0.810***	0.889***
	(0.062)	(0.037)	(0.019)	(0.013)	(0.009)	(0.006)	(0.005)	(0.004)	(0.002)	(0.002)
unemployment	0.137***	-1.052***	-0.342***	-0.080***	0.011**	0.126***	0.176***	0.004***	-0.183***	-0.309***
	(0.037)	(0.026)	(0.014)	(0.009)	(0.006)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Asset type: Prime	0.000	-1.048***	-0.805***	-0.748***	-0.621***	-0.299***	-0.248***	-0.669***	-1.456***	-1.640***
	(.)	(0.245)	(0.100)	(0.068)	(0.049)	(0.035)	(0.028)	(0.024)	(0.018)	(0.011)
Asset type: Second Lien	0.000	-3.509***	-5.936***	-4.410***	-1.984***	-0.213***	0.471***	0.877***	0.639***	0.616***
	(.)	(1.012)	(1.002)	(0.271)	(0.062)	(0.027)	(0.018)	(0.012)	(0.007)	(0.005)
Asset type: Subprime	2.872***	0.939***	0.213***	0.147***	0.274***	0.438***	0.742***	1.027***	0.777***	0.710***
	(0.311)	(0.128)	(0.062)	(0.040)	(0.028)	(0.019)	(0.014)	(0.010)	(0.005)	(0.004)
Obs	230,631	516,866	865,545	1,435,035	2,630,290	4,307,739	5,766,680	6,014,866	6,014,866	6,014,866

Standard errors in parentheses

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

Table 9: This table shows estimates using the maximum likelihood estimation of a complementary log-log specification, using a dummy indicator for default, on the loan level data available from ABSNet for private label loans (purchases only). For each year, variables are taken at the measurement point (either default time, if defaulted, or observation time, which is the end of the given year).

	(1)	(2)	(3)	(4)
	AAA balance at origination as share of deal issuance			
Opacity index	-0.104***	-0.0835***	-0.101***	-0.0259*
	(0.0154)	(0.0153)	(0.0151)	(0.0151)
Observations	1,902	1,902	1,902	1,902
Model-implied PD	N	Y	Y	Y
Vintage year	N	N	Y	Y
Asset type	N	N	N	Y

Standard errors in parentheses

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

Table 10: Regression results from running a linear regression at deal level of AAA origination (as share of total) on the deal opacity index. Controls include model-implied PD, vintage year (we include vintages up to June 2005) and asset type.

	downgrade		
	(1)	(2)	(3)
	All	AAA only	Non-AAA only
Deal average correlation	0.211		
	(0.189)		
Average correlation within rating bucket		0.721*	0.369*
		(0.427)	(0.217)
Price	-0.0185***	-0.0446***	-0.00937***
	(0.00152)	(0.00305)	(0.00150)
Coupon	-0.123***	-0.0402	-0.183***
	(0.0178)	(0.0245)	(0.0240)
Subordination	-3.168***	-4.144***	-4.113***
	(0.271)	(0.601)	(0.316)
Observations	26,242	14,034	12,206
Rating at first transaction	Y	N	Y
Vintage year	Y	Y	Y
Asset type	Y	Y	Y

Standard errors in parentheses

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

Table 11: Regression results from running logit regression 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default, as estimated in section 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Column (1) includes all issues; columns (2) and (3) split the sample between bonds rated AAA at origination and the rest, respectively. Independent variables include deal level average correlation (column 1), AAA average correlation (column 2), sub-AAA average correlation (column 3) and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Errors are clustered at deal level.

	(1) [0, 0.25)	(2) [0.25, 0.5)	(3) [0.5, 0.75)	(4) [0.75, 1]
	Downgrade indicator			
Deal average correlation	-0.581	-0.155	-0.0162	0.901**
	(0.750)	(0.522)	(0.367)	(0.395)
Price	-0.0165***	-0.0202***	-0.0110***	-0.0167***
	(0.00631)	(0.00343)	(0.00269)	(0.00356)
Coupon	-0.134**	-0.0369	-0.117***	-0.0749
	(0.0649)	(0.0310)	(0.0442)	(0.0465)
Subordination	-0.0512	-1.869***	-4.013***	-5.858***
	(0.852)	(0.653)	(0.489)	(0.959)
Observations	2,489	5,513	7,073	5,049
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 12: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default as estimated in section 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include deal level average implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the value of the documentation index corresponding to the given deal. Errors are clustered at deal level.

the month (all tranches linked to the deal involved) are removed so as to ensure computations of the tranching structure are correct.[21] We follow Adelino (2009) in removing Interest Only, Principal Only, Inverse Floater and Fixed to Variable bonds from the sample.

Notice that the most aggressive cleaning step is the removal of observations where price is missing. As dicussed in section 1, this is due to the data gap that covers late (2005 and more recent) vintages.

B.2 Loan level data

We start with a set of 22,008,610 loan originations. Of our originations set, 21,759,836 map to one of our deal IDs. Below is a summary of deal level averages of certain covariates (FICO score, LTV, private mortgage insurance coverage percentage) are computed.[22]

Historic data are contained in monthly reports. From the input 21,996,382 facilities we have at least one observation for 17,350,072 of them. We recover a total 792,664,139 loan-month observations from payment history (on average 45.7 obs per loan). From there we can compute default rates at deal level. We have loss event data for 3,986,974 observations, linked to 5,965 deal IDs. From there we can compute LGDs at deal level or vintage level.

[21] I manually computed subordination percentages on a random sample of deals to check the calculations by ABSNet.

[22] Simple averages were preferred over weighted averages (weighted by e.g. the initial securitized balance) as this reduces the number of missing observations.

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator - AAA only			
AAA average correlation	5.310	0.701	1.074	3.019***
	(3.508)	(1.594)	(0.962)	(0.857)
Price	-0.0309***	-0.0351***	-0.0328***	-0.0535***
	(0.00991)	(0.00542)	(0.00655)	(0.0118)
Coupon	0.0452***	0.0544	0.0898	0.149**
	(0.0173)	(0.0453)	(0.0580)	(0.0726)
Subordination	0.610	-2.791**	-2.283	-11.79***
	(1.530)	(1.260)	(1.851)	(4.377)
Observations	1,325	3,073	3,272	2,926
Rating at first transaction	N	N	N	N
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y
	Downgrade indicator - not AAA			
Below-AAA average correlation	-1.266	0.527	-0.0727	0.195
	(0.886)	(0.548)	(0.431)	(0.574)
Price	-0.0154**	-0.0127***	-0.00783***	-0.0115***
	(0.00704)	(0.00370)	(0.00249)	(0.00362)
Coupon	-0.357***	-0.174***	-0.201***	-0.155***
	(0.103)	(0.0496)	(0.0529)	(0.0599)
Subordination	0.133	-2.715***	-4.482***	-4.278***
	(1.905)	(0.878)	(0.541)	(0.852)
Observations	1,038	2,248	3,757	2,111
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 13: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default as estimated in section 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include AAA average correlation (upper panel), sub-AAA average correlation (lower panel) and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the value of the documentation index corresponding to the given deal. Errors are clustered at deal level.

	(1)	(2)	(3)	(4)
	AAA balance at origination as share of deal issuance			
Opacity index	-0.104***	-0.0835***	-0.101***	-0.0259*
	(0.0154)	(0.0153)	(0.0151)	(0.0151)
Observations	1,902	1,902	1,902	1,902
Model-implied PD	N	Y	Y	Y
Vintage year	N	N	Y	Y
Asset type	N	N	N	Y

Standard errors in parentheses

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

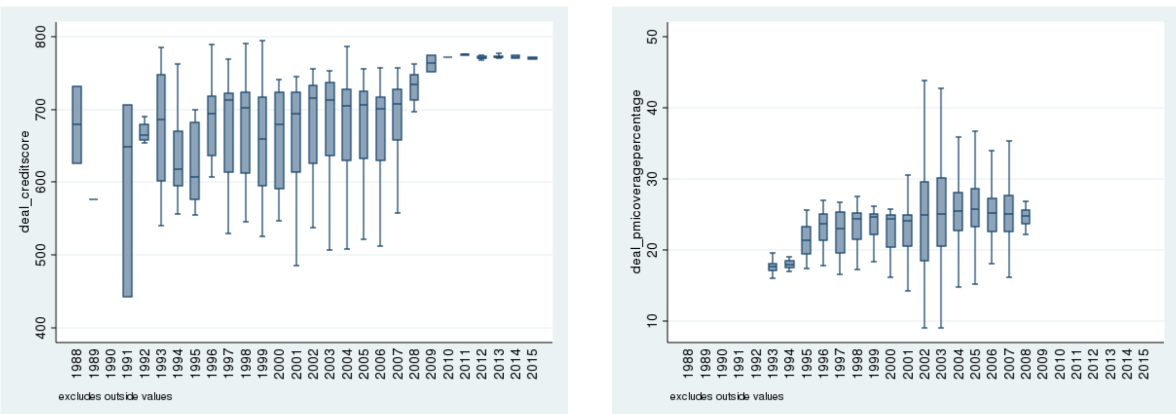
Table 14: Regression results from running a linear regression at deal level of AAA origination (as share of total) on the deal opacity index. Controls include model-implied PD, vintage year (we include vintages up to June 2005) and asset type.

Stage	Tranches left
Remove deals that are entirely made of mixed asset types	119,215
Remove deals where one tranche has subordination >1	119,215
Remove observations with missing price	74,307
Remove mixed-type asset pools	74,253
Remove PO, IO, IF and FtV	71,950

Table 15: Data cleaning stages with number of tranches outstanding at the end of each step.

S&P rating	Code	Coarse rating	Code
AAA	1	AAA	1
AA+	2	AA	2
AA	3	AA	2
AA-	4	AA	2
A+	5	A	3
A	6	A	3
A-	7	A	3
BBB+	8	BBB	4
BBB	9	BBB	4
BBB-	10	BBB	4
BB+	11	BB	5
BB	12	BB	5
BB-	13	BB	5
B+	14	B	6
B	15	B	6
B-	16	B	6
CCC	17	C	7
CCC-	18	C	7
CC	19	C	7
C	20	C	7
D	21	D	8
NR	-	NR	-

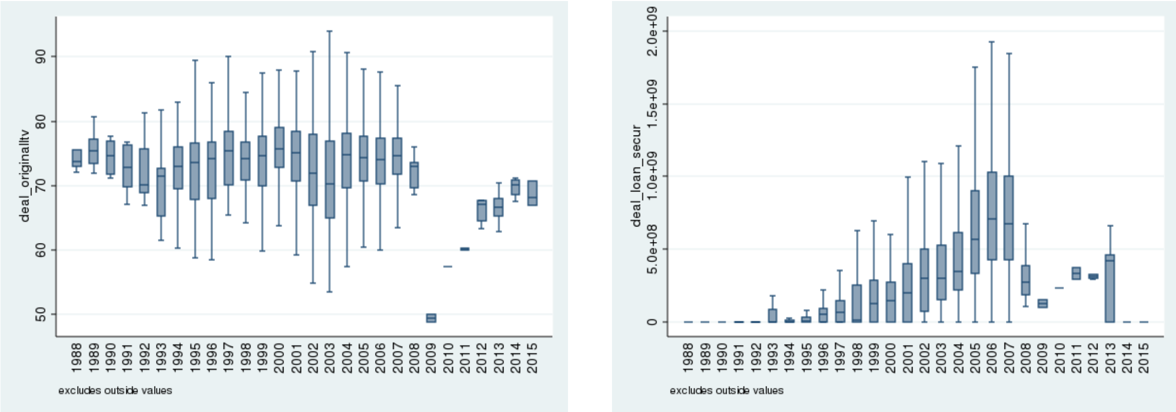
Table 16: Mapping of ratings - fine and coarse level (with numbering code)



(a) FICO score

(b) PMI coverage

Figure B.1: Distribution of covariates over time (vintage year).



(a) LTV

(b) Original securitized balance

Figure B.2: Distribution of covariates over time (vintage year).

At the loan level, we keep only loans having purchase as purpose. This reduces the sample to 8,862,561 loans. Aside minor cleaning (originations before 1980, errors in time stamps) we arrive to 7,145,251. From these we discard asset types other than Alt-A, Prime, Second Lien or Subprime to arrive at the initial sample composed of early and late vintages.

C Variations on the baseline model

C.1 Pricing results with constant default probability and prepayment speed

In this section we use a constant PD, by asset type, given as the

	(1)	(2)	(3)
Asset type	Early vintages	Late vintages	Model PD
Alt-A	7.5%	19.5%	24.5%
Prime	2.3%	6.6%	6.4%
Second Lien	7.2%	25.8%	21.1%
Subprime	14.8%	30.5%	30%
Observations	4,060,698	631,793	2,112

Table 17: Liquidation rates from the loan sample, and PD used for baseline estimation. Column (1) calculates the percentage of loans linked to early vintage deals (before June 2005) that are liquidated. Column (2) calculates the same ratio for late vintage loans. Column (3) shows the PD parameters used for the pricing model, calculated as the average of the deal level liquidation rates for both early and late deals.

After the collapse of private label securitization in 2007, most securitization conduits are insured against default risk by the Government-Sponsored Entities (Fannie Mae and Freddie Mac), making prepayment risk the most significant risk. Schwartz and Torous (1989) and Stanton (1995) measure the value of the prepayment option in default-free securities (guaranteed by the Government-Sponsored Entities). Downing, Stanton, and Wallace (2005) propose a two-factor valuation model that distinguishes the separate, competing risks carried by the default and the prepayment options. Sugimura (2004) develops an intensity model to price RMBS (pass-through) bonds not insured against default risk, and thus exposed to both prepayment and default risk (but credit events in his approach are assumed to be uncorrelated). We seek an accurate measure of prepayment while keeping the focus on default risk, hence the choice of the PSA schedule (see Figure N.8).

In order to chose the PSA factor we look at the class balance. Class balance factor, which measures balance over time relative to the tranche initial balance, reflects both losses and prepayments, thus is an upper bound for prepayments. The results in Figure N.7 suggest that 150% is an appropriate upper bound. Gorton (2009) states that subprime deals were mostly linked to ARMs (see Figure N.11), those being a priori subject to higher prepayment rates.[23] The evolution of class factor over time does not suggest a radically different prepayment rate for subprime deals in our sample. In this section we will apply the PSA schedule, with a factor of 150%, to all tranches within the same deal.

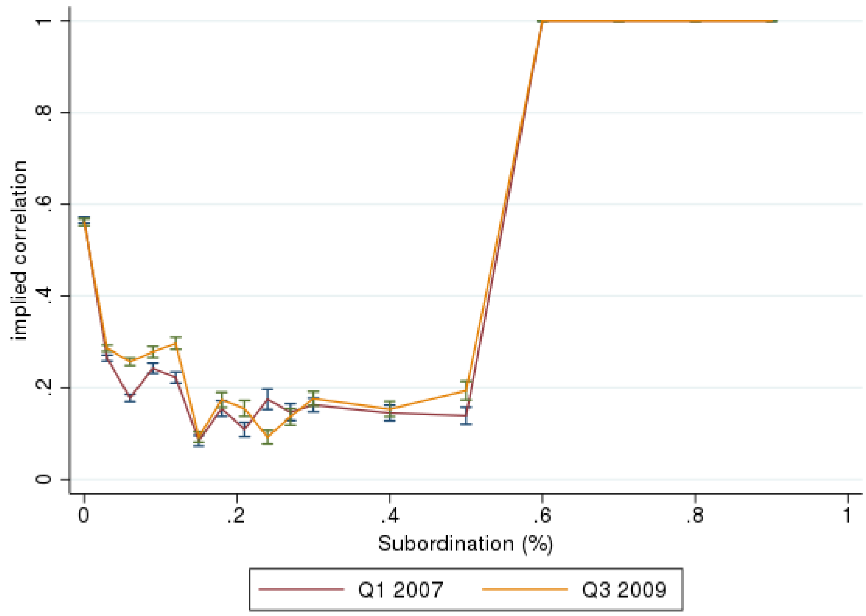


Figure C.1: Average correlation plotted against tranche subordination percentage, on two given dates. We use the sample of early vintage bonds (originated prior to June 2005). Subordinations are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.

Cornaggia, Cornaggia, and Hund (2017) find that ratings are not comparable across broad asset types (corporate, CDO, ABS and RMBS). Within RMBS we emphasize the difference across asset types (prime, subprime and Alt-A), and in this section document a difference in information across asset types, namely between Alt-A and other types. Breaking the change by asset type we see an increase for Alt-A tranches (from 0.36 to 0.40), no change for prime ones (0.30) and a decrease for subprime deals (from 0.59 to 0.49, significant at 99%) so that the upward adjustment over the crisis seems to have mainly affected Alt-A issues.

[23] He finds that the shift to subprime deals happened for the later cohorts. Similarly, we find that later cohorts see faster reductions in balance.

In terms of seniorities, the difference observed by Buzková and Teplỳ (2012) over the crisis is mainly driven by mezzanine tranches (i.e. subordinations between 7%-10% and 10%-15%). Figure K.1 also suggests the increase in correlations is larger among intermediate seniorities, though not as large as the one they observe on the CDX tranches. We now look at average correlation over time (see Figure K.2).

The regression results on price informativeness are similar to those obtained in Section 3.2: implied default correlations are informative when they are linked to well-documented deals, which happens both for AAA and non-AAA tranches. First, the results in Table 40 confirm those of Table 29 in that implied correlations are informative about bond downgrades, except for AAA tranches. Second, the split by opacity index (see Table 42) yields a similar results to that in Table 30. Finally, the further split by rating in Table 42 yields results that are consistent with those in Table 30.

	downgrade		
	(1)	(2)	(3)
	All	AAA only	Non-AAA only
Correlation at first transaction	0.248***	0.0378	0.138**
	(0.0531)	(0.114)	(0.0561)
Observations	29,938	17,234	12,702
Rating at first transaction	Y	N	Y
Vintage year	Y	Y	Y

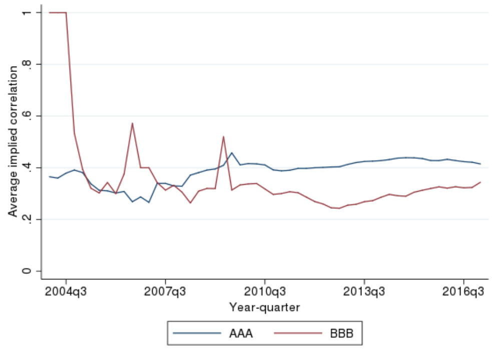
Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Regression results from running logit regression 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Column (1) includes all issues; columns (2) and (3) split the sample between bonds rated AAA at origination and the rest, respectively. Errors are clustered at deal level.

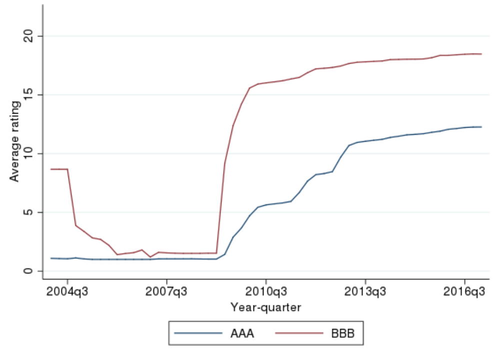
	(1)	(2)	(3)	(4)
	Alt-A	Prime	Second Lien	Subprime
	Downgrade indicator			
Correlation at first transaction	0.198*	0.293**	-0.907	0.266***
	(0.101)	(0.130)	(0.910)	(0.0693)
Observations	8,766	11,862	60	8,620
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

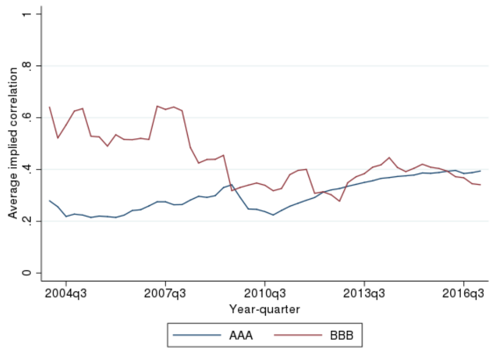
Table 19: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to a given asset type. Errors are clustered at deal level.



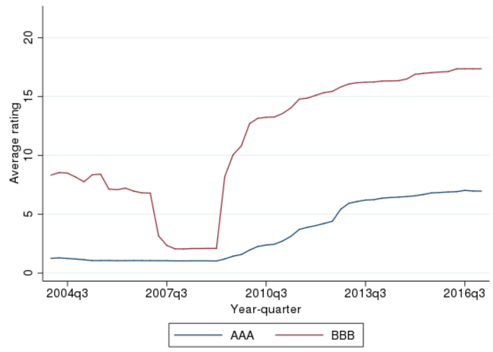
(a) Implied correlation - Alt-A



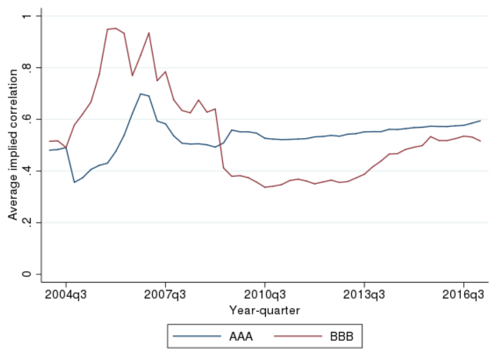
(b) Rating - Alt-A



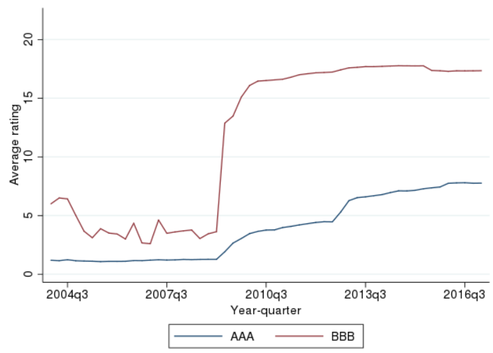
(c) Implied correlation - prime



(d) Rating - prime



(e) Implied correlation - subprime



(f) Rating - subprime

Figure C.2: Performance of early vintage tranches: average implied correlation and average rating for bonds originated before June 2005. For a given we compute the implied correlation, at each point in time. The average is taken by transaction period, by coarse rating at origination (AAA=1,..., BBB=4,..., D=8).

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator			
Correlation at first transaction	-0.142	0.237**	0.357***	0.281**
	(0.203)	(0.105)	(0.0871)	(0.130)
Observations	3,149	7,274	8,824	7,096
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 20: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator - AAA only			
Correlation at first transaction	-0.539	-0.0777	0.378*	0.595**
	(0.356)	(0.168)	(0.210)	(0.297)
Observations	1,760	4,544	4,538	4,369
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y
	Downgrade indicator - not AAA			
Correlation at first transaction	-0.147	0.106	0.221**	0.0974
	(0.271)	(0.126)	(0.0908)	(0.138)
Observations	1,204	2,704	4,222	2,701
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 21: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

C.2 The information content of news in prices

Using the partial observations we recover from the ABSNet data (namely, observations post June 2009) we study the effect of news in prices across the cycle. A number of cleaning stages (see Table 37 in the appendix) are applied to attain the final sample, which contains 6,322,690 panel observations -close to 64 transactions per tranche-. We illustrate the overall numbers in Figure L.1.

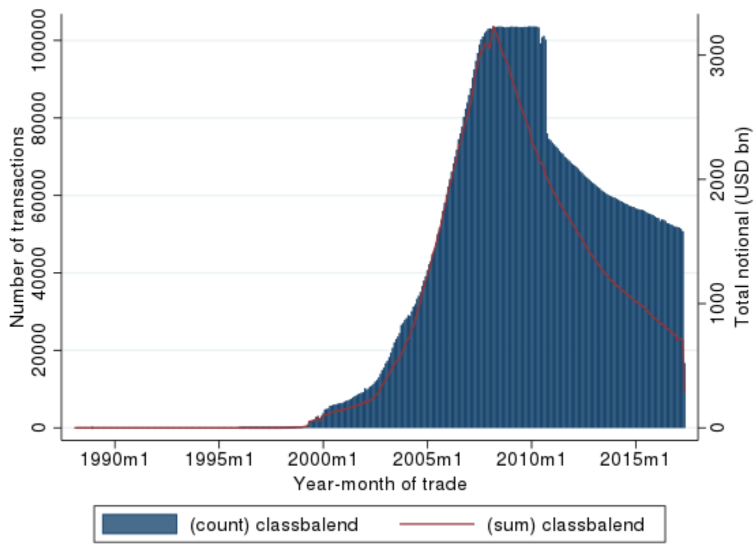


Figure C.3: Tranche balance and number of bonds outstanding by transaction year and month.

The results up to now suggest implied correlation at origination is predictive of downgrades to the extent that the loans have full documentation. Having seen the role of initial signals, our next question is about the role of price news both from rating agencies and the market. While initial ratings rely on an a priori assessment, its evolution over time reflects progressively more of the bond performance, implying that updated rating values should in principle absorb the information that was initially private. We estimate panel 16 using a linear model, with random effects in order to control for tranche-invariants such as first rating and first implied correlation. The advantage of the panel specification 16 is that we can incorporate the partial information coming from the late vintages (after June 2005).

$$outcome_{it} = \alpha_{it} + \beta_0 \rho_{i,0} + \eta_0 rating_{i0} + \beta_1 \rho_{i,t-1} + \eta_1 rating_{i,t-1} + \gamma X_{i,t} + \varepsilon_{it}. \tag{13}$$

In equation 16 $outcome_{it}$ is the month-on-month rating change in notches. Table 44 shows that updates in signals contain information about future bond performance, but the signal is not statistically sufficient for prices. This suggests that investors retain private information over the life of the bond, besides the information given by agency ratings. The second finding is that Alt-A investors do not learn over the life of the bond, so that news in ratings remain statistically sufficient for news in correlation in terms of bond performance.

To see the effect of the crisis on the information content of prices, we will use interactions with an indicator dummy for post-2007 transaction to split estimates between before and after the crisis. The regression specification is the following:

$$\Delta rating_{it} = \alpha_{it} + \beta_0 \rho_{i,0} + \eta_0 rating_{i0} + 1_{post-07} \tag{14}$$

$$+\beta_1 \rho_{i,t-1} \times 1_{post-07} + \eta_1 rating_{i,t-1} \times 1_{post-07} + \gamma X_{i,t} + \varepsilon_{it} \tag{15}$$

Table 45 presents the results of estimating equation (18). It shows that most of the effect of news about default correlation shown in Table 44 comes from the post-crisis period. Griffin and Nickerson (2016) discuss how rating agencies improved their methodologies following the crisis. Under such improvement, the expectation would be that ratings become sufficient for implied correlations, but this is not what we observe. An improvement in rating methodology is consistent with more statistical information coming from prices if ratings are now following the market more closely. In that case changes in implied correlation have more statistical power to predict future downgrades by construction of the downgrade process. The other possibility is that investors learned more from the crisis than the rating agencies, but if this is so it is rational for ratings to follow the market more closely.

	(1)	(2)	(3)	(4)
	Alt-A	Prime	Second Lien	Subprime
	One-month change in rating (notches)			
Lagged correlation (1 month)	0.004	0.007**	-0.025***	-0.007***
	(0.004)	(0.003)	(0.008)	(0.003)
Lagged rating (1 month)	-0.026***	-0.012***	-0.026***	-0.038***
	(0.001)	(0.001)	(0.003)	(0.001)
Correlation at first transaction	0.003	0.027***	-0.004	0.008***
	(0.005)	(0.003)	(0.008)	(0.003)
Rating at first transaction	0.017***	0.017***	0.001	0.010***
	(0.001)	(0.001)	(0.005)	(0.001)
Subordination	0.084***	0.090***	-0.119***	-0.148***
	(0.022)	(0.019)	(0.019)	(0.008)
Observations	2,032,055	1,773,020	55,293	1,452,760
Vintage	Y	Y	Y	Y
Year-quarter	Y	Y	Y	Y

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

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

Table 22: Regression results from running the panel regression 16, by GLS with tranche random effects. The first line gives the coefficient for the change over 1 month (lagged 1 month) of the correlation coefficient, and the second one the coefficient for the change over 1 month (lagged 1 month) of the change in rating (in notches). Errors are clustered at deal level.

In DeMarzo (2005), two factors drive the benefits and drawbacks of securitization: private information by the issuer, on one hand, and asset correlation, on the other. Like CDOs, CMOs are a priori affected by it.[24] Our measure of beliefs about default correlation reflects in part adverse selection concerns on the part of the investors. Because we can't disentangle these two components

[24]Beltran et al. (2017) show that, under asymmetric information, even a modest percentage of bad securities can push security prices far below fundamentals -even to a market meltdown-.

	(1) Alt-A	(2) Prime	(3) Second Lien	(4) Subprime
	Size of downgrade (notches)			
Lagged correlation	-0.001	-0.018***	0.002	0.005
	(0.005)	(0.004)	(0.017)	(0.003)
Lagged correlation × post-07=1	0.005	0.027***	-0.028	-0.013***
	(0.006)	(0.004)	(0.018)	(0.004)
Lagged rating	-0.055***	-0.095***	-0.082***	-0.026***
	(0.016)	(0.009)	(0.014)	(0.008)
Lagged rating × post-07=1	0.029*	0.083***	0.056***	-0.012
	(0.016)	(0.009)	(0.014)	(0.008)
post-07=1	0.016	-0.076***	-0.013	0.130***
	(0.019)	(0.010)	(0.034)	(0.011)
Correlation at first transaction	0.003	0.028***	-0.001	0.008***
	(0.005)	(0.003)	(0.008)	(0.003)
Rating at first transaction	0.017***	0.017***	0.003	0.009***
	(0.001)	(0.001)	(0.005)	(0.001)
Subordination	0.084***	0.089***	-0.120***	-0.148***
	(0.022)	(0.019)	(0.019)	(0.008)
Observations	2,032,055	1,773,020	55,293	1,452,760
Vintage	Y	Y	Y	Y
Year-quarter	Y	Y	Y	Y

Marginal effects; Standard errors in parentheses

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

Table 23: Regression results from running the panel regression 16, by GLS with tranche random effects. The first line gives the coefficient for the change over 1 month (lagged 1 month) of the correlation coefficient, and the second one the coefficient for the change over 1 month (lagged 1 month) of the change in rating (in notches). Errors are clustered at deal level.

empirically, our implied correlation measure is a proxy for market conservatism vis-à-vis information asymmetry. In line with this, Alt-A deals being more reliant on ratings (not only for the first transaction, but also for subsequent ones) potentially reflects a concern for asymmetric information as in Adelino et al. (2016).

D Additional causes of default clustering: frailty and contagion

Following Azizpour et al. (2016), defaults are driven by three factors: systemic risk[25] as captured by macroeconomic variables (Bullard et al., 2009; Khandani et al., 2013)[26], an unobserved frailty factor (Duffie et al., 2009; Kau et al., 2011) and a contagion factor, which captures the extent to which more defaults increase the conditional intensity of default arrival.

A given loan n has a default time T_n. Defaults have a conditional mean of arrival λ given by

$$\lambda_t = \exp\left(a_0 + \sum_{i=1}^{d} a_i X_{i,t}\right) + Y_t + Z_t$$

where X represents a vector of macroeconomic variables. Unobservable frailty Z_t follows the CIR process

$$dZ_t = k(z - Z_t)dt + \sigma\sqrt{Z_t}dW_t$$
$$Z_0 \sim \Gamma\left(\frac{2kz}{\sigma^2}, \frac{\sigma^2}{2k}\right)$$

Defaults are self-exciting, in the sense that the mass of defaults at a given time increases the rate of arrival. This is captured by means of a contagion factor Y such that

$$Y_t = b \sum_{n:T_n \leq t} e^{-\kappa(t-T_n)} U_n$$
$$U_n = \max(0, \log u_n)$$

where u_n is the sum of defaulted debt at time T_n. This implies that larger defaults are followed by more defaults.

The estimation of $\theta = (a, k, z, \sigma, b, \kappa)$ is a filtered likelihood problem (the likelihood is a posterior mean of the complete-data likelihood), and can be solved following Giesecke and Schwenkler (2016).

[25] Bisias et al. (2012) provides a survey of systemic risk measures. See also Chan-Lau et al. (2009). Other approaches include macro measures (costly asset-price boom/bust cycles, property-price, equity-price, credit-gap indicators), forward-looking measures (e.g. absorption rate as in Kritzman, Li, Page, and Rigobon (2010)), cross-sectional measures (CoVaR, Co-Risk, marginal and systemic expected shortfall, see Acharya, Pedersen, Philippon, and Richardson (2012)), stress tests (e.g. Duffie (2011)), illiquidity and insolvency (e.g. Brunnermeier, Gorton, and Krishnamurthy (2011)). Giglio, Kelly, Pruitt, and Qiao (2013) use predictive quantile regression to provide an empirical assessment of 17 of them. Their main finding is that, overall, the compendium of systemic risk measures contains useful predictive information. Instead individual measures tend to fail in capturing systematic risk.

[26] The characterization of systemic risk as deterioration of macroeconomic indicators leaves aside the widely discussed view that the pre-crisis mortgage system was systemically vulnerable (Hellwig, 2009; Poitras and Zanotti, 2016).

The likelihood is written as a product of two terms, one that depends on event data (defaults) and one that depends on factor data. The decomposition is based on a change of measure, which resolves the interaction between the point process and the factors of λ.[27] One of the terms is a point process filter, which makes the computation difficult. Giesecke and Schwenkler (2016) propose an approximation based on a quadrature method, from which the posterior mean can be computed. They write an algorithm and derive conditions for convergence.

E Supplemental regression results

	downgrade		
	(1)	(2)	(3)
	All	AAA only	Non-AAA only
Price	-0.0187***	-0.0457***	-0.00932***
	(0.00151)	(0.00299)	(0.00149)
Coupon	-0.123***	-0.0365	-0.184***
	(0.0178)	(0.0245)	(0.0240)
Subordination	-3.130***	-3.944***	-3.978***
	(0.268)	(0.565)	(0.310)
Observations	26,242	14,034	12,206
Rating at first transaction	Y	N	Y
Vintage year	Y	Y	Y

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 24: Regression results from running logit regression 1 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include price, subordination, coupon and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Column (1) includes all issues; columns (2) and (3) split the sample between bonds rated AAA at origination and the rest, respectively. Errors are clustered at deal level.

F Supplemental graphs and tables

G Data cleaning

H Bond data

We start with 16,397,826 panel observations, corresponding to 127,963 tranches. I remove data entry errors such as subordination percentages larger than one. In those cases all observations for the month (all tranches linked to the deal involved) are removed so as to ensure computations of the

[27]An alternative is to apply the expectation maximization (EM) algorithm. Giesecke and Schwenkler (2016) compare the two approaches.

	(1) [0, 0.25)	(2) [0.25, 0.5)	(3) [0.5, 0.75)	(4) [0.75, 1]
		Downgrade indicator		
Price	-0.0159***	-0.0200***	-0.0110***	-0.0169***
	(0.00606)	(0.00333)	(0.00267)	(0.00354)
Coupon	-0.142**	-0.0380	-0.117***	-0.0780*
	(0.0640)	(0.0304)	(0.0441)	(0.0466)
Subordination	0.00163	-1.857***	-4.016***	-5.722***
	(0.864)	(0.657)	(0.489)	(0.943)
Observations	2,489	5,513	7,073	5,049
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 25: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include price, subordination, coupon and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

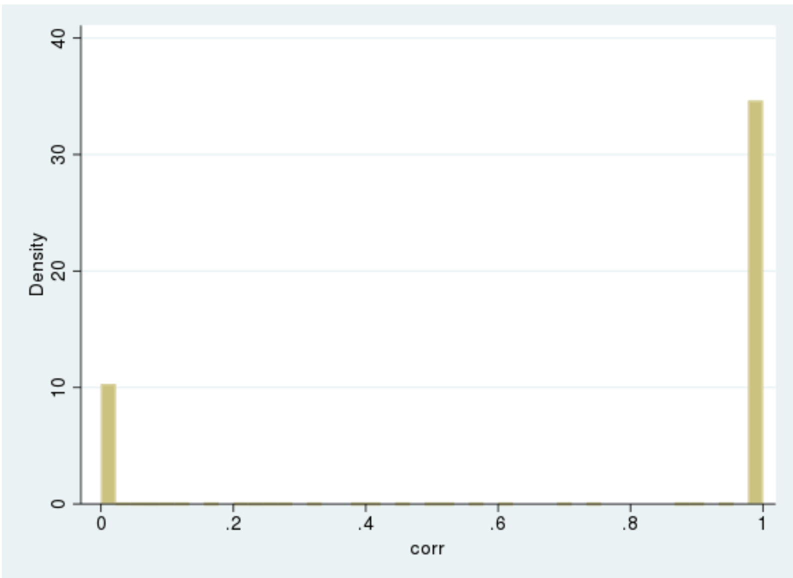


Figure E.1: Histogram plotting all outcomes from the pricing model.

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator - AAA only			
Price	-0.0352***	-0.0360***	-0.0347***	-0.0539***
	(0.00900)	(0.00529)	(0.00632)	(0.0127)
Coupon	0.0508***	0.0546	0.0919	0.118*
	(0.0161)	(0.0451)	(0.0575)	(0.0625)
Subordination	-0.0174	-2.774**	-2.014	-9.907***
	(1.622)	(1.229)	(1.881)	(3.612)
Observations	1,325	3,073	3,272	2,926
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y
	Downgrade indicator - not AAA			
Price	-0.0163**	-0.0129***	-0.00786***	-0.0113***
	(0.00714)	(0.00371)	(0.00250)	(0.00358)
Coupon	-0.367***	-0.167***	-0.201***	-0.156***
	(0.102)	(0.0475)	(0.0529)	(0.0603)
Subordination	-0.309	-2.648***	-4.501***	-4.193***
	(1.881)	(0.880)	(0.538)	(0.784)
Observations	1,038	2,248	3,757	2,111
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 26: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include price, subordination, coupon and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

	with data up to 2004		with data up to 2007	
	(1)	(2)	(3)	(4)
	Default	Prepayment	Default	Prepayment
log(FICO)	-1.468***	1.408***	-2.076***	0.305**
	-0.157	-0.155	-0.199	-0.12
owner occupied	0.039	-0.024	-0.098*	0.024
	-0.05	-0.02	-0.054	-0.02
original r - original 10 year rate	0.475***	0.249***	0.252***	0.066***
	-0.01	-0.017	-0.011	-0.006
log(original amount)	0.421***	0.257***	0.143***	0.02
	-0.043	-0.031	-0.041	-0.026
log(original LTV)	0.439***	-0.007	0.183***	0.069***
	-0.043	-0.036	-0.033	-0.02
prepayment penalty	-1.866***	-1.034***	-0.914***	-0.950***
	-0.08	-0.073	-0.031	-0.025
adjustable rate mortgage	0.655***	0.493***	0.367***	0.467***
	-0.062	-0.047	-0.038	-0.015
log(Cumulative HPA)	-8.398***	-7.780***	-6.482***	-2.474***
	-1.041	-0.963	-0.652	-0.41
coupon gap	0.400***	0.120*	-0.255***	-0.144**
	-0.05	-0.062	-0.04	-0.06
unemployment	0.330***	0.320***	0.201***	0.319***
	-0.072	-0.075	-0.068	-0.075
Asset type: Prime	-1.008***	-0.147***	-1.130***	-0.603***
	-0.078	-0.027	-0.078	-0.033
Asset type: Second Lien	-0.580***	0.124	0.843***	0.385***
	-0.142	-0.079	-0.064	-0.028
Asset type: Subprime	0.504***	-0.021	1.113***	0.201***
	-0.053	-0.05	-0.037	-0.02
CBSA FE	Y	Y	Y	Y
Month since origination FE	Y	Y	Y	Y
Observations	68,634,789	76,206,672	121,236,208	126,625,633

Standard errors in parentheses

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

Table 27: This table shows estimates using the maximum likelihood estimation of the complementary log-log specification in (11), using a nonparametric baseline hazard, on the loan level data available from ABSNet for private label loans (purchases only). The model treats competing risks independently, indicating 1 for failure and 0 for censoring. Each coefficient is the effect of the corresponding variable on the log hazard rate for either the default or prepayment of a mortgage. The conditional hazard is captured by performance month dummies, where performance is tracked over the first 60 months of the sample. The sample is truncated at December 2004 for columns (1) and (2), and at June 2007 for columns (3) and (4). Errors are clustered at CBSA level.

	(1) default	(2) prepayment
log(FICO)	-2.481***	0.448***
	-0.064	-0.018
owner occupied	0.025*	0.372***
	-0.014	-0.005
original r - original 10 year rate	0.429***	-0.011***
	-0.004	-0.001
log(original amount)	0.137***	0.324***
	-0.01	-0.003
log(original LTV)	0.572***	0.183***
	-0.012	-0.005
adjustable rate mortgage	0.487***	0.579***
	-0.016	-0.004
log(Cumulative HPA)	-1.826***	-1.581***
	-0.051	-0.011
coupon gap	0.848***	-0.261***
	-0.007	-0.002
unemployment	0.080***	0.001
	-0.004	-0.001
Asset type: Prime	-0.808***	-2.719***
	-0.044	-0.014
Asset type: Second Lien	-0.794***	0.298***
	-0.038	-0.011
Asset type: Subprime	0.402***	1.079***
	-0.025	-0.005
CBSA FE	N	N
Month since origination FE	N	Y
Observations	2,630,290	76,374,400

Standard errors in parentheses

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

Table 28: This table shows estimates using the maximum likelihood estimation of a complementary log-log specification, using a hazard specification for prepayments and an dummy indicator for default, on the loan level data available from ABSNet for private label loans (purchases only). The hazard model treats default risk as censored. Each coefficient is the effect of the corresponding variable on the log hazard rate for prepayment or the log probability of default of a mortgage. The conditional hazard is captured by performance month dummies, where performance is tracked over the first 60 months of the sample. The sample is truncated at December 2004.

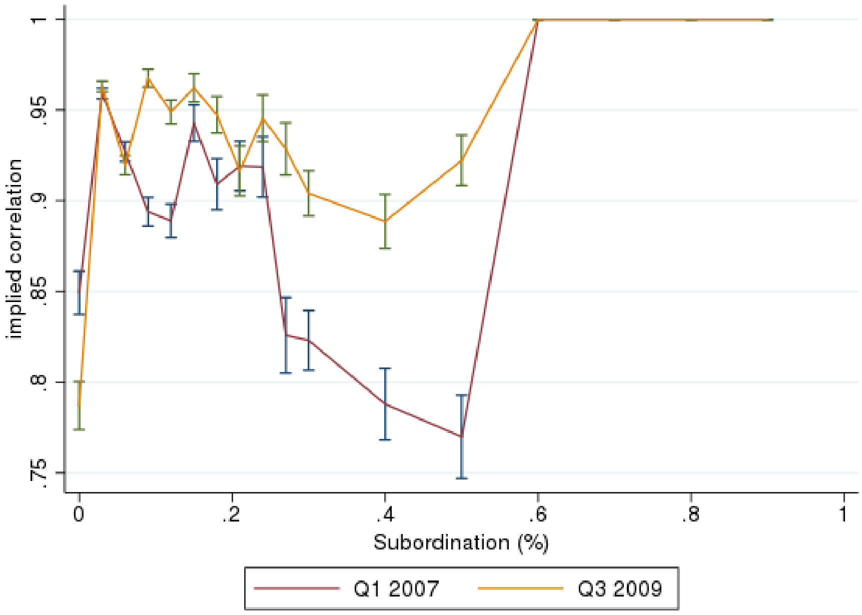


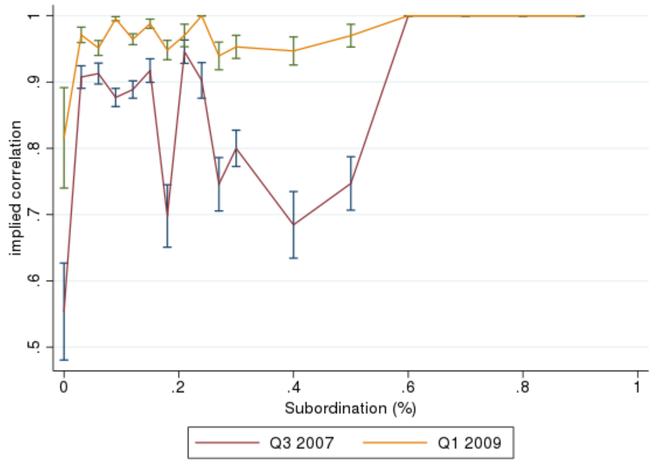
Figure E.2: Average correlation plotted against tranche subordination percentage, on two given dates. We use the sample of early vintage bonds (originated prior to June 2005). Subordinations are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.

	downgrade		
	(1)	(2)	(3)
	All	AAA only	Non-AAA only
Correlation at first transaction	0.414***	0.299	0.268***
	(0.0629)	(0.201)	(0.0644)
Observations	28,991	16,618	12,371
Rating at first transaction	Y	N	Y
Vintage year	Y	Y	Y
Model-implied PD	Y	Y	Y
Asset type	Y	Y	Y

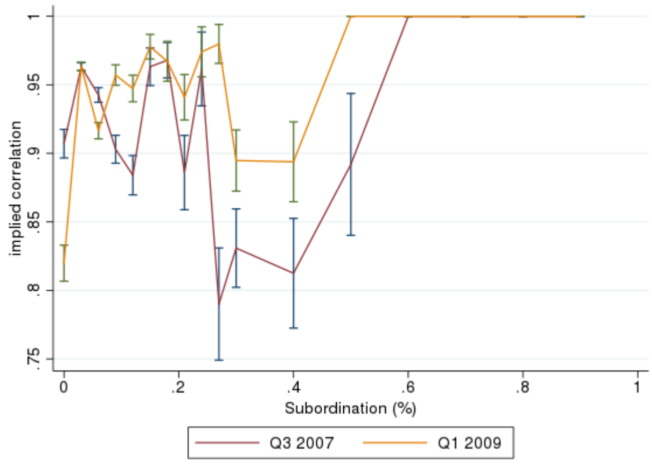
Standard errors in parentheses

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

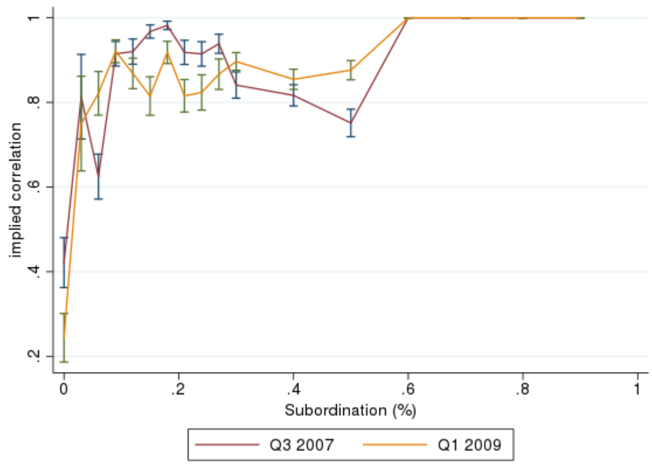
Table 29: Regression results from running logit regression 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default, as estimated in section 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Column (1) includes all issues; columns (2) and (3) split the sample between bonds rated AAA at origination and the rest, respectively. Errors are clustered at deal level.



(a) Alt-A



(b) Prime



(c) Subprime

Figure E.3: Average correlation plotted against tranche subordination percentage, on two given dates. Subordination values are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator			
Correlation at first transaction	0.243	0.605***	0.476***	0.569***
	(0.250)	(0.200)	(0.102)	(0.135)
Observations	2,723	6,285	7,808	5,565
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Model-implied PD	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 30: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default as estimated in section 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the value of the documentation index corresponding to the given deal. Errors are clustered at deal level.

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator - AAA only			
Correlation at first transaction	1.018	0.430	1.647***	0.842***
	(0.703)	(0.599)	(0.627)	(0.321)
Observations	1,529	3,765	3,975	3,429
Rating at first transaction	N	N	N	N
Vintage year	Y	Y	Y	Y
Model-implied PD	Y	Y	Y	Y
Asset type	Y	Y	Y	Y
	Downgrade indicator - not AAA			
Correlation at first transaction	0.0485	0.370**	0.314***	0.353**
	(0.283)	(0.155)	(0.109)	(0.158)
Observations	1,045	2,289	3,787	2,124
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Model-implied PD	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 31: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year (vintages up to June 2005) and model-implied probability of default as estimated in section 3.1. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the value of the documentation index corresponding to the given deal. Errors are clustered at deal level.

	(1)	(2)	(3)	(4)
	AAA balance at origination as share of deal issuance			
Opacity index	-0.104***	-0.0835***	-0.101***	-0.0259*
	(0.0154)	(0.0153)	(0.0151)	(0.0151)
Observations	1,902	1,902	1,902	1,902
Model-implied PD	N	Y	Y	Y
Vintage year	N	N	Y	Y
Asset type	N	N	N	Y

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 32: Regression results from running a linear regression at deal level of AAA origination (as share of total) on the deal opacity index. Controls include model-implied PD, vintage year (we include vintages up to June 2005) and asset type.

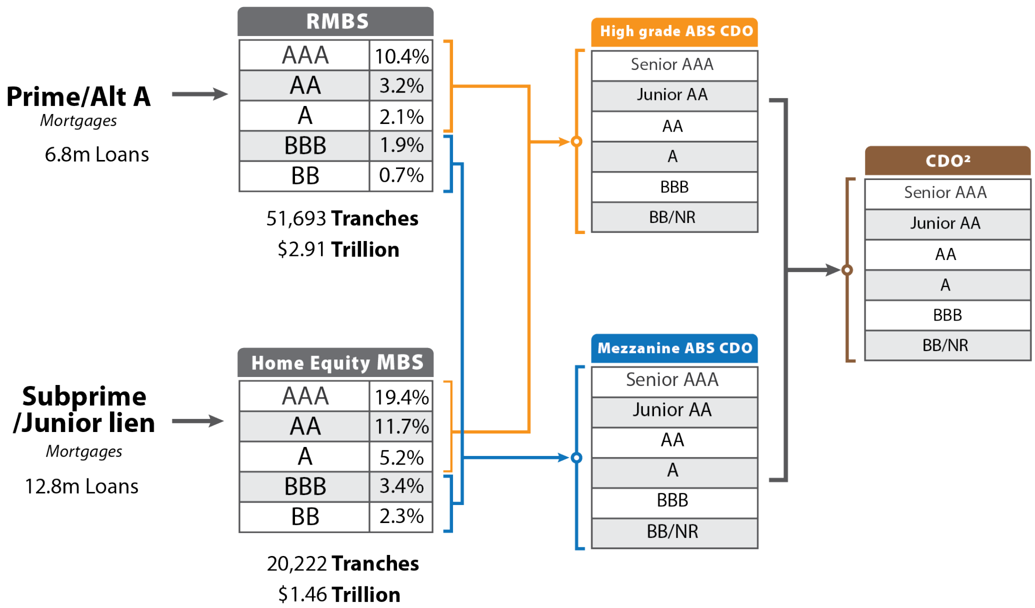
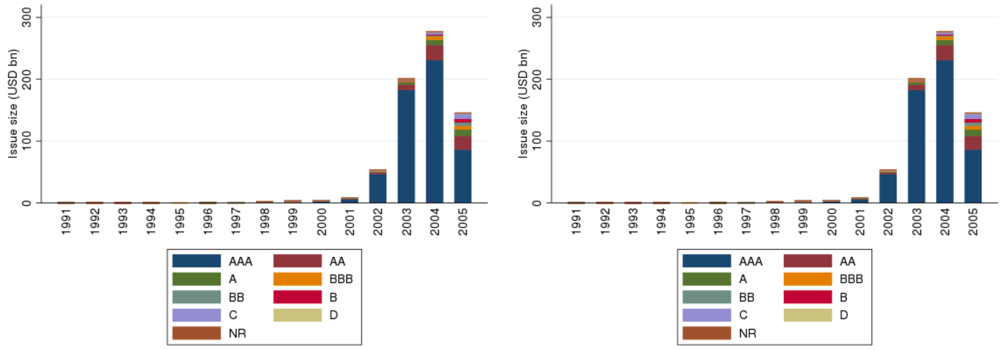


Figure F.1: Diagram: from loans to RMBS CMO, from CMO to CDO, from CDO to CDO^2. Details are reported on the total number of loans recorded by ABSNet, the universe of securities issued and the average subordination percentage by Standard & Poor's rating, as explained in Section 1

Asset type	After Jun-05		Before Jun-05	
	Origination ($bn)	Count	Origination ($bn)	Count
Alt-A	1,179.0	16,837	557.7	11,000
Prime	621.7	9,097	557.9	14,759
Second Lien	64.7	478	19.0	408
Subprime	660.0	9,811	720.2	9,525
Total	2,525.4	36,223	1,854.8	35,692

Table 33: Issued amounts and counts by asset type.



(a) Number of tranches

(b) Total issue

Figure F.2: Number of tranches and amount issued by vintage year for private label collateralized mortgage obligations. Source: ABSNet bond data. For our sample of early vintages (prior to June 2005) we provide the distribution by (coarse, see Table 38) initial rating.

	Our sample		Cordell et al. (2012)	
rating	Prime/Alt-A	Second Lien/Subprime	Prime/Alt-A	Second Lien/Subprime
AAA	10.8%	25.7%	6%	23%
AA	3.4%	14.3%	3%	13%
A	3.0%	5.9%	2%	8%
BBB	2.9%	4.0%	1%	4%

Table 34: Subordination percentage by tranche rating - comparison. The figures computed using ABSNet data are derived by aggregating the subordination percentages at orgination as given in Table 1.4. Our sample contains only early vintages (prior to June 2005) while Cordell et al. (2012) use late vintages as well.

Year	ABSNet sample		Adelino (2009)	
	Origination ($bn)	Count	Origination ($bn)	Count
≤2002	319.3	5,438		
2003	470.5	10,120	496.5	8,574
2004	677.4	12,519	767.3	11,460
2005	904.5	16,684	1,058.5	17,135
2006	1,038.0	15,022	1,080.4	18,206
2007	939.4	11,716	802.1	12,037
≥2008	31.2	177		
Total	4,380.3	71,676	4,204.8	67,412

Table 35: Origination amounts and counts at origination, by vintage year, compared to the sample in Adelino (2009).

	(1)	(2)
Asset type	Early vintages	Late vintages
Alt-A	7.5%	19.5%
Prime	2.3%	6.6%
Second Lien	7.2%	25.8%
Subprime	14.8%	30.5%
Observations	4,060,698	631,793

Table 36: Liquidation rates from the loan sample. Column (1) calculates the percentage of loans linked to early vintage deals (before June 2005) that are liquidated. Column (2) calculates the same ratio for late vintage loans.

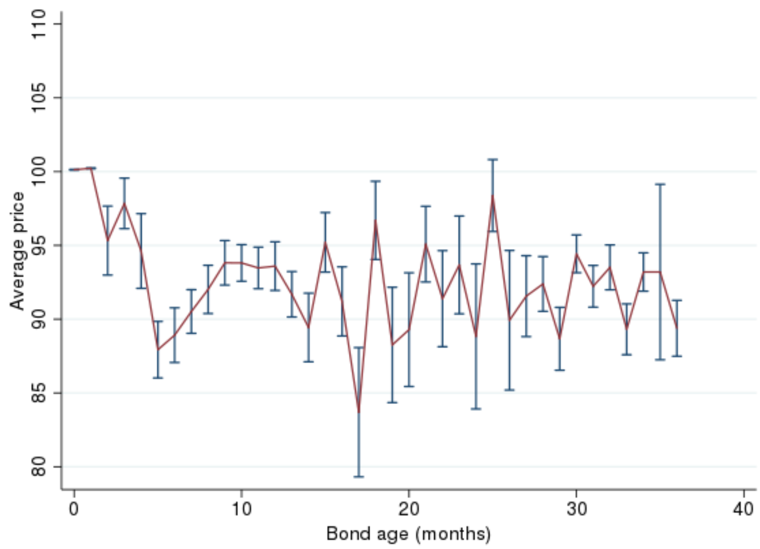


Figure F.3: Average tranche price by age of the bond in months. For our sample of bonds originated in 2004 and 2005 we compute the average price by the time elapsed (in months) since the bond issue. Vertical whiskers show the standard errors.

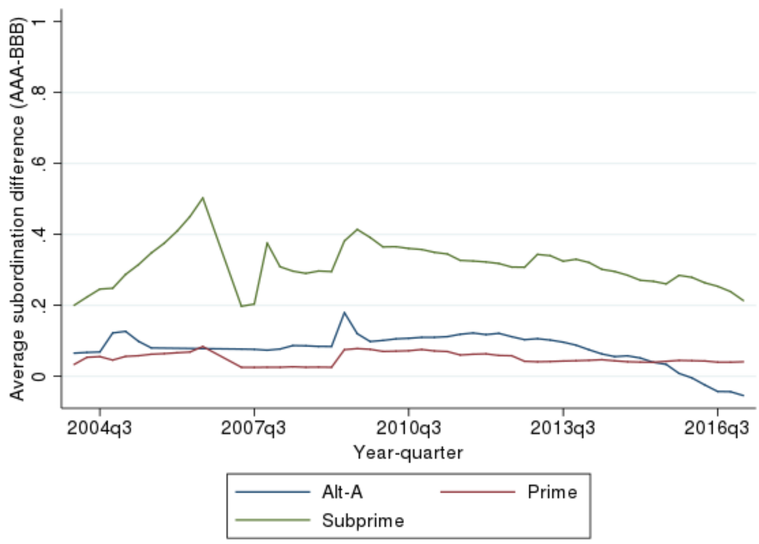


Figure F.4: Average subordination difference between AAA and BBB bonds. Source: ABSNet bond data.The figure presents the difference between the average AAA and average BBB subordination over trading time (for early vintages, prior to June 2005) using the rating at the given trading time. The difference is computed by asset type.

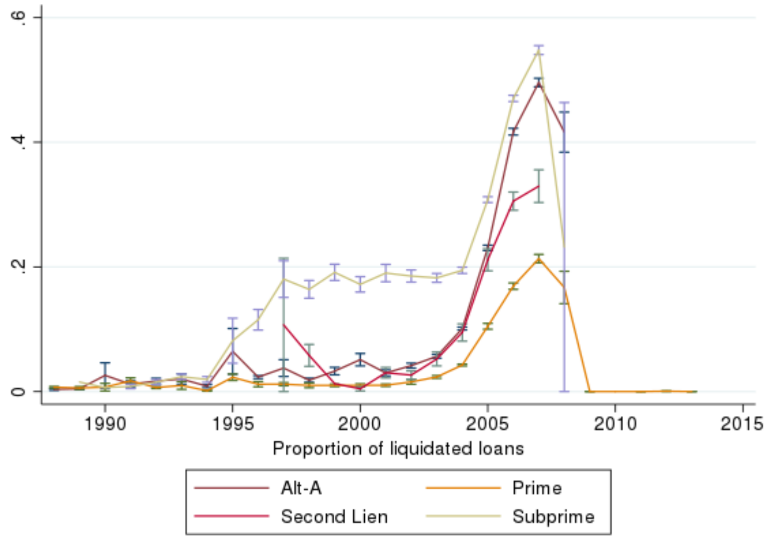


Figure F.5: Probability of default by vintage year. We compute the default rate for each of the deals that compose our population, and then average by vintage year and asset type. The results are presented here along with standard error bands around the average.

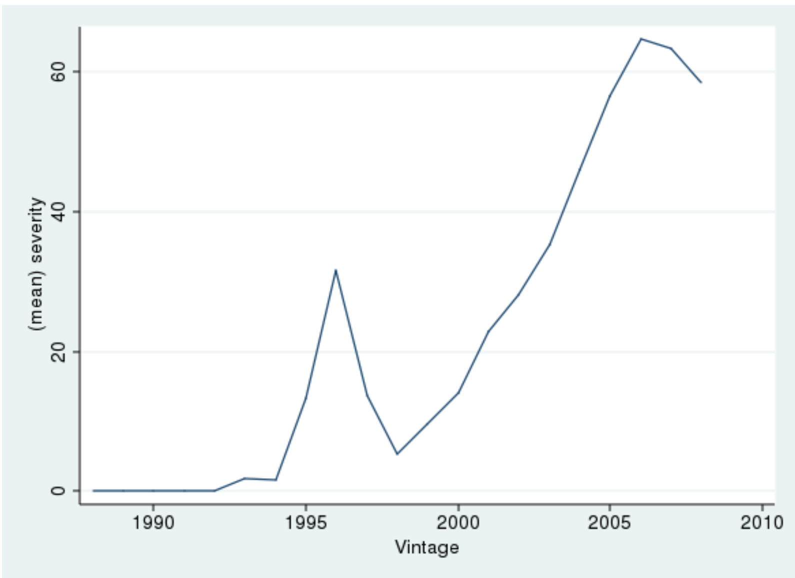


Figure F.6: Percentage loss given default by vintage year. The aggregate loss given default is computed from the sample of loans associated to the deals that compose our population of CMOs.

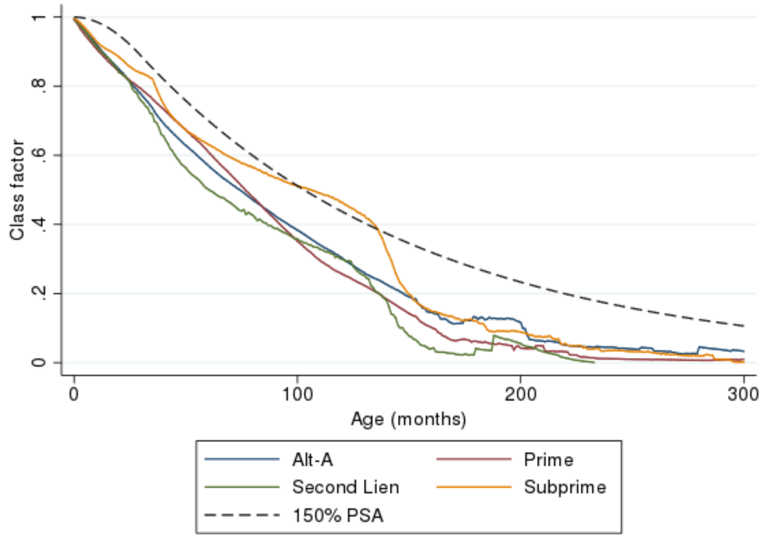


Figure F.7: Average class balance factor by asset class over tranche age. Alongside the averages, we compute the balance factor that results from a 150% payment schedule alone (excluding planned amortization).

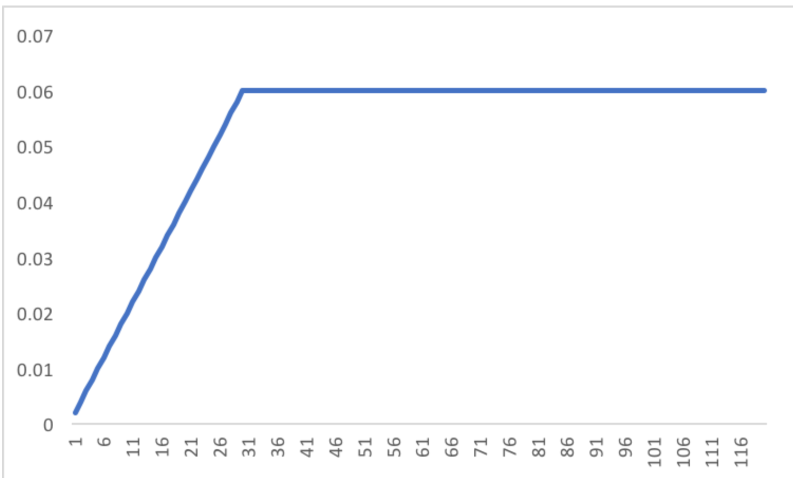


Figure F.8: Standard Prepayment Model of The Bond Market Association. Prepayment percentage for each month in the life of the underlying mortgages, expressed on an annualized basis.

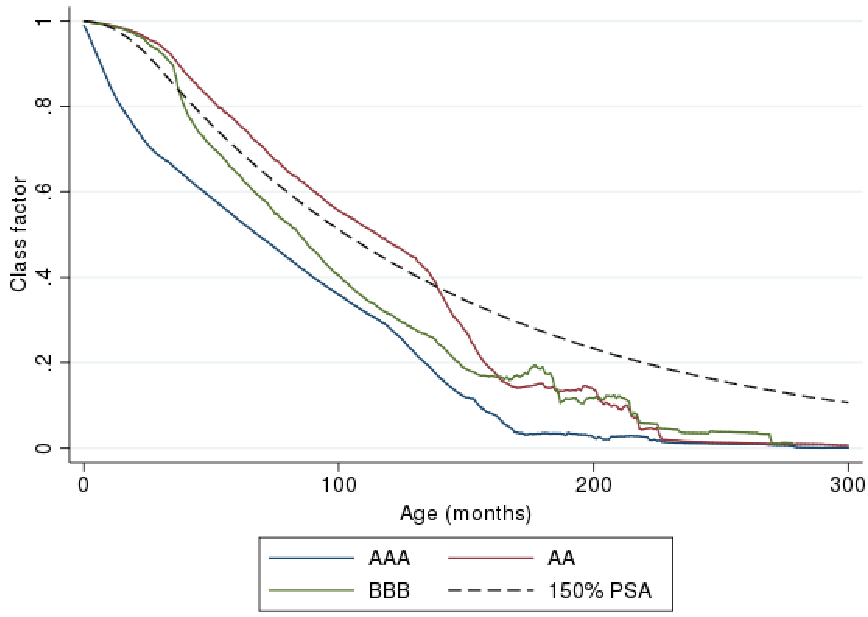


Figure F.9: Plot of average class factor against tranche age by tranche initial rating.

tranching structure are correct.[28] We follow Adelino (2009) in removing Interest Only, Principal Only, Inverse Floater and Fixed to Variable bonds from the sample.

Stage	Tranches left
Remove deals that are entirely made of mixed asset types	119,215
Remove deals where one tranche has subordination >1	119,215
Remove observations with missing price	74,307
Remove mixed-type asset pools	74,253
Remove PO, IO, IF and FtV	71,950

Table 37: Data cleaning stages with number of tranches outstanding at the end of each step.

Notice that the most aggressive cleaning step is the removal of observations where price is missing. As dicussed in section 1, this is due to the data gap that covers late (2005 and more recent) vintages.

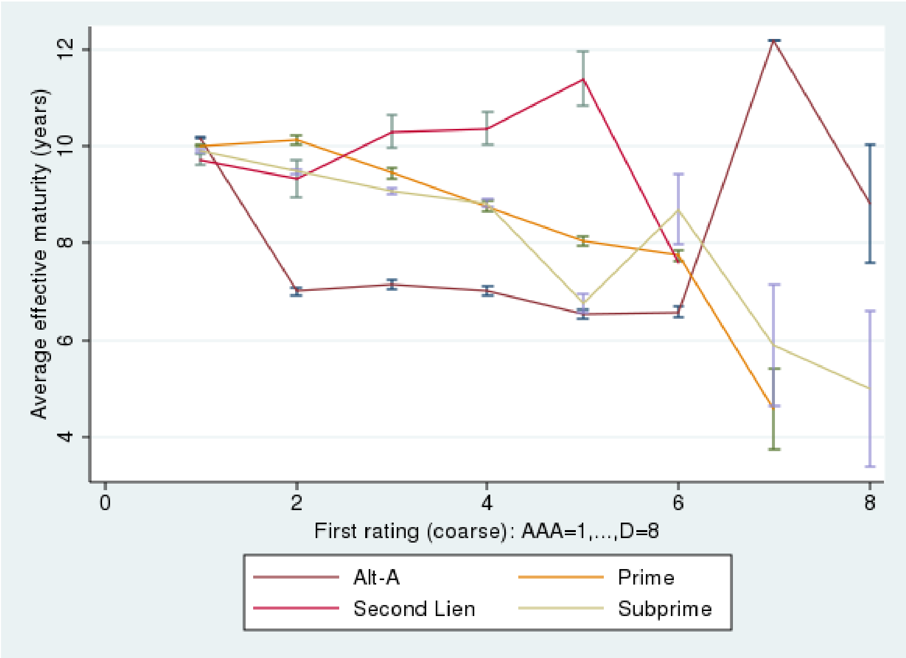
I Loan level data

We start with a set of 22,008,610 loan originations. Of our originations set, 21,759,836 map to one of our deal IDs. Below is a summary of deal level averages of certain covariates (FICO score, LTV, private mortgage insurance coverage percentage) are computed.[29]

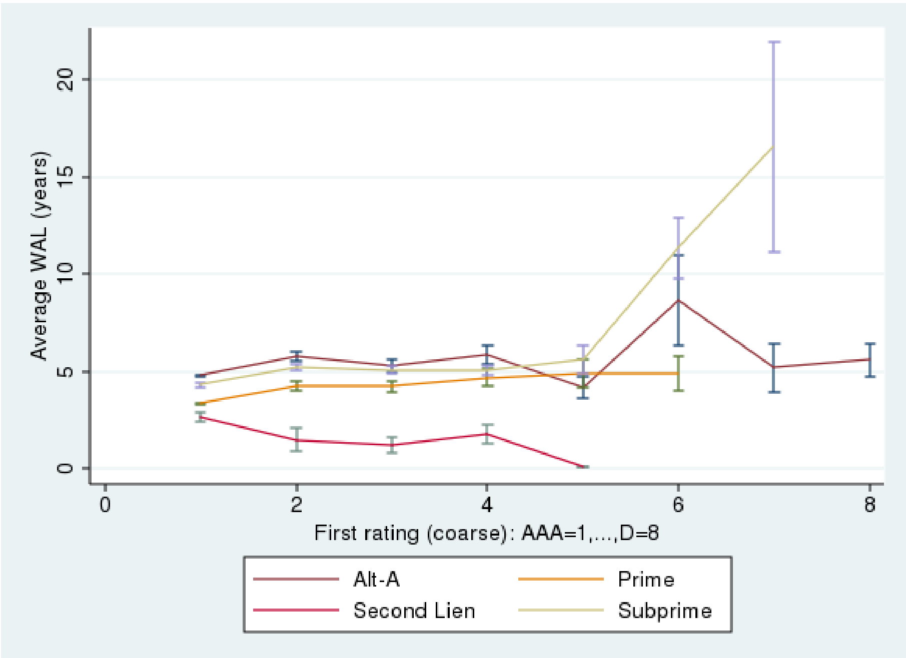
Historic data are contained in monthly reports. From the input 21,996,382 facilities we have at least one observation for 17,350,072 of them. We recover a total 792,664,139 loan-month observations

[28] I manually computed subordination percentages on a random sample of deals to check the calculations by ABSNet.

[29] Simple averages were preferred over weighted averages (weighted by e.g. the initial securitized balance) as this reduces the number of missing observations.



(a) Average realized



(b) WAL

Figure F.10: Average realized and weighted average life by coarse rating and asset type. The second panel includes observations where we found a matching WAL in Bloomberg.

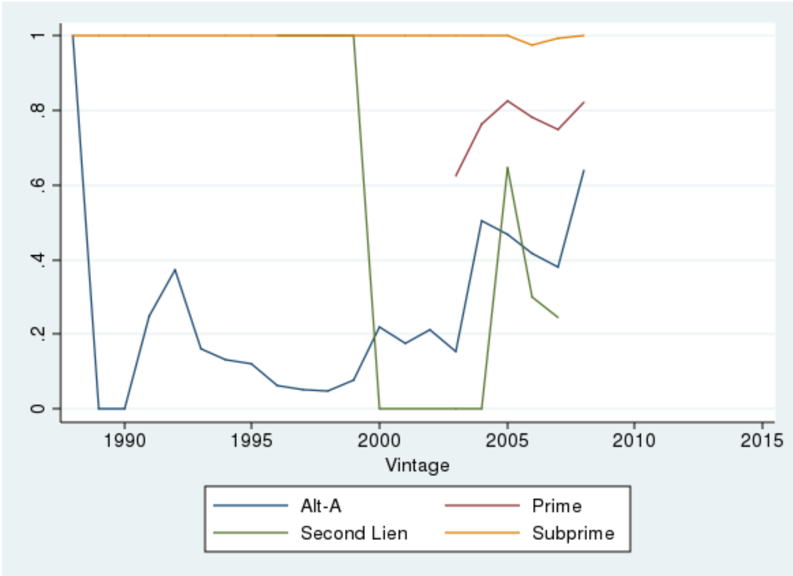


Figure F.11: Proportion of ARM loans by vintage and asset type.

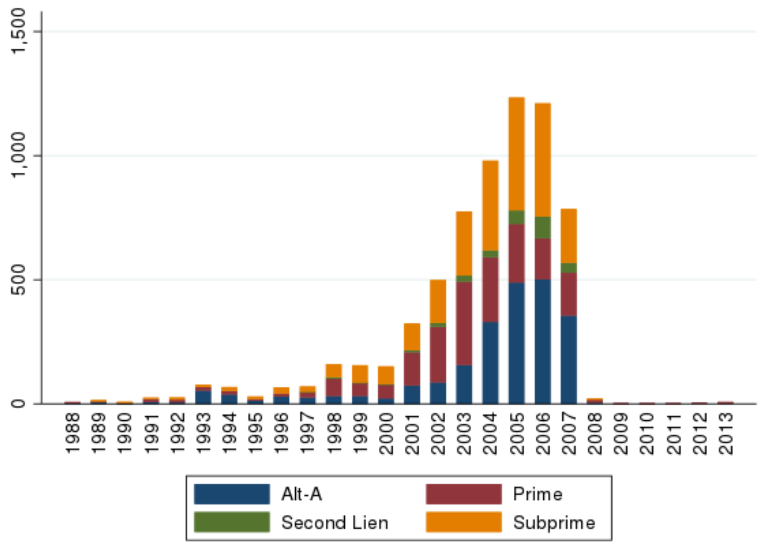


Figure F.12: Number of deals originated by asset type and vintage year.

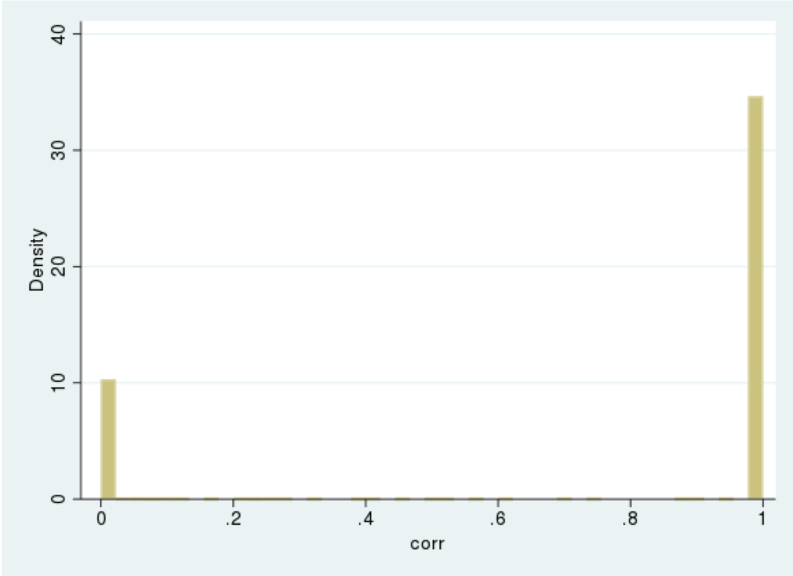


Figure F.13: Histogram plotting all outcomes from the pricing model.

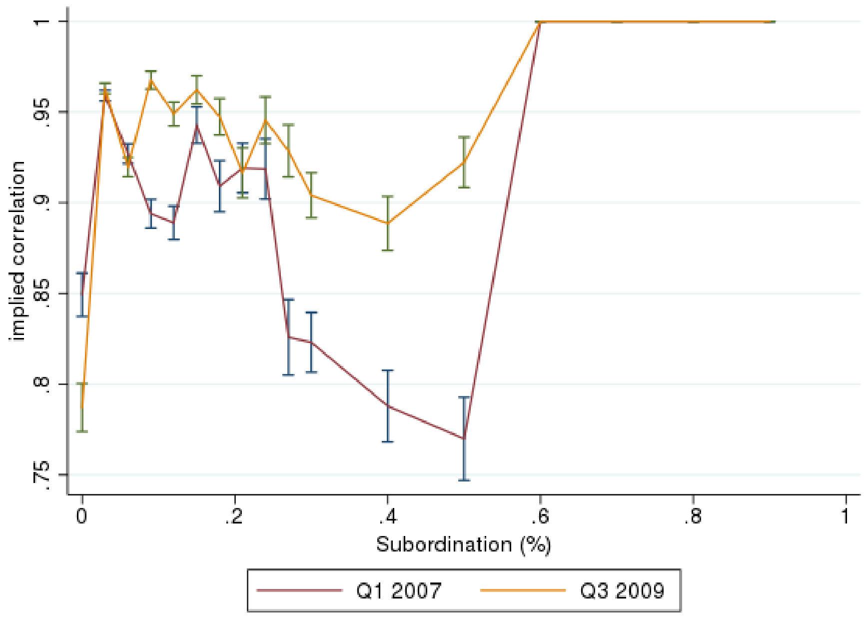
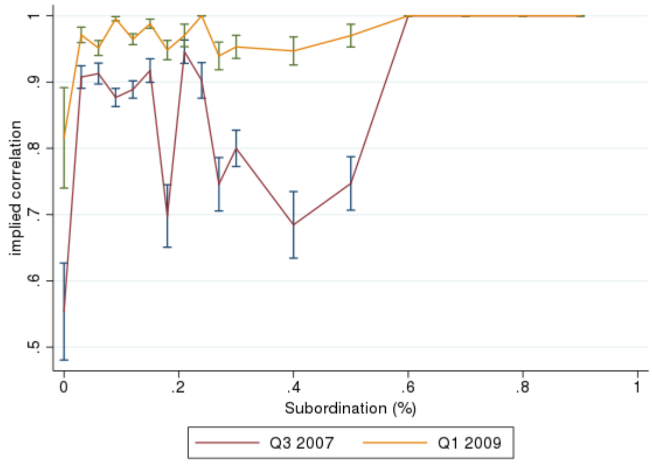
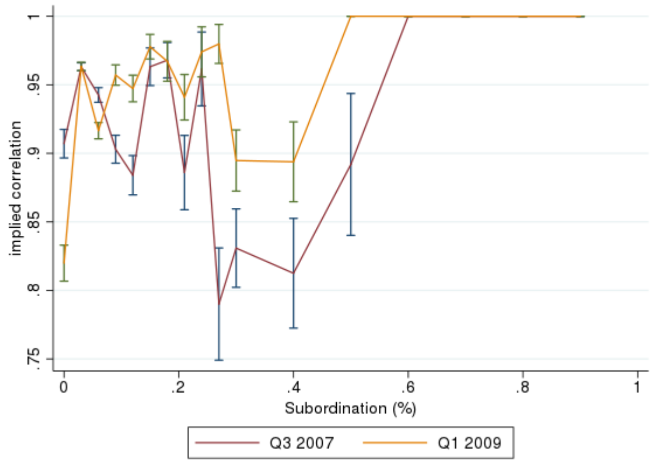


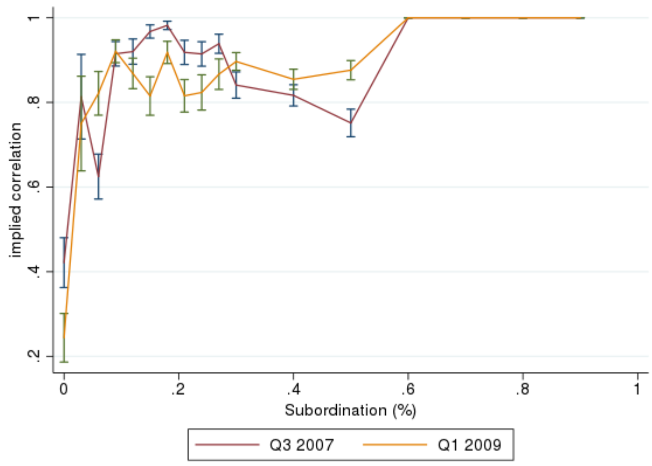
Figure F.14: Average correlation plotted against tranche subordination percentage, on two given dates. We use the sample of early vintage bonds (originated prior to June 2005). Subordinations are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.



(a) Alt-A



(b) Prime

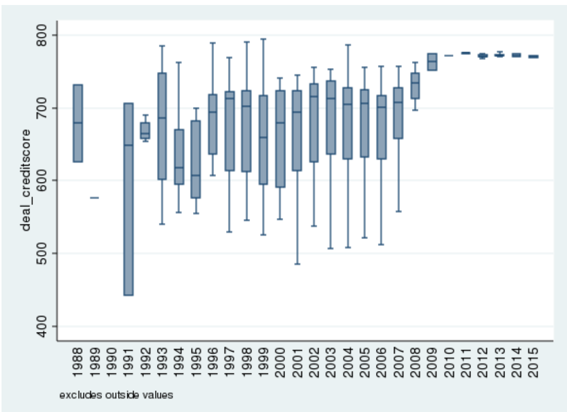


(c) Subprime

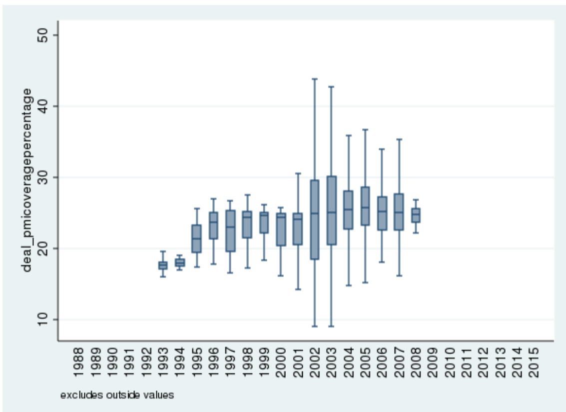
Figure F.15: Average correlation plotted against tranche subordination percentage, on two given dates. Subordination values are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.

S&P rating	Code	Coarse rating	Code
AAA	1	AAA	1
AA+	2	AA	2
AA	3	AA	2
AA-	4	AA	2
A+	5	A	3
A	6	A	3
A-	7	A	3
BBB+	8	BBB	4
BBB	9	BBB	4
BBB-	10	BBB	4
BB+	11	BB	5
BB	12	BB	5
BB-	13	BB	5
B+	14	B	6
B	15	B	6
B-	16	B	6
CCC	17	C	7
CCC-	18	C	7
CC	19	C	7
C	20	C	7
D	21	D	8
NR	-	NR	-

Table 38: Mapping of ratings - fine and coarse level (with numbering code)

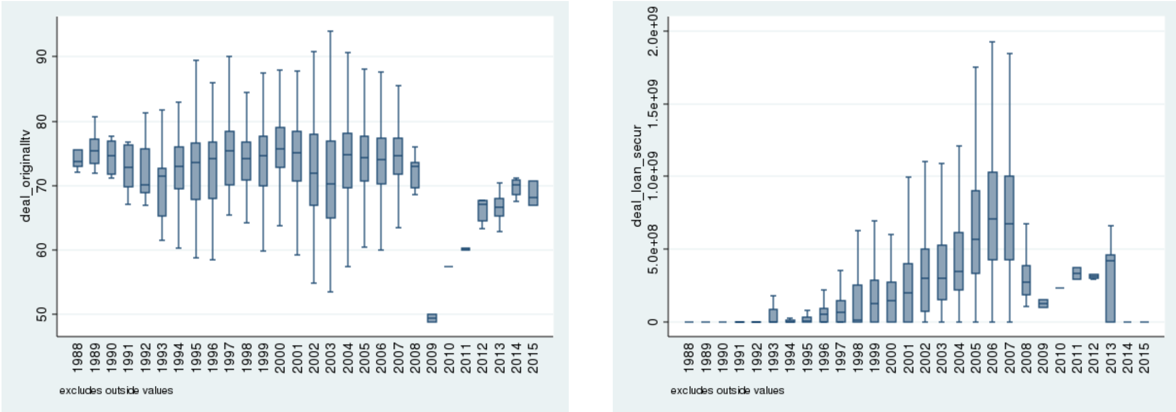


(a) FICO score



(b) PMI coverage

Figure I.1: Distribution of covariates over time (vintage year).



(a) LTV

(b) Original securitized balance

Figure I.2: Distribution of covariates over time (vintage year).

from payment history (on average 45.7 obs per loan). From there we can compute default rates at deal level. We have loss event data for 3,986,974 observations, linked to 5,965 deal IDs. From there we can compute LGDs at deal level or vintage level.

At the loan level, we keep only loans having purchase as purpose. This reduces the sample to 8,862,561 loans. Aside minor cleaning (originations before 1980, errors in time stamps) we arrive to 7,145,251. From these we discard asset types other than Alt-A, Prime, Second Lien or Subprime to arrive at the initial sample composed of early and late vintages.

J Variations on the baseline model

K Pricing results with constant default probability and prepayment speed

In this section we use a constant PD, by asset type, given as the

	(1)	(2)	(3)
Asset type	Early vintages	Late vintages	Model PD
Alt-A	7.5%	19.5%	24.5%
Prime	2.3%	6.6%	6.4%
Second Lien	7.2%	25.8%	21.1%
Subprime	14.8%	30.5%	30%
Observations	4,060,698	631,793	2,112

Table 39: Liquidation rates from the loan sample, and PD used for baseline estimation. Column (1) calculates the percentage of loans linked to early vintage deals (before June 2005) that are liquidated. Column (2) calculates the same ratio for late vintage loans. Column (3) shows the PD parameters used for the pricing model, calculated as the average of the deal level liquidation rates for both early and late deals.

After the collapse of private label securitization in 2007, most securitization conduits are insured against default risk by the Government-Sponsored Entities (Fannie Mae and Freddie Mac), making prepayment risk the most significant one in the literature. Schwartz and Torous (1989) and Stanton (1995) measure the value of prepayment option in default-free securities (guaranteed by the Government-Sponsored Entities). Downing et al. (2005) propose a two-factor valuation model that distinguishes the separate, competing risks carried by the default and the prepayment options. Sugimura (2004) develops an intensity model to price RMBS (pass-through) bonds not insured against default risk, and thus exposed to both prepayment and default risk (but credit events in his approach are assumed to be uncorrelated). We seek an accurate measure of prepayment while keeping the focus on default risk, hence the choice of the PSA schedule (see Figure N.8).

In order to chose the PSA factor we look at the class balance. Class balance factor, which measures balance over time relative to the tranche initial balance, reflects both losses and prepayments, thus is an upper bound for prepayments. The results in Figure N.7 suggest that 150% is an appropriate upper bound. Gorton (2009) states that subprime deals were mostly linked to ARMs (see Figure N.11), those being a priori subject to higher prepayment rates.[30] The evolution of class factor over time does not suggest a radically different prepayment rate for subprime deals in our sample. In this section we will apply the PSA schedule, with a factor of 150%, to all tranches within the same deal.

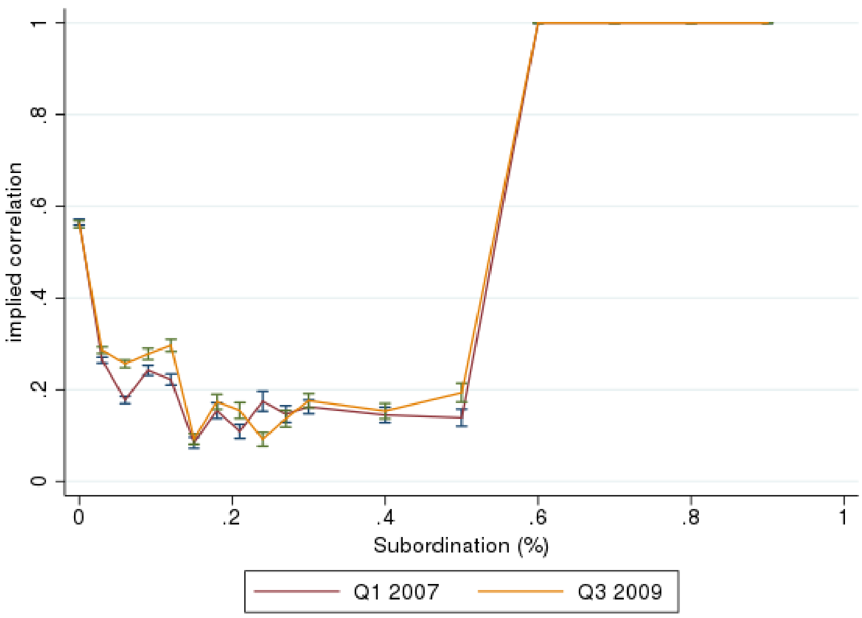


Figure K.1: Average correlation plotted against tranche subordination percentage, on two given dates. We use the sample of early vintage bonds (originated prior to June 2005). Subordinations are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.

Cornaggia et al. (2017) find that ratings are not comparable across broad asset types (corporate, CDO, ABS and RMBS). Within RMBS we emphasize the difference across asset types (prime,

[30]He finds that the shift to subprime deals happened for the later cohorts. Similarly, we find that later cohorts see faster reductions in balance.

subprime and Alt-A), and in this section document a difference in information across asset types, namely between Alt-A and other types.

Breaking the change by asset type we see an increase for Alt-A tranches (from 0.36 to 0.40), no change for prime ones (0.30) and a decrease for subprime deals (from 0.59 to 0.49, significant at 99%) so that the upward adjustment during seems to have mainly affected Alt-A issues.

In terms of seniorities, the difference observed by Buzková and Teplý (2012) over the crisis is mainly driven by mezzanine tranches (7%-10% and 10%-15%). Figure K.1 also suggests the increase in correlations is larger among intermediate seniorities, though not as large as the one they observe on the CDX tranches. We now look at average correlation over time (see Figure K.2).

The regression results on price informativeness are similar to those obtained in Section 3.2: implied default correlations are informative when they are linked to well-documented deals, which happens both for AAA and non-AAA tranches. First, the results in Table 40 confirm those of Table 29 in that implied correlations are informative about bond downgrades, except for AAA tranches. Second, the split by opacity index (see Table 42) yields a similar results to that in Table 30. Finally, the further split by rating in Table 42 yields results that are consistent with those in Table 30.

	downgrade		
	(1)	(2)	(3)
	All	AAA only	Non-AAA only
Correlation at first transaction	0.248^{***}	0.0378	0.138^{**}
	(0.0531)	(0.114)	(0.0561)
Observations	29,938	17,234	12,702
Rating at first transaction	Y	N	Y
Vintage year	Y	Y	Y

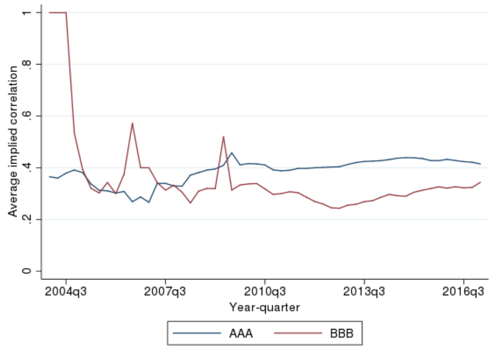
Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 40: Regression results from running logit regression 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Column (1) includes all issues; columns (2) and (3) split the sample between bonds rated AAA at origination and the rest, respectively. Errors are clustered at deal level.

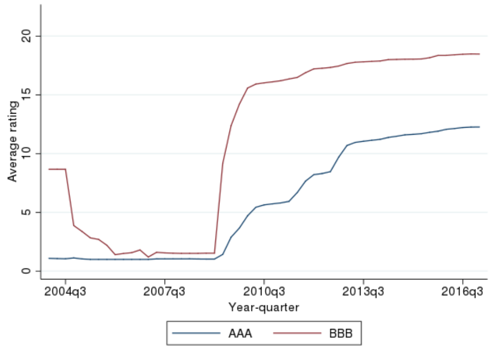
L The information content of news in prices

Using the partial observations we recover from the ABSNet data (namely, observations post June 2009) we study the effect of news in prices across the cycle. A number of cleaning stages (see Table 37 in the appendix) are applied to attain the final sample, which contains 6,322,690 panel observations -close to 64 transactions per tranche-. We illustrate the overall numbers in Figure L.1.

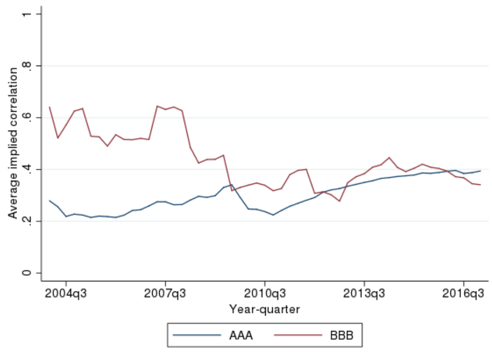
The results up to now suggest implied correlation at origination is predictive of downgrades to the extent that the loans have full documentation. Having seen the role of initial signals, our next question is about the role of price news both from rating agencies and the market. While initial



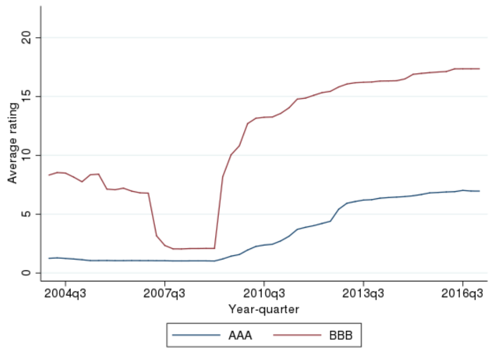
(a) Implied correlation - Alt-A



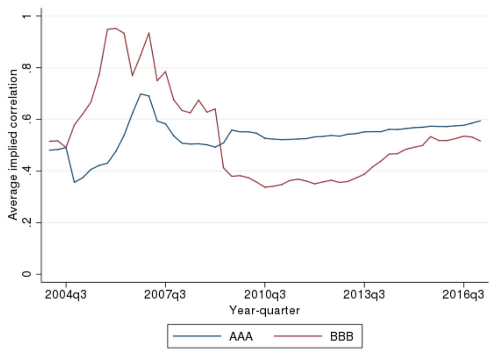
(b) Rating - Alt-A



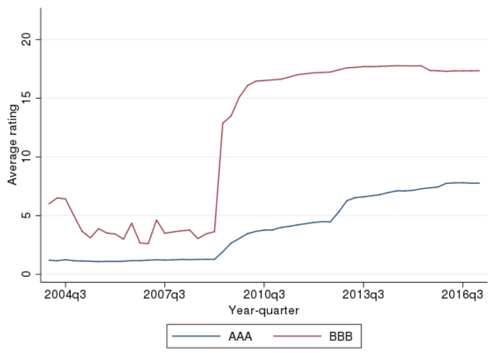
(c) Implied correlation - prime



(d) Rating - prime



(e) Implied correlation - subprime



(f) Rating - subprime

Figure K.2: Performance of early vintage tranches: average implied correlation and average rating for bonds originated before June 2005. For a given we compute the implied correlation, at each point in time. The average is taken by transaction period, by coarse rating at origination (AAA=1,..., BBB=4,..., D=8).

	(1) Alt-A	(2) Prime	(3) Second Lien	(4) Subprime
		Downgrade indicator		
Correlation at first transaction	0.198*	0.293**	-0.907	0.266***
	(0.101)	(0.130)	(0.910)	(0.0693)
Observations	8,766	11,862	60	8,620
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y

Standard errors in parentheses

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

Table 41: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to a given asset type. Errors are clustered at deal level.

	(1) [0, 0.25)	(2) [0.25, 0.5)	(3) [0.5, 0.75)	(4) [0.75, 1]
		Downgrade indicator		
Correlation at first transaction	-0.142	0.237**	0.357***	0.281**
	(0.203)	(0.105)	(0.0871)	(0.130)
Observations	3,149	7,274	8,824	7,096
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 42: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

	(1)	(2)	(3)	(4)
	[0, 0.25)	[0.25, 0.5)	[0.5, 0.75)	[0.75, 1]
	Downgrade indicator - AAA only			
Correlation at first transaction	-0.539	-0.0777	0.378*	0.595**
	(0.356)	(0.168)	(0.210)	(0.297)
Observations	1,760	4,544	4,538	4,369
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y
	Downgrade indicator - not AAA			
Correlation at first transaction	-0.147	0.106	0.221**	0.0974
	(0.271)	(0.126)	(0.0908)	(0.138)
Observations	1,204	2,704	4,222	2,701
Rating at first transaction	Y	Y	Y	Y
Vintage year	Y	Y	Y	Y
Asset type	Y	Y	Y	Y

Standard errors in parentheses

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

Table 43: Regression results from running logit specification 12 by maximum likelihood, controlling for vintage year, for vintages up to June 2005. The dependent variable is a dummy indicator for whether there was a downgrade by December 2009. Independent variables include implied correlation and coarse rating dummy indicator at the time of the first transaction. The dependent variable is the downgrade indicator. Each column presents the results on a subset of the data corresponding to the average documentation index corresponding to the given deal. Errors are clustered at deal level.

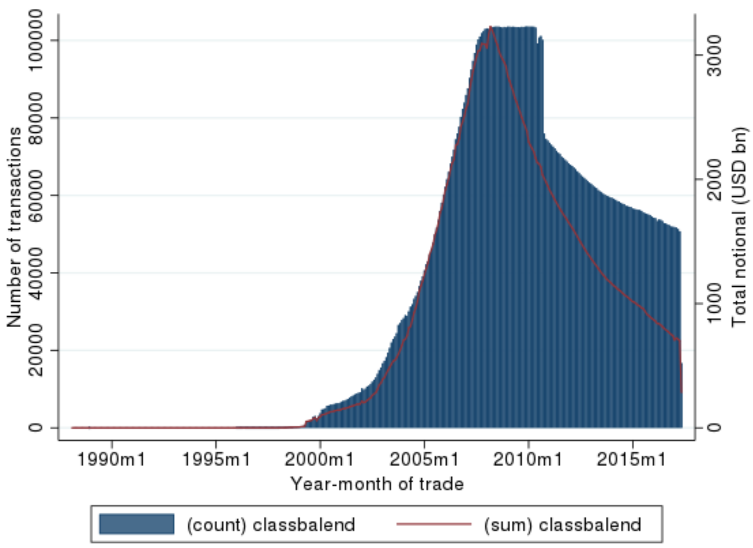


Figure L.1: Tranche balance and number of bonds outstanding by transaction year and month.

ratings rely on an a priori assessment, its evolution over time reflects progressively more of the bond performance, implying that updated rating values should in principle absorb the information that was initially private. We estimate panel 16 using a linear model, with random effects in order to control for tranche-invariants such as first rating and first implied correlation. The advantage of the panel specification 16 is that we can incorporate the partial information coming from the late vintages (after June 2005).

$$outcome_{it} = \alpha_{it} + \beta_0 \rho_{i,0} + \eta_0 rating_{i0} + \beta_1 \rho_{i,t-1} + \eta_1 rating_{i,t-1} + \gamma X_{i,t} + \varepsilon_{it}. \tag{16}$$

In equation 16 $outcome_{it}$ is the month-on-month rating change in notches. Table 44 shows that updates in signals contain information about future bond performance, but the signal is not statistically sufficient for prices. This suggests that investors retain private information over the life of the bond, besides the information given by agency ratings. The second finding is that Alt-A investors do not learn over the life of the bond, so that news in ratings remain statistically sufficient for news in correlation in terms of bond performance.

To see the effect of the crisis on the information content of prices, we will use interactions with an indicator dummy for post-2007 transaction to split estimates between before and after the crisis. The regression specification is the following:

$$\Delta rating_{it} = \alpha_{it} + \beta_0 \rho_{i,0} + \eta_0 rating_{i0} + 1_{post-07} \tag{17}$$

$$+\beta_1 \rho_{i,t-1} \times 1_{post-07} + \eta_1 rating_{i,t-1} \times 1_{post-07} + \gamma X_{i,t} + \varepsilon_{it} \tag{18}$$

Table 45 presents the results of estimating equation (18). It shows that most of the effect of news about default correlation shown in Table 44 comes from the post-crisis period. Griffin and Nickerson (2016) discuss how rating agencies improved their methodologies following the crisis. Under such improvement, the expectation would be that ratings become sufficient for implied correlations, but this is not what we observe. An improvement in rating methodology is consistent with more statistical information coming from prices if ratings are now following the market more closely. In that case changes in implied correlation have more statistical power to predict future downgrades by construction of the downgrade process. The other possibility is that investors learned more from the crisis than the rating agencies, but if this is so it is rational for ratings to follow the market more closely.

In DeMarzo (2005), two factors drive the benefits and drawbacks of securitization: private information by the issuer, on one hand, and asset correlation, on the other. Like CDOs, CMOs are a priori affected by it.[31] Our measure of beliefs about default correlation reflects in part adverse selection concerns on the part of the investors. Because we can't disentangle these two components empirically, our implied correlation measure is a proxy for market conservatism vis-à-vis information asymmetry. In line with this, Alt-A deals being more reliant on ratings (not only for the first

[31] Beltran et al. (2017) show that, under asymmetric information, even a modest percentage of bad securities can push security prices far below fundamentals -even to a market meltdown-.

	(1)	(2)	(3)	(4)
	Alt-A	Prime	Second Lien	Subprime
	One-month change in rating (notches)			
Lagged correlation (1 month)	0.004	0.007**	-0.025***	-0.007***
	(0.004)	(0.003)	(0.008)	(0.003)
Lagged rating (1 month)	-0.026***	-0.012***	-0.026***	-0.038***
	(0.001)	(0.001)	(0.003)	(0.001)
Correlation at first transaction	0.003	0.027***	-0.004	0.008***
	(0.005)	(0.003)	(0.008)	(0.003)
Rating at first transaction	0.017***	0.017***	0.001	0.010***
	(0.001)	(0.001)	(0.005)	(0.001)
Subordination	0.084***	0.090***	-0.119***	-0.148***
	(0.022)	(0.019)	(0.019)	(0.008)
Observations	2,032,055	1,773,020	55,293	1,452,760
Vintage	Y	Y	Y	Y
Year-quarter	Y	Y	Y	Y

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

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

Table 44: Regression results from running the panel regression 16, by GLS with tranche random effects. The first line gives the coefficient for the change over 1 month (lagged 1 month) of the correlation coefficient, and the second one the coefficient for the change over 1 month (lagged 1 month) of the change in rating (in notches). Errors are clustered at deal level.

	(1)	(2)	(3)	(4)
	Alt-A	Prime	Second Lien	Subprime
	Size of downgrade (notches)			
Lagged correlation	-0.001	-0.018***	0.002	0.005
	(0.005)	(0.004)	(0.017)	(0.003)
Lagged correlation × post-07=1	0.005	0.027***	-0.028	-0.013***
	(0.006)	(0.004)	(0.018)	(0.004)
Lagged rating	-0.055***	-0.095***	-0.082***	-0.026***
	(0.016)	(0.009)	(0.014)	(0.008)
Lagged rating × post-07=1	0.029*	0.083***	0.056***	-0.012
	(0.016)	(0.009)	(0.014)	(0.008)
post-07=1	0.016	-0.076***	-0.013	0.130***
	(0.019)	(0.010)	(0.034)	(0.011)
Correlation at first transaction	0.003	0.028***	-0.001	0.008***
	(0.005)	(0.003)	(0.008)	(0.003)
Rating at first transaction	0.017***	0.017***	0.003	0.009***
	(0.001)	(0.001)	(0.005)	(0.001)
Subordination	0.084***	0.089***	-0.120***	-0.148***
	(0.022)	(0.019)	(0.019)	(0.008)
Observations	2,032,055	1,773,020	55,293	1,452,760
Vintage	Y	Y	Y	Y
Year-quarter	Y	Y	Y	Y

Marginal effects; Standard errors in parentheses

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

Table 45: Regression results from running the panel regression 16, by GLS with tranche random effects. The first line gives the coefficient for the change over 1 month (lagged 1 month) of the correlation coefficient, and the second one the coefficient for the change over 1 month (lagged 1 month) of the change in rating (in notches). Errors are clustered at deal level.

transaction, but also for subsequent ones) potentially reflects a concern for asymmetric information as in Adelino et al. (2016).

M Additional causes of default clustering: frailty and contagion

Following Azizpour et al. (2016), defaults are driven by three factors: systemic risk[32] as captured by macroeconomic variables (Bullard et al., 2009; Khandani et al., 2013)[33], an unobserved frailty factor (Duffie et al., 2009; Kau et al., 2011) and a contagion factor, which captures the extent to which more defaults increase the conditional intensity of default arrival.

A given loan n has a default time T_n. Defaults have a conditional mean of arrival λ given by

$$\lambda_t = \exp\left(a_0 + \sum_{i=1}^{d} a_i X_{i,t}\right) + Y_t + Z_t$$

where X represents a vector of macroeconomic variables. Unobservable frailty Z_t follows the CIR process

$$\begin{aligned} dZ_t &= k(z - Z_t)dt + \sigma\sqrt{Z_t}dW_t \\ Z_0 &\sim \Gamma\left(\frac{2kz}{\sigma^2}, \frac{\sigma^2}{2k}\right) \end{aligned}$$

Defaults are self-exciting, in the sense that the mass of defaults at a given time increases the rate of arrival. This is captured by means of a contagion factor Y such that

$$\begin{aligned} Y_t &= b \sum_{n:T_n \leq t} e^{-\kappa(t-T_n)} U_n \\ U_n &= \max(0, \log u_n) \end{aligned}$$

where u_n is the sum of defaulted debt at time T_n. This implies that larger defaults are followed by more defaults.

The estimation of $\theta = (a, k, z, \sigma, b, \kappa)$ is a filtered likelihood problem (the likelihood is a posterior mean of the complete-data likelihood), and can be solved following Giesecke and Schwenkler (2016). The likelihood is written as a product of two terms, one that depends on event data (defaults) and one that depends on factor data. The decomposition is based on a change of measure, which

[32] Bisias et al. (2012) provides a survey of systemic risk measures. See also Chan-Lau et al. (2009). Other approaches include macro measures (costly asset-price boom/bust cycles, property-price, equity-price, credit-gap indicators), forward-looking measures (e.g. absorption rate as in Kritzman et al. (2010)), cross-sectional measures (CoVaR, Co-Risk, marginal and systemic expected shortfall, see Acharya et al. (2012)), stress tests (e.g. Duffie (2011)), illiquidity and insolvency (e.g. Brunnermeier et al. (2011)). Giglio et al. (2013) use predictive quantile regression to provide an empirical assessment of 17 of them. Their main finding is that, overall, the compendium of systemic risk measures contains useful predictive information. Instead individual measures tend to fail in capturing systematic risk.

[33] The characterization of systemic risk as deterioration of macroeconomic indicators leaves aside the widely discussed view that the pre-crisis mortgage system was systemically vulnerable (Hellwig, 2009; Poitras and Zanotti, 2016).

resolves the interaction between the point process and the factors of λ.[34] One of the terms is a point process filter, which makes the computation difficult. Giesecke and Schwenkler (2016) propose an approximation based on a quadrature method, from which the posterior mean can be computed. They write an algorithm and derive conditions for convergence.

N Supplemental graphs and tables

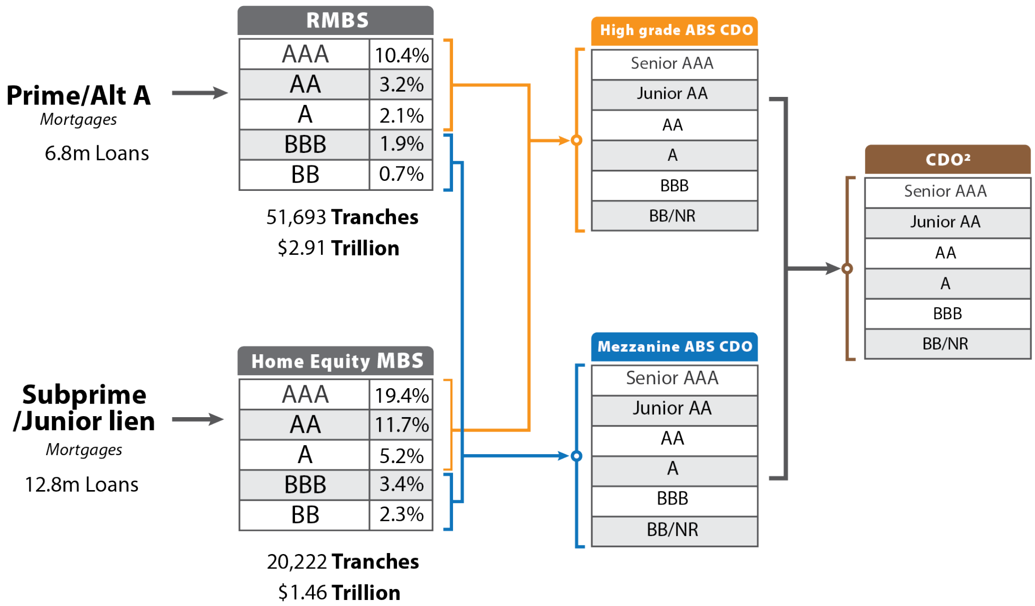
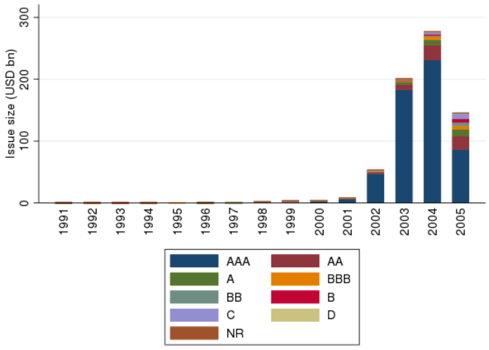


Figure N.1: Diagram: from loans to RMBS CMO, from CMO to CDO, from CDO to CDO^2. Details are reported on the total number of loans recorded by ABSNet, the universe of securities issued and the average subordination percentage by Standard & Poor's rating, as explained in Section 1

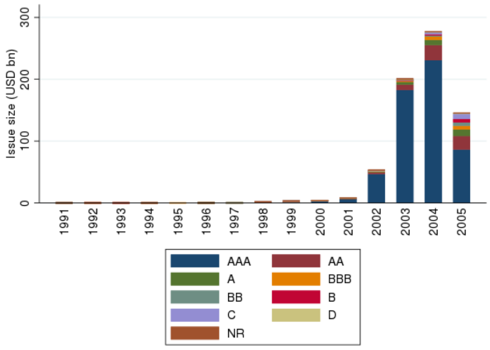
Asset type	After Jun-05		Before Jun-05	
	Origination ($bn)	Count	Origination ($bn)	Count
Alt-A	1,179.0	16,837	557.7	11,000
Prime	621.7	9,097	557.9	14,759
Second Lien	64.7	478	19.0	408
Subprime	660.0	9,811	720.2	9,525
Total	2,525.4	36,223	1,854.8	35,692

Table 46: Issued amounts and counts by asset type.

[34] An alternative is to apply the expectation maximization (EM) algorithm. Giesecke and Schwenkler (2016) compare the two approaches.



(a) Number of tranches



(b) Total issue

Figure N.2: Number of tranches and amount issued by vintage year for private label collateralized mortgage obligations. Source: ABSNet bond data. For our sample of early vintages (prior to June 2005) we provide the distribution by (coarse, see Table 38) initial rating.

	Our sample		Cordell et al. (2012)	
rating	Prime/Alt-A	Second Lien/Subprime	Prime/Alt-A	Second Lien/Subprime
AAA	10.8%	25.7%	6%	23%
AA	3.4%	14.3%	3%	13%
A	3.0%	5.9%	2%	8%
BBB	2.9%	4.0%	1%	4%

Table 47: Subordination percentage by tranche rating - comparison. The figures computed using ABSNet data are derived by aggregating the subordination percentages at orgination as given in Table 1.4. Our sample contains only early vintages (prior to June 2005) while Cordell et al. (2012) use late vintages as well.

Year	ABSNet sample		Adelino (2009)	
	Origination ($bn)	Count	Origination ($bn)	Count
≤2002	319.3	5,438		
2003	470.5	10,120	496.5	8,574
2004	677.4	12,519	767.3	11,460
2005	904.5	16,684	1,058.5	17,135
2006	1,038.0	15,022	1,080.4	18,206
2007	939.4	11,716	802.1	12,037
≥2008	31.2	177		
Total	4,380.3	71,676	4,204.8	67,412

Table 48: Origination amounts and counts at origination, by vintage year, compared to the sample in Adelino (2009).

	(1)	(2)
Asset type	Early vintages	Late vintages
Alt-A	7.5%	19.5%
Prime	2.3%	6.6%
Second Lien	7.2%	25.8%
Subprime	14.8%	30.5%
Observations	4,060,698	631,793

Table 49: Liquidation rates from the loan sample. Column (1) calculates the percentage of loans linked to early vintage deals (before June 2005) that are liquidated. Column (2) calculates the same ratio for late vintage loans.

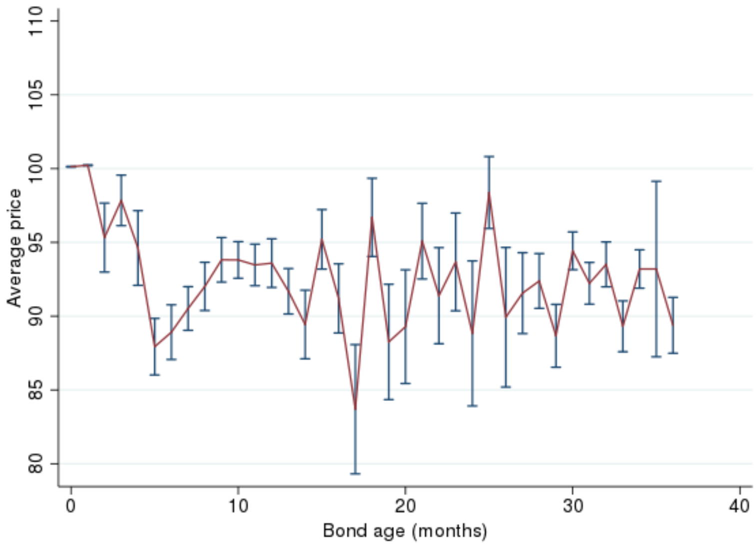


Figure N.3: Average tranche price by age of the bond in months. For our sample of bonds originated in 2004 and 2005 we compute the average price by the time elapsed (in months) since the bond issue. Vertical whiskers show the standard errors.

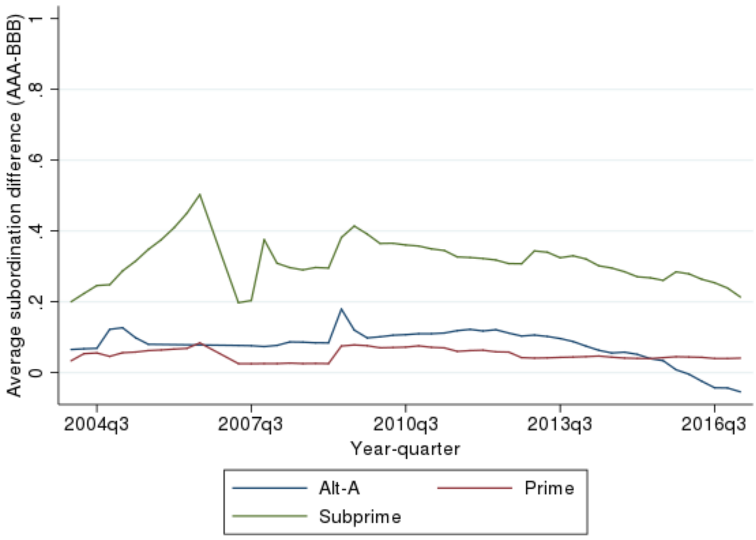


Figure N.4: Average subordination difference between AAA and BBB bonds. Source: ABSNet bond data.The figure presents the difference between the average AAA and average BBB subordination over trading time (for early vintages, prior to June 2005) using the rating at the given trading time. The difference is computed by asset type.

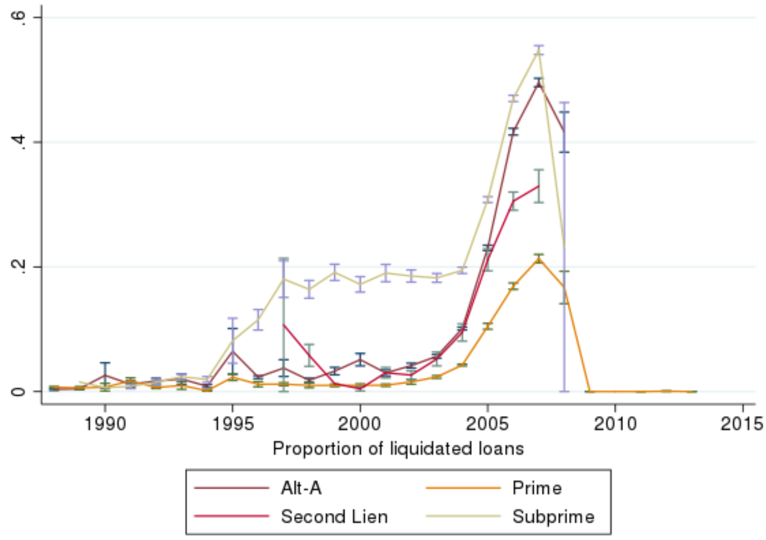


Figure N.5: Probability of default by vintage year. We compute the default rate for each of the deals that compose our population, and then average by vintage year and asset type. The results are presented here along with standard error bands around the average.

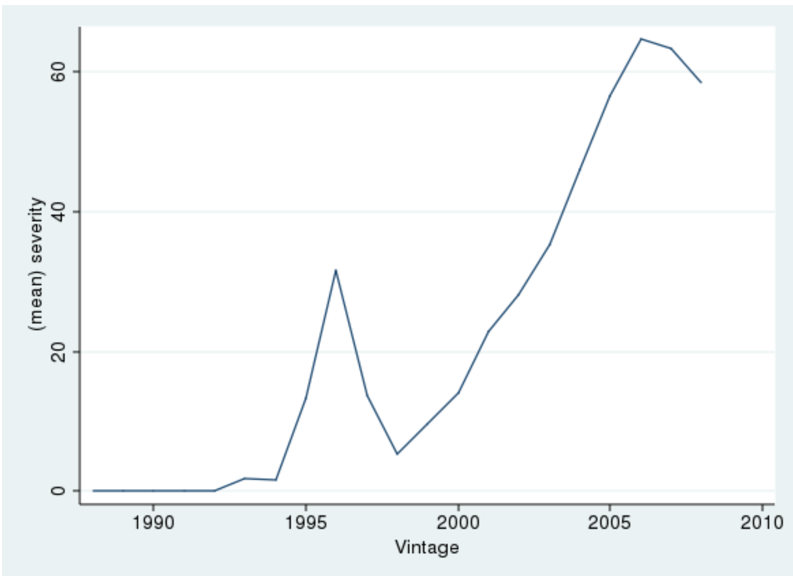


Figure N.6: Percentage loss given default by vintage year. The aggregate loss given default is computed from the sample of loans associated to the deals that compose our population of CMOs.

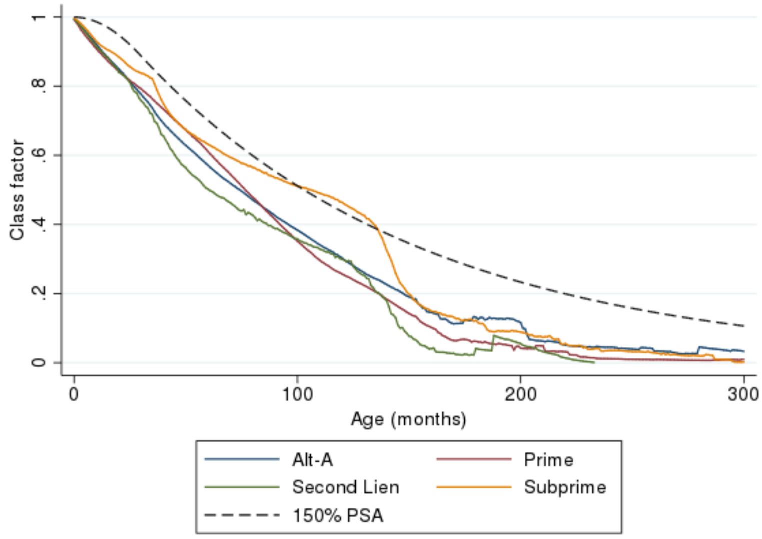


Figure N.7: Average class balance factor by asset class over tranche age. Alongside the averages, we compute the balance factor that results from a 150% payment schedule alone (excluding planned amortization).

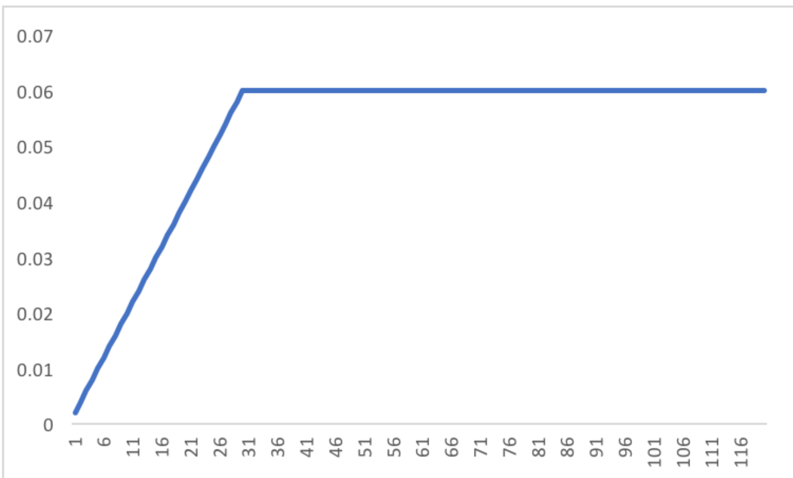


Figure N.8: Standard Prepayment Model of The Bond Market Association. Prepayment percentage for each month in the life of the underlying mortgages, expressed on an annualized basis.

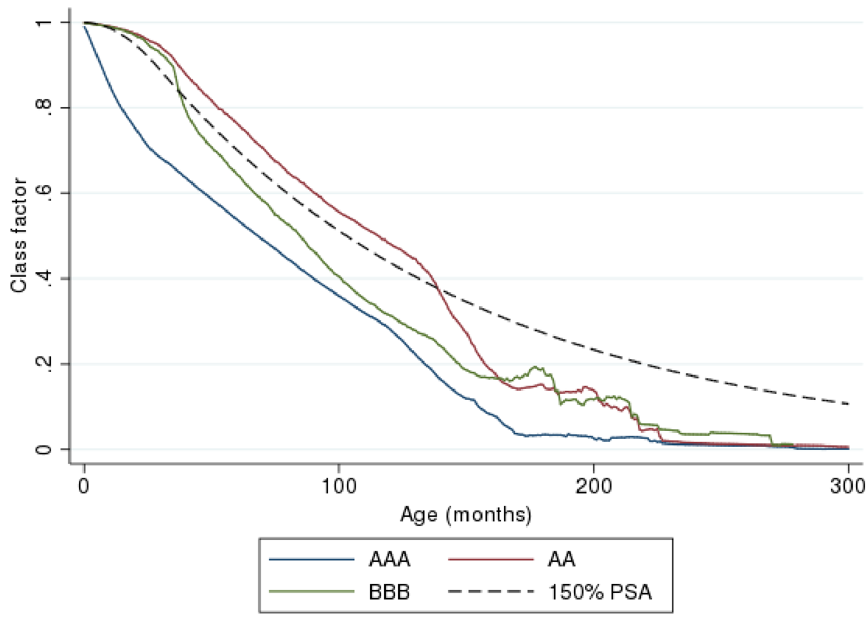
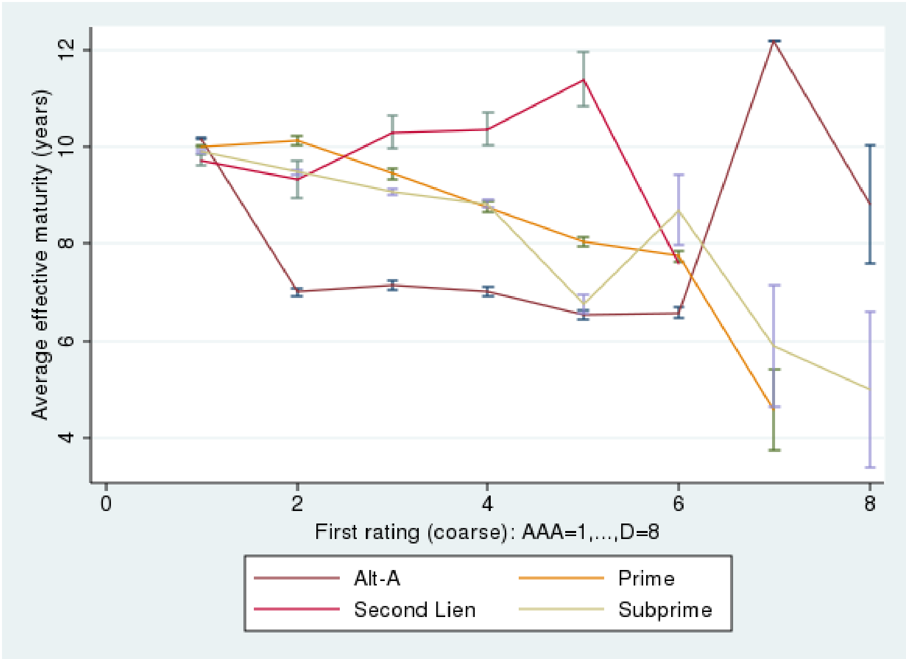
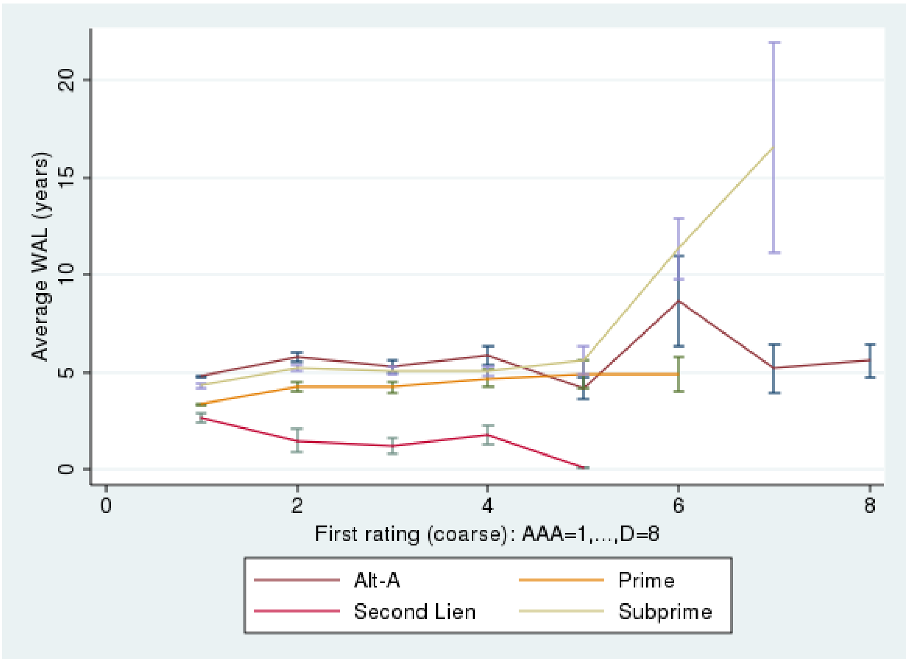


Figure N.9: Plot of average class factor against tranche age by tranche initial rating.



(a) Average realized



(b) WAL

Figure N.10: Average realized and weighted average life by coarse rating and asset type. The second panel includes observations where we found a matching WAL in Bloomberg.

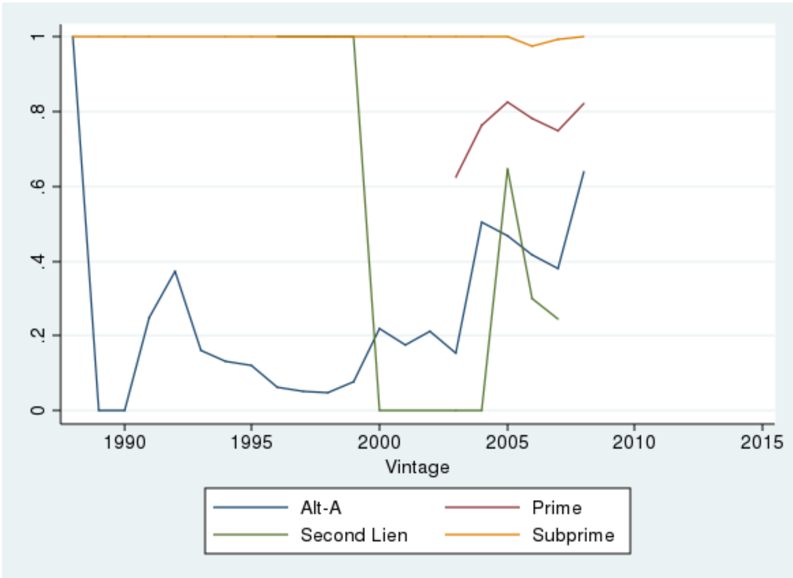


Figure N.11: Proportion of ARM loans by vintage and asset type.

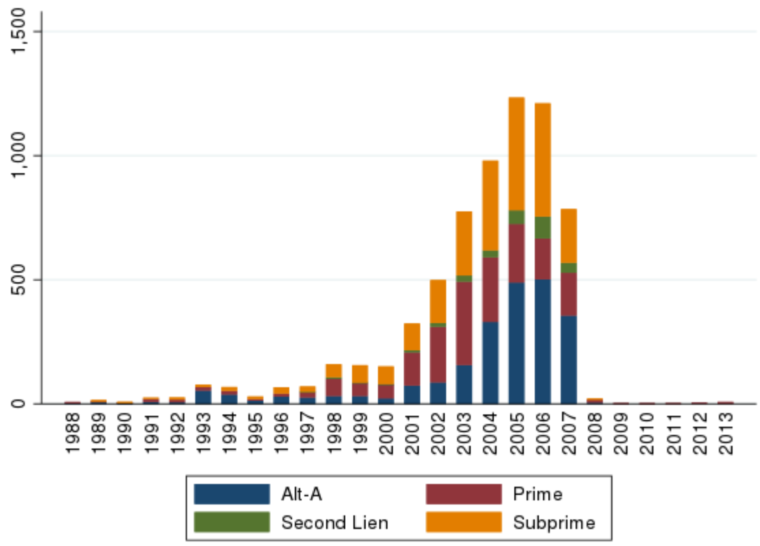


Figure N.12: Number of deals originated by asset type and vintage year.

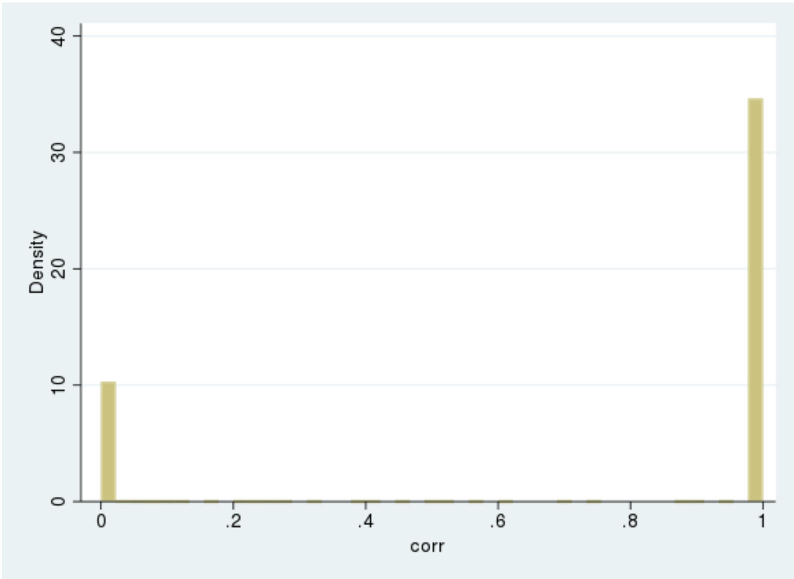


Figure N.13: Histogram plotting all outcomes from the pricing model.

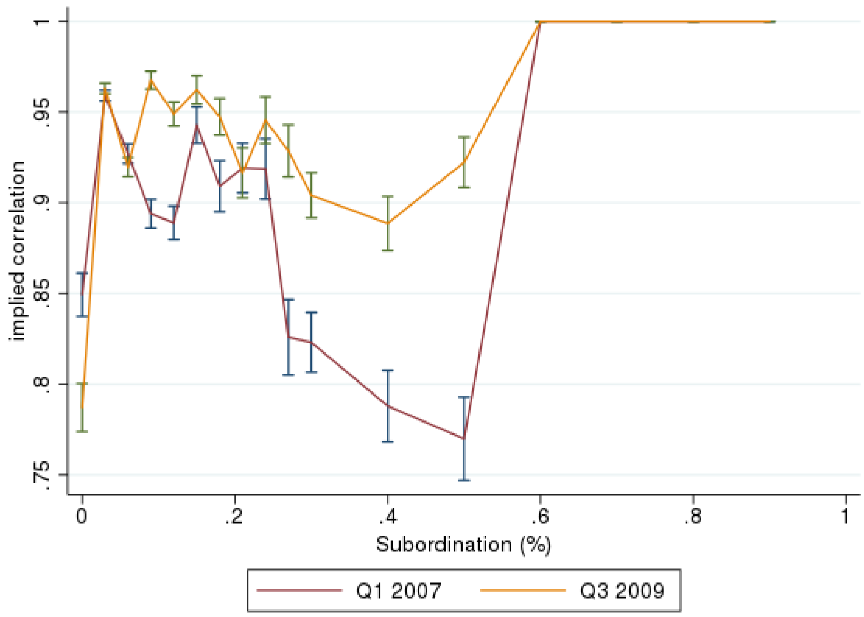
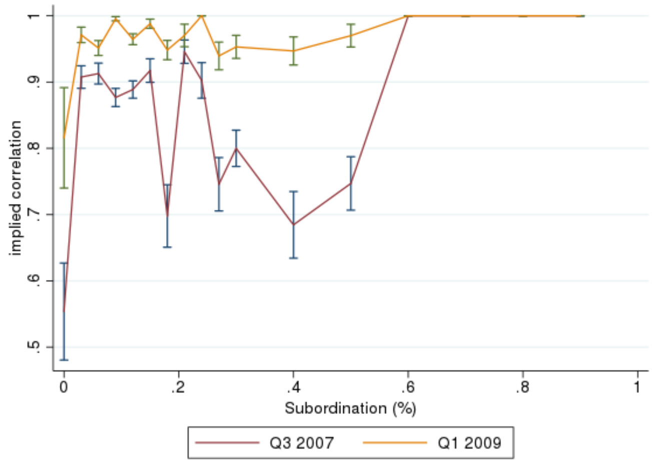
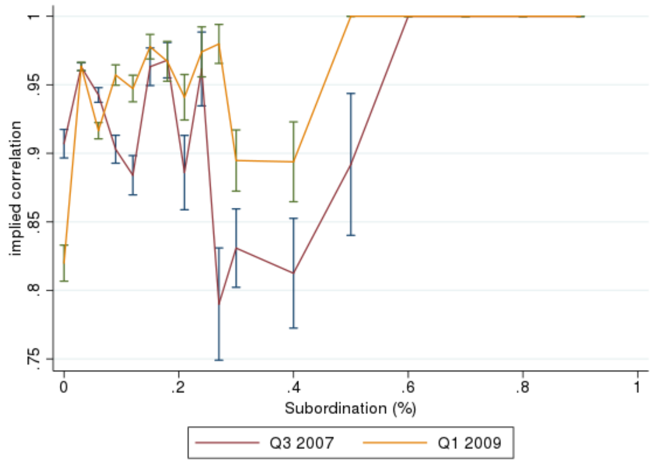


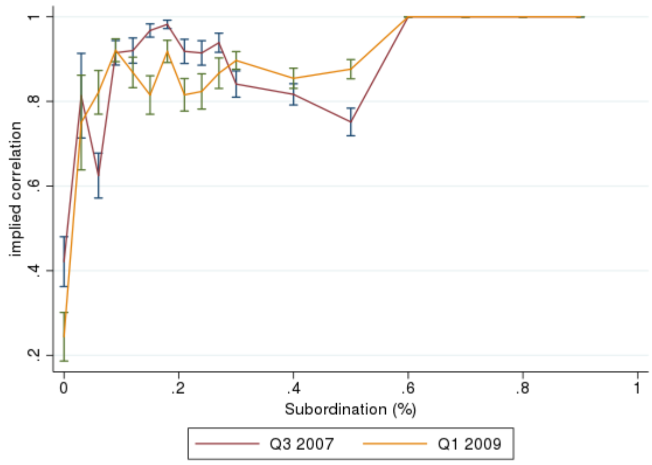
Figure N.14: Average correlation plotted against tranche subordination percentage, on two given dates. We use the sample of early vintage bonds (originated prior to June 2005). Subordinations are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.



(a) Alt-A



(b) Prime



(c) Subprime

Figure N.15: Average correlation plotted against tranche subordination percentage, on two given dates. Subordination values are assigned to 10 equally spaced bins. Within each subordination bin we plot the average correlation, along with vertical whiskers representing the standard error of the average.

References

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. P. Richardson (2012, Feb). Measuring systemic risk.

Adelino, M. (2009). Do investors rely only on ratings? the case of mortgage-backed securities.

Adelino, M., K. Gerardi, and B. Hartman-Glaser (2016). Are lemons sold first? dynamic signaling in the mortgage market.

Altman, E. I. (2006). Default recovery rates and lgd in credit risk modeling and practice: an updated review of the literature and empirical evidence. *New York University, Stern School of Business*.

An, X., Y. Deng, J. B. Nichols, and A. B. Sanders (2015). What is subordination about? credit risk and subordination levels in commercial mortgage-backed securities (CMBS). *The Journal of Real Estate Finance and Economics 51*(2), 231–253.

Andreoli, A., L. V. Ballestra, and G. Pacelli (2016). From insurance risk to credit portfolio management: a new approach to pricing CDOs. *Quantitative Finance 16*(10), 1495–1510.

Ashcraft, A., P. Goldsmith-Pinkham, P. Hull, and J. Vickery (2011). Credit ratings and security prices in the subprime mbs market. *The American Economic Review 101*(3), 115–119.

Ashcraft, A., P. Goldsmith-Pinkham, and J. Vickery (2010). Mbs ratings and the mortgage credit boom.

Ashcraft, A. B. and T. Schuermann (2008). Understanding the securitization of subprime mortgage credit. *Foundations and Trends® in Finance 2*(3), 191–309.

Azizpour, S., K. Giesecke, and G. Schwenkler (2016). Exploring the sources of default clustering.

Bai, J., P. Collin-Dufresne, R. S. Goldstein, and J. Helwege (2015). On bounding credit-event risk premia. *The Review of Financial Studies 28*(9), 2608.

Beltran, D. O., L. Cordell, and C. P. Thomas (2017). Asymmetric information and the death of ABS CDOs. *Journal of Banking & Finance 76*, 1 – 14.

Benmelech, E. and J. Dlugosz (2010). The credit rating crisis. *NBER Macroeconomics Annual 24*(1), 161–208.

Berndt, A., R. Douglas, D. Duffie, M. Ferguson, and D. Schranz (2005). Measuring default risk premia from default swap rates and edfs.

Bisias, D., M. Flood, A. W. Lo, and S. Valavanis (2012). A survey of systemic risk analytics. *Annual Review of Financial Economics 4*(1), 255–296.

Boot, A. W. and A. V. Thakor (1993). Security design. *The Journal of Finance 48*(4), 1349–1378.

Brunne, T. (2006). Implied correlation of synthetic CDOs with liquid markets.

Brunnermeier, M. K., G. Gorton, and A. Krishnamurthy (2011). *Risk Topography*, pp. 149–176. National Bureau of Economic Research, Inc.

Bullard, J., C. J. Neely, D. C. Wheelock, et al. (2009). Systemic risk and the financial crisis: a primer. *Federal Reserve Bank of St. Louis Review 91*(5 Part 1), 403–18.

Buzková, P. and P. Teplý (2012). Collateralized debt obligations valuation using the one factor gaussian copula model. *Prague Economic Papers 21*(1), 30–49.

Carlin, B. I., F. A. Longstaff, and K. Matoba (2014). Disagreement and asset prices. *Journal of Financial Economics 114*(2), 226 – 238.

Chan-Lau, J. A., M. Espinosa, K. Giesecke, and J. A. Sole (2009). Assessing the systemic implications of financial linkages. *IMF Global Financial Stability Report 2*.

Cordell, L., Y. Huang, and M. Williams (2012). Collateral damage: Sizing and assessing the subprime CDO crisis.

Cornaggia, J. N., K. J. Cornaggia, and J. E. Hund (2017). Credit ratings across asset classes: A long-term perspective*. *Review of Finance 21*(2), 465–509.

Coval, J., J. Jurek, and E. Stafford (2009a). The economics of structured finance. *The Journal of Economic Perspectives 23*(1), 3–25.

Coval, J. D., J. W. Jurek, and E. Stafford (2009b). Economic catastrophe bonds. *The American Economic Review 99*(3), 628–666.

Cowan, A. M. and C. D. Cowan (2004). Default correlation: An empirical investigation of a subprime lender. *Journal of Banking & Finance 28*(4), 753 – 771. Retail Credit Risk Management and Measurement.

Crouhy, M., D. Galai, and R. Mark (2000). A comparative analysis of current credit risk models. *Journal of Banking & Finance 24*(1), 59–117.

D'Amato, J. and J. Gyntelberg (2005). CDS index tranches and the pricing of credit risk correlations. *BIS Quarterly Review*.

Dang, T. V., G. Gorton, and B. Holmström (2013). The information sensitivity of a security. *Unpublished working paper, Yale University*.

de Servigny, A. and O. Renault (2002). Default correlation: empirical evidence. *Standard and Poor's*.

Dell'Ariccia, G., D. Igan, and L. Laeven (2012, 03). Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market. *Journal of Money, Credit and Banking 44*, 367–384.

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

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

Deng, Y., S. A. Gabriel, and A. B. Sanders (2011). CDO market implosion and the pricing of subprime mortgage-backed securities. *Journal of Housing Economics 20*(2), 68 – 80. Special Issue: Housing and the Credit Crunch.

Denzler, S. M., M. M. Dacorogna, U. A. Müller, and A. J. McNeil (2006). From default probabilities to credit spreads: Credit risk models do explain market prices. *Finance Research Letters 3*(2), 79–95.

Diep, P., A. L. Eisfeldt, and S. A. Richardson (2016). Prepayment risk and expected mbs returns.

Downing, C., R. Stanton, and N. Wallace (2005). An empirical test of a two-factor mortgage valuation model: How much do house prices matter? *Real Estate Economics 33*(4), 681–710.

Duffie, D. (2008, July). Innovations in credit risk transfer: Implications for financial stability. *BIS Working Paper* (255).

Duffie, D. (2011, August 2011). Systemic risk exposures: A 10-by-10-by-10 approach. *National Bureau of Economic Research Working Paper Series No. 17281.*

Duffie, D., A. Eckner, G. Horel, and L. Saita (2009). Frailty correlated default. *The Journal of Finance 64*(5), 2089–2123.

Duffie, D. and N. Gârleanu (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts Journal 57*(1), 41–59.

Duffie, D. and K. J. Singleton (2012). *Credit risk: pricing, measurement, and management.* Princeton University Press.

Elizalde, A. (2005). Credit risk models iv: Understanding and pricing cdos. *CEMFI and Universidad Publica de Navarro, download: www. cemfi. es/ elizalde.*

Finkelstein, D., A. Strzodka, and J. Vickery (2018). Credit risk transfer and de facto gse reform. Staff Reports 838, Federal Reserve Bank of New York.

Frye, J. (2008). Correlation and asset correlation in the structural portfolio model. *The Journal of Credit Risk 4*(2), 75–96.

Geidosh, M. (2014, 06). Asset correlation in residential mortgage-backed security reference portfolios. *Journal of Credit Risk 10*(2), 71–95.

Ghent, A. C., W. N. Torous, and R. Valkanov (2016). Complexity in structured finance. *The Review of Economic Studies*.

Giesecke, K. and G. Schwenkler (2016). Filtered likelihood for point processes.

Giglio, S., B. Kelly, S. Pruitt, and X. Qiao (2013). Systemic risk and the macroeconomy: An empirical evaluation. *Chicago Booth Research Paper* (12-49).

Gordy, M. B. (2000). A comparative anatomy of credit risk models. *Journal of Banking & Finance 24*(1), 119–149.

Gorton, G. (2009). The subprime panic*. *European Financial Management 15*(1), 10–46.

Gorton, G. and G. Pennacchi (1990). Financial intermediaries and liquidity creation. *The Journal of Finance 45*(1), 49–71.

Gorton, G. B. and G. Ordonez (2013). The supply and demand for safe assets. Technical report, National Bureau of Economic Research.

Griffin, J. M. and J. Nickerson (2016). Debt correlations in the wake of the financial crisis: What are appropriate default correlations for structured products?

Griffin, J. M. and D. Y. Tang (2012). Did subjectivity play a role in cdo credit ratings? *The Journal of Finance 67*(4), 1293–1328.

Gupta, A. (2016). Foreclosure contagion and the neighborhood spillover effects of mortgage defaults. Technical report, Working Paper.

Hellwig, M. F. (2009). Systemic risk in the financial sector: An analysis of the subprime-mortgage financial crisis. *De Economist 157*(2), 129–207.

Heynderickx, W., J. Cariboni, W. Schoutens, and B. Smits (2016). The relationship between risk-neutral and actual default probabilities: the credit risk premium. *Applied Economics 48*(42), 4066–4081.

Hull, J. and A. White (2008). Dynamic models of portfolio credit risk: A simplified approach. *Journal of Derivatives 15*(4), 9–28.

Hull, J. C. and A. D. White (2004). Valuation of a CDO and an n-th to default CDS without monte carlo simulation. *The Journal of Derivatives 12*(2), 8–23.

Hull, J. C. and A. D. White (2006). Valuing credit derivatives using an implied copula approach. *The Journal of Derivatives 14*(2), 8–28.

IOSCO (2008, March). The role of credit rating agencies in structured finance markets. Technical report, International Organization of Securities Commissions.

Jarrow, R. A. (2011). Risk management models: construction, testing, usage. *The Journal of Derivatives 18*(4), 89–98.

JEC (2007). The subprime lending crisis: The economic impact on wealth, property values and tax revenues, and how we got here. *World Wide Web page¡ http://www. jec. senate. gov/Documents/Reports/10.25. 07OctoberSubprimeReport. pdf¿(accessed January 15, 2008).*

Kau, J. B., D. C. Keenan, and X. Li (2011). An analysis of mortgage termination risks: a shared frailty approach with msa-level random effects. *The Journal of Real Estate Finance and Economics 42*(1), 51–67.

Kau, J. B., D. C. Keenan, C. Lyubimov, and V. C. Slawson (2011). Subprime mortgage default. *Journal of Urban Economics 70*(2), 75–87.

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

Khandani, A. E., A. W. Lo, and R. C. Merton (2013). Systemic risk and the refinancing ratchet effect. *Journal of Financial Economics 108*(1), 29–45.

Kritzman, M., Y. Li, S. Page, and R. Rigobon (2010). Principal components as a measure of systemic risk.

Li, D. X. (2000). On default correlation: A copula function approach. *The Journal of Fixed Income 9*(4), 43–54.

Liu, H. (2016). Do government guarantees inhibit risk management? evidence from Fannie Mae and Freddie Mac.

McGinty, L., E. Beinstein, R. Ahluwalia, and M. Watts (2004). Credit correlation: A guide. Technical report.

Nagpal, K. and R. Bahar (2001). Measuring default correlation. *Risk 14*(3), 129–132.

O'Kane, D. and M. Livesey (2004). Base correlation explained. *Lehman Brothers, Fixed Income Quantitative Credit Research*.

Palmer, C. (2015). Why did so many subprime borrowers default during the crisis: Loose credit or plummeting prices?

Poitras, G. and G. Zanotti (2016). Mortgage contract design and systemic risk immunization. *International Review of Financial Analysis 45*, 320–331.

Rapisarda, G. and D. Echeverry (2013). A nonparametric approach to incorporating incomplete workouts into loss given default estimates. *The Journal of Credit Risk 9*(2), 47.

Schwartz, E. S. and W. N. Torous (1989). Prepayment and the valuation of mortgage-backed securities. *The Journal of Finance 44*(2), 375–392.

Simsek, A. (2013). Belief disagreements and collateral constraints. *Econometrica 81*(1), 1–53.

Sirignano, J., A. Sadhwani, and K. Giesecke (2016). Deep learning for mortgage risk.

Skreta, V. and L. Veldkamp (2009). Ratings shopping and asset complexity: A theory of ratings inflation. *Journal of Monetary Economics 56*(5), 678–695.

Stanton, R. (1995). Rational prepayment and the valuation of mortgage-backed securities. *The Review of Financial Studies 8*(3), 677.

Stanton, R. and N. Wallace (2011). The bear's lair: Index credit default swaps and the subprime mortgage crisis. *Review of Financial Studies 24*(10), 3250–3280.

Sugimura, T. (2004). Valuation of residential mortgage-backed securities with default risk using an intensity-based approach. *Asia-Pacific Financial Markets 11*(2), 185–214.

Tzani, R. and A. Polychronakos (2008, 12). Correlation breakdown, copula credit default models and arbitrage. *Global Association of Risk Professionals Magazine*.