The CDS-Bond Basis

Abstract

We investigate the cross-sectional variation in the CDS-bond basis, which measures the difference between credit default swap (CDS) spread and cash-bond implied credit spread. We test several explanations for the violation of the arbitrage relation between cash bond and CDS contract, which states that the basis should be zero in normal conditions. The evidence is consistent with 'limits to arbitrage' theories in that deviations are larger for bonds with higher frictions as measured by trading liquidity, funding cost, counterparty risk, and collateral quality. Surprisingly however, we find that the basis is more negative when the bond lending fee is higher, suggesting that arbitrageurs are unwilling to engage in a negative basis trade when short interest on the bond is high.

1 Introduction

Financial markets experienced tremendous disruptions during the 2007-2009 financial crisis. Credit spreads across all asset classes and rating categories widened to unprecedented levels.[1] Perhaps even more surprising, many relations that were considered to be text-book arbitrage before the crisis were severely violated. For example, in currency markets, violations of covered interest rate parity occurred for currency pairs involving the US dollar (Coffey, Hrung, and Sarkar, 2009). In interest rate markets, the swap spread that measures the difference between Treasury bond yields and Libor swap rates turned negative. In interbank markets, basis swaps that exchange different tenor Libor rates (e.g., 3-month for 6-month) deviated from zero. In inflation markets, break-even inflation rates turned negative implying an arbitrage with inflation swaps (Fleckenstein, Longstaff, and Lustig, 2014). In credit markets, the CDS-bond basis that measures the difference between credit default swap (CDS) spreads and cash-bond implied credit spreads turned negative.

These anomalies suggest that such relations are not, in fact, arbitrage opportunities in the traditional textbook sense. Indeed, the arbitrage profits may be difficult to realize in practice. Many of these relations involve a fully funded (e.g., cash) instrument and one or more unfunded derivative positions. Thus, counterparty risk of the derivative issuer may have rendered the 'arbitrage' risky. Furthermore, funding cost differentials between the cash instrument and derivative positions may have made the arbitrage costly to implement for an investor requiring funding. In the latter case, the arbitrage violations persist because of 'limits to arbitrage' such as the inability of arbitrageurs to raise capital quickly and/or their unwillingness to take large positions in these 'arbitrage' trades because of mark-to-market risk. These apparent arbitrage violations thus provide an interesting opportunity to test several of

[1]For example, investment-grade corporate credit spreads as measured by the CDX.IG index rose from 50 basis points (bps) in early 2007 to more than 250 bps at the end of 2008. Even at the safest end of the spectrum the widening was dramatic. AAA-rated synthetic debt products, that would have been deemed virtually risk-free before the crisis, saw their spreads widen dramatically: CDX.IG super senior tranche widened from 5 bps to 100 bps, CMBX AAA "super duper" widened from 2 bps to 700 bps, ABS-HEL AAA tranche price rose from 0 to 20% upfront plus 500 bps running. These numbers illustrate that it became much more expensive to insure AAA-rated debt across various markets (corporate, residential and commercial real estate).

the 'limits to arbitrage' theories (as surveyed, for example, by Gromb and Vayanos, 2010).

In this paper, we focus on the CDS-bond basis, which measures the difference between the CDS spread of a specific company and the credit spread paid on a bond of the same company. Figure 1 plots the time series of the CDS-bond basis for investment-grade (IG) and high yield (HY) bonds, The plots show that the average basis for IG firms, which hovers usually around −17 basis points (bps) before the crisis, fell to −243 bps, and the average basis for HY firms dropped from 12 bps to −560 bps. Also, the bases for both IG and HY firms remain negative even after the financial crisis. At first sight, a large negative basis smacks of arbitrage since it suggests that an investor can purchase the bond, fund it at Libor, and insure the default risk on the bond by buying protection via the CDS contract. The resulting trade is 'virtually' risk-free and yet, as the plots show, it generates between 243 bps and 560 bps in guaranteed return per annum.

Studying the CDS-bond basis during the crisis is interesting for several reasons. First, early studies of this basis found that the arbitrage relation between CDS and cash-bond spreads holds fairly well during the pre-crisis period (Hull, Predescu, and White, 2004,Blanco, Brennan, and Marsh, 2005, Longstaff, Mithal, and Neis, 2005). In fact, if anything, these studies typically conclude that the basis should be slightly positive. Indeed, the arbitrage is, in general, not perfect (Duffie, 1999), and there are a few technical reasons (such as the difficulty in short-selling bonds and the cheapest-to-deliver option) that tend to push the basis into the positive domain (Blanco, Brennan, and Marsh, 2005). However, during the crisis the bases were tremendously negative, which suggests the need for alternative explanations.

Second, there is a large cross-sectional variation in the observed bases across individual firms. This cross-sectional variation makes the basis arbitrage an interesting laboratory to test various 'limits to arbitrage' theories. It is the focus of our paper.

There are several reasons why one might expect the basis to become negative during the financial crisis of 2007−2009. Anecdotes for the negative basis claim that several major financial institutions, pressed to free up their balance sheet and to improve their cash balance, reduced their leverage by selling off bonds. The argument is that deleveraging exerted downward pressure on bond prices and thus upward pressure on credit spreads relative to CDS spreads that represent the 'fair' value of the default risk insurance. This however cannot be the whole story since in a perfect frictionless market, investors would simply borrow cash to buy the bonds, buy protection and finance the position until maturity or default. For deleveraging to have a persistent impact on the basis, there must be some 'limits to arbitrage' (Shleifer and Vishny, 1997). In particular, if risk capital is limited then the mark-to-market basis trade becomes risky and investors will tend to buy the bonds-basis packages that are (ex-ante) most attractive from a risk-return tradeoff.

In this paper we analyze the risk-return tradeoff in a basis trade for an investor with limited capital. We find that the investor is exposed to (i) increased collateral values of the underlying bonds, (ii) increased illiquidity of bond trading, (iii) counterparty risk in buying CDS contracts, and (iv) increased funding constraint faced by investors who want to explore the arbitrage. For a given level of the basis, we expect investors with limited arbitrage capital to prefer those trades with better collateral, higher trading liquidity, less counterparty risk, and lower exposures to funding costs. Put differently, if arbitrageurs have limited capital, then in equilibrium we expect to see a larger basis (which can be thought of as an expected return) for firms with worse collateral quality, lower trading liquidity, more counterparty risk, and higher funding liquidity risk. To disentangle which explanation is most relevant in explaining the deviations in the CDS-bond basis, we construct measures of collateral quality, bond liquidity risk, counterparty risk, and funding liquidity risk for each firm, and explore whether they can explain the cross-section of the bases first for the individual bond bases and second for a set of 24 portfolios sorted on rating and size.

We find that all our 'limits to arbitrage' measures matter and together they 'explain' about 80% of the cross-sectional variation in the bases of 24 size and rating sorted portfolios and about 40% of the cross-sectional variation of individual firm bases. During the crisis, where the basis became significantly negative, all measures of collateral quality, bond illiquidity, counterparty risk and funding liquidity risk are statistically significantly correlated with the basis and with the expected sign, suggesting that the basis was more negative for a bond more costly to engage in the arbitrage trade. We find one exception, namely that the bond lending fee is associated negatively with the negative bond basis. This is surprising since a high lending fee is typically associated with high collateral value and thus should make the arbitrage trade easier to implement (since the negative basis arbitrage involves buying the bond, the lending fee could be earned by the arbitrageur). In contrast, and as expected, bond lending fee is associated positively with the basis when the bond basis is positive (a high lending fee indicates the bond is costly to short rendering the positive basis arbitrage trade more costly). We speculate that arbitrageurs refrain from engaging in a negative basis trade with a high-fee bond, because the high lending fee proxies the high short interest on the bond and thus signals that the basis might further widen. So arbitrageurs might be avoiding the 'catching-a-falling-knife' trade. In fact, we find that for firms with a negative basis, the basis tends to be more negative if the underlying bond has worse rating, higher CDS, and smaller size. All of these characteristics typically correlate negatively with collateral quality. So for firms with a negative basis, the lending fee seems to capture different information than collateral quality.

In sum, our cross-sectional results strongly support the hypothesis that limits to arbitrage prevented arbitrageurs from closing the basis gap.

The negative basis has attracted considerable attention in the practitioner literature[2]. These papers emphasize the role of financing risk in generating the negative basis, as well as deleveraging in generating downward price pressure on cash-bonds. In the academic literature, Garleanu and Pedersen (2011) provide a theoretical model, where leverage constraints can generate a pricing difference between two otherwise identical financial securities that differ in terms of their margin requirements or haircuts. Specifically, their theory predicts that the

[2] For example, D.E. Shaw Group, 2009, JP Morgan "The bond-CDS basis handbook" (2009), Mitchell and Pulvino, 2012.

difference between two bases should be related to the difference in margin requirements (i.e., haircuts) times the difference between the collateralized and uncollateralized borrowing rate.

Our study differs from previous papers in that we focus on the cross-sectional variation in individual firm bases (rather than on the average basis level) and try to relate it to firm, bond and CDS characteristics. Two contemporaneous papers have also investigated the CDS-bond basis. Fontana (2011) studies the time-series variation in the average basis using cointegration techniques. Kim, Li, and Zhang (2016, 2017) examine a long-short basis risk factor and study their implications on corporate bond returns.

The paper proceeds as follows. The next section discusses practical issues regarding an actual basis trade and isolate the various sources of risk in such a trade. Section 3 introduces the data. In Section 4 and 5 we construct limits-to-arbitrage measures and examine their impact in explaining the cross-sectional variation of the bases. Section 6 concludes.

2 The CDS-Bond Basis

A credit default swap is essentially an insurance contract against a credit event of a specific reference entity. It is an over-the-counter transaction between two parties in which the protection buyer makes periodic coupon payments to the protection seller until maturity or until some credit event happens. When a credit event occurs,[3] typically the protection buyer delivers a bond from an eligible pool to the protection seller in exchange for its par value.[4]

The contract is designed so that the owner of a particular bond can hedge her credit risk exposure to the issuer of that bond by buying CDS protection on that counterparty. As a result we would expect CDS spreads to be similar to credit spreads observed on corporate bonds that are deliverable into the CDS contract. In fact, under some conditions, an exact arbitrage relation exists which implies that the CDS spread should equal the credit spread

[3]In the 2003 definition, the International Swap and Derivative Association (ISDA) lists six items as credit events: (1) bankruptcy, (2) failure to pay, (3) repudiation/moratorium, (4) obligation acceleration, (5) obligation default, and (6) restructuring. For more detail, see "2003 ISDA Credit Derivatives Definitions," released on 11 February 2003.

[4]See Duffie and Singleton (2003) for a detailed description.

on the deliverable corporate bond.[5] This leads to the theoretical definition of the CDS-bond basis as the CDS spread minus the corporate bond credit spread.

While the CDS spread is observable in the market, it is not obvious how to compute the appropriate corporate bond spread. As discussed by Duffie (1999) the ideal corporate bond spread would be the spread over Libor of a floating rate note with the same maturity as the CDS referenced on the same firm. In practice, this spread is often not observable as firms rarely issue floating rate notes. Instead, we have to rely on other available fixed rate corporate bond prices. Several methodologies have been proposed in the literature. Following Elizalde, Doctor, and Saltuk (2009), we adopt the par equivalent CDS (PECDS) methodology. This method, which we present for completeness in Appendix A, essentially amounts to extracting the default intensity consistent with the prices of the corporate bonds observed in the market and using the Libor swap curve as the risk-free benchmark curve. Then one can calculate the fair CDS spread consistent with the bond implied default intensity and the risk-free benchmark curve, given a standard recovery assumption. It is this theoretical bond-implied CDS spread, called the PECDS spread, that we compare to the quoted CDS spread on the same reference entity to define the CDS-bond basis:

$$Basis_i(\tau) = CDS_i(\tau) - PECDS_i(\tau), \tag{1}$$

where τ is the maturity and i indicates the reference entity. This methodology has several advantages, as reviewed in Elizalde, Doctor, and Saltuk (2009). It has also been used by previous academic studies such as Nashikkar, Subrahmanyam, and Mahanti (2011).

Another important issue for the measurement of the basis is the funding or 'risk-free' rate benchmark (Hull, Predescu, and White, 2004). Several authors have argued that the Treasury curve is not the appropriate risk-free benchmark and indeed, that it is lower than the typical funding cost an investor can achieve via collateralized borrowing.[6] In fact, Hull, Predescu,

[5]Duffie (1999) discusses the specific conditions and shows why this relation might not exactly hold in practice.

[6]Studies that document the special status of the US Treasury curve, –presumably due to its greater

and White (2004) use the basis package (a portfolio long several corporate bonds and long CDS protection) to define a risk-free asset available to any investor. They argue that since the average CDS-bond basis is zero when measuring funding cost using swap rate minus 10 bps and the CDS-bond basis exhibits little cross-sectional variation, this is evidence that the 'true' shadow risk-free rate for a typical investor is around swap minus 10 bps (or approximately Treasury plus 50 bps).

We emphasize that the very large cross-sectional variation in the basis (across rating categories) documented in Figure 1 allows us to immediately dismiss the fact that mis-measurement of the risk-free rate benchmark is the explanation for the puzzling behavior of the CDS-bond basis during the crisis. If we were simply mis-measuring the risk-free benchmark we would observe an approximately constant CDS-Bond basis across firms reflecting the spread between our benchmark risk-free curve and the true (unobserved) risk-free curve. Let us stress, however, that since we do not observe the true risk-free benchmark curve, it is important to focus on the cross-sectional variation in the basis, rather than focusing on the average level, which could be affected by the 'flight-to-quality' effect documented in the Treasury and swap literature.

When the basis is positive, the credit default swap spread is larger than the bond spread. An investor could then short the bond and sell CDS protection to capture the basis. When the basis is negative, the credit default swap spread is lower than the bond spread. By buying the bond and buying CDS protection, investors could lock in a risk-free annuity equal to the absolute value of the basis.

As discussed in the introduction, during normal times the CDS-bond basis tends to be very small and, if anything, slightly positive. This has been studied extensively by Blanco, Brennan, and Marsh (2005), Longstaff, Mithal, and Neis (2005), and recently by Nashikkar, Subrahmanyam, and Mahanti (2011). However, Figure 1 reveals that the CDS-bond basis was significantly and persistently negative during the financial crisis, and even so after the

liquidity– include Longstaff (2004), Feldhütter and Lando (2008) among others.

crisis. Furthermore, there was substantial cross-sectional variation in the negative basis as we can see from the conspicuous difference in the bases between IG and HY bonds.

While a positive basis can often be traced back to some inability to implement the 'arbitrage' trade because either bonds are difficult to short (see liquidity concern in Nashikkar, Subrahmanyam, and Mahanti (2011)), or there exists a cheapest-to-deliver option (see Blanco, Brennan, and Marsh (2005)), a negative basis is harder to explain. Indeed, in the negative basis case, the 'arbitrage' trade requires buying the bond, financing its purchase, and buying protection to hedge against the default event. Figure 1 suggests that the return to the 'negative basis' trade would have been between 243 bps and 560 bps for IG and HY bonds respectively. These seem like very high arbitrage profits. So it is important to review the details of such a basis trade implementation to better understand where the 'limits to arbitrage' may arise.

2.1 Negative Basis Trade

In practice, there are several reasons why a negative basis trade is not a pure arbitrage. These risks are discussed in detail in Elizalde, Doctor, and Saltuk (2009) (see, in particular, their Table 2 on page 23). The main issues when implementing a negative basis trade have to do with funding risk, sizing the long CDS position, liquidity risk, and counterparty risk.

Suppose we find a bond with negative basis that trades at a price B below its notional of N. A negative basis trade requires buying the bond. The purchase is funded via the repo market where investors face a haircut h. This effectively implies that arbitrageurs will have to provide hB dollars of 'risk-capital' funded at $libor + f$ where f is the funding spread over Libor faced by the arbitrageur. The repo contract is typically overnight (up to a few months at most) with an agreed-upon repo rate, and needs to be rolled over repeatedly until the maturity of the basis trade, which is the lesser of default and maturity (e.g., 5 years).

At the same time, the investor buys the protection in the CDS market to offset the default risk. A question arises as to how to size the CDS position. A conservative approach from a viewpoint of minimizing exposure to a jump to default is to buy protection on the full notional N of the bond.

Market participants typically prefer to buy less protection to improve the carry profile of the trade (pay less in insurance premium). The justification is that the maximum capital at risk in the transaction is the initial purchase price of B.[7] In fact, a customary approach is to make an assumption about recovery (for example, assume that in case of bankruptcy a fraction R of the notional of the bond is recovered) and buy protection on a CDS notional of N_{CDS} so as to cover the loss in capital, i.e., such that $B - N\,R = N_{CDS}(1 - R)$. This will increase the carry of the trade (since the CDS premia are now reduced), but expose the investor to a jump to default in case the recovery is smaller than expected. An alternative approach is to choose the notional of the CDS position to match the spread duration on the risky bond (this approach tries to minimize mark-to-market differences between the bond and CDS position over the life of the bond as opposed to thinking about the jump-to-default risk). As explained in Duffie (1999), there is no perfect arbitrage when the underlying bond is not a floating rate note with the same maturity as the CDS contract.[8]

For illustration, suppose the investor buys protection on a notional N_{CDS}. This requires a margin payment of M and periodic mark-to-market margin calls. The margin has to be funded at $Libor + f$.[9]

After one day, the profit or loss (P&L) on the trade can be written as:

$$\begin{aligned} P\&L(t+1) &= B^{\text{bid}}_{t+1} - B^{\text{ask}}_{t} + N_{CDS} D_{CDS} * (CDS^{\text{bid}}_{t+1} - CDS^{\text{ask}}_{t}) \\ &\quad - B^{\text{ask}}_{t} * [h(Libor + f) + (1-h) * (Repo)] - M_t(Libor + f), \end{aligned} \tag{2}$$

where D_{CDS} is the duration of the CDS such that the P&L on the CDS is the product of the duration with the change in CDS rate—note that if CDS increases, the short-credit/long-

[7]For bonds that trade at a premium one may in fact buy more protection than the nominal!

[8]In fact, even in that case the arbitrage is not perfect (e.g, Lando (2004)).

[9]We assume conservatively that the investor does not earn any interest on the posted margin. If the investor was paid interest on the Margin then the margin would be funded at $Libor + f_m$ where f_m would be the funding spread net of the interest earned on the posted margin.

protection position makes money. In the P&L analysis, we explicitly consider trading costs in the form of bid-ask spreads on the bond and the CDS contract. For illustration, suppose we size our position in the CDS contract to match the Libor-spread duration on the corporate bond, then we can rewrite the P&L as: [10]

$$\begin{aligned} P\&L(t+1) &\approx D_B * \Delta Basis_t - D_B \Delta BAS_t \\ &\quad - B_t^{ask} * [h(Libor + f) + (1-h) * (Repo)] - M_t(Libor + f). \end{aligned} \tag{3}$$

BAS is the average of the bond yield bid-ask spread and the CDS bid-ask spread. Specifically, this relation shows that a typical basis trade, when rolled over repeatedly, is exposed to:

- An increase in funding cost as measured by the benchmark Libor rate.
- An increase in the arbitrageur's own credit risk, which would lead to a larger markup (f). We note that if the arbitrageur has a large position in basis trades, then this could be tied to the basis becoming more negative (i.e., the trade running away from him).
- A worsening of the collateral quality of the bond, which would lead to an increase in the haircut (h) and the repo rate ($Repo$)
- An increase in the margin requirements on the CDS position (M_t).
- An increase in trading costs as measured by the average bid-ask spread on the bond and CDS. This component is important for two reasons. First, it affects the daily mark-to-market positions. Second, if the position does need to be unwound before the maturity of the contract (or default), trading liquidity matters.

[10] We use the approximation $\Delta B^{\text{mid}} \approx -D_B(\Delta Libor + \Delta YS^{\text{mid}})$, where YS is the bond yield credit spreads. We size the position in the CDS so that $N_{CDS} D_{CDS} = D_B$. Further, we define the CDS bid-ask spread $BAS_t^C = 2(CDS_t^{\text{ask}} - CDS^{\text{mid}}) = (CDS^{\text{ask}} - CDS^{\text{bid}})$. Then we assume $B_t^{\text{ask}} - B_t^{\text{bid}} \approx D_B * BAS_t^B$. Therefore $B_{t+1}^{\text{bid}} - B_t^{\text{ask}} = -D_B(\Delta Libor + \Delta YS^{\text{mid}} + 0.5 \Delta BAS_t^B)$. Putting all together we get the above expression with $BAS_t = 0.5(BAS_t^B + BAS_t^C)$ is the average of the bid-ask spread on the bond and on the CDS.

Finally, the trade is also affected by counterparty risk in the sense that if a default on a bond occurs at time $(\tau)_B$, then the P&L will be:

$$P\&L(\tau_B) = R\,N + N_{CDS}(1-R)\mathbf{1}_{\tau_C > \tau_B}, \tag{4}$$

where τ_C denotes the default time of the counterparty selling protection and R denotes the realized recovery on the bond. Specifically, if the counterparty defaults (or has defaulted) when the underlying firm defaults then the CDS protection expires worthless. This highlights the fact that from an ex ante perspective counterparty risk depends on the correlation between the default risk of the underlying name and the counterparty selling the protection, which is typically a large bank such as J.P. Morgan, Lehman Brothers, and Goldman Sachs.

It is important to stress that counterparty risk is typically viewed as likely to be small, since if the counterparty defaults prior to the default event (i.e., $\tau_C < \tau_B$) and if the marking to market were perfect, then the investor could reopen a new position at no cost with another counterparty. Thus, in theory, counterparty risk only affects the investor if the counterparty defaults on the exact same day as the underlying bond ($\tau_C = \tau_B$). In practice however, it is likely that the failure of the counterparty, especially during an extraordinary period like the financial crisis, would be associated with more substantial costs and risks for the investor. These losses would typically be related to the likely mark-to-market loss in the position on the day of the counterparty default as well as more technical considerations, which have to do with the specific bankruptcy provisions in the ISDA covering the CDS trade (e.g., if the mark-to-market limits were insufficient, or if the collateral posted with the counterparty was rehypothecated, or if the cash settlement upon bankruptcy of the counterparty is based on mid-market quotes).

Below we try to use the cross-sectional variation in individual bond bases to disentangle the effects of various risks outlined above that affect the risk-return tradeoff of a basis trade. Our working hypothesis is that an arbitrageur having limited access to capital will try to exploit the basis trade opportunities that offer the best expected return per unit of risk capital. So she will choose basis trades that have the most negative basis (highest expected return) but controlling for ex ante measures of exposure to market and funding liquidity. All else equal she will prefer basis trades on bonds with low haircuts, low exposure to funding cost (in the sense that for two bonds with equally negative basis, the one which correlates more with funding costs is more attractive, since the basis trade converges when funding costs rise), low counterparty risk (in the sense that the probability of the underlying firm defaulting at the same time as the counterparty in the CDS is lower) and low trading liquidity risk (i.e., lower current transaction costs and lower future risk-adjusted transaction costs). If this hypothesis is correct then we expect that the risk characteristics of the basis trade (counterparty risk, funding risk, liquidity risk, collateral quality) should be related to the cross-sectional variation of the bases.[11]

3 Data

The data used to study the CDS-bond basis come from several sources. We start with the universe of firms whose single-name CDS is traded in the derivative market and transactions are recorded in the Markit database. Then we identify corporate bonds issued by these firms from the Mergent Fixed Income database and collect bond characteristics. We further download corporate bond transaction records from the enhanced version of the Trade Reporting and Compliance Engine (TRACE). Finally we match each firm's credit default swap and bond spread to corresponding equity returns in the Center for Research in Security Prices (CRSP). All data are in daily frequency from July 1, 2006 through December 30, 2014. The whole sample is further partitioned into four phases: Phase 1 is the period before the subprime credit crisis, named 'Before Crisis' (7/1/2006 - 6/30/2007);[12] Phase 2 is the period between the subprime credit crisis and the bankruptcy of Lehman Brothers, called 'Crisis I' (7/1/2007

[11]A more sophisticated analysis would be to solve the optimal capital allocation decision of the arbitrageur to the available basis trades and test her first order condition.

[12]There is not a unanimously agreed day for the beginning of the subprime crisis. Popular opinion is that the subprime crisis started in August 2007. Here we take a conservative stance by starting the crisis period in July 2007.

- 8/31/2008); Phase 3 is the period after Lehman Brothers' failure, 'Crisis II' (9/1/2008 - 9/30/2009);[13] and Phase 4 is the period after the financial crisis, 'Post-Crisis' (10/1/2009 - 12/30/2014).

Our goal is to examine the arbitrage relationship between the corporate bond cash and derivative market. The cornerstone of our analysis is an accurate measure of the CDS-bond basis. As shown in Section 2 and Appendix A, the basis is constructed from the credit default swap spread, the corporate bond transaction price, and the reference interest rate.

3.1 Corporate Bond

Corporate bond transaction data are downloaded from Trace. Bond characteristics are collected from Mergent Fixed Income Databases including coupon, rating, interest rate frequency, option features, and so on. For bond-level rating information, if a bond is rated only by Moody's or by Standard and Poor's, we use that rating. If a bond is rated by both rating agencies, we take the average rating as the final one.

We highlight the following filtering criteria in order to choose qualified bonds. First, we remove bonds that are not listed or traded in the U.S. public market, which include bonds issued through private placement, bonds issued under the 144A rule, bonds that do not trade in US dollars, and bond issuers not in the jurisdiction of the United States. Second, we focus on corporate bonds that are not structured notes, not mortgage backed or asset backed. We also remove the bonds that are agency-backed or equity-linked. Third, we exclude convertible bonds since this option feature distorts the basis calculation and makes it impossible to compare the basis of convertible and non-convertible bonds.[14] Fourth, we remove bonds with floating rate; that means the sample comprises only bonds with fixed or zero coupon. This rule is applied based on the consideration of the accuracy in the basis calculation, given the

[13] It's unclear of the earliest ending date of the U.S. financial crisis. According to NBER business cycle dates, the recession ended by June 2009. We allow for another three months in the Crisis II period.

[14] Bonds also contain other option features such as putable, redeemable/callable, exchangeable, and fungible. Except callable bonds, bonds with other option features are relatively a small portion in the sample. However, callable bonds constitute about 67% of the whole sample. Hence, we keep the callable bonds in our final sample, but we also conduct robustness check for a smaller sample filtering out the bonds with option features.

challenge in tracking floating-coupon bond's cash flows.

The Enhanced TRACE provides the information on bond transactions at the intraday frequency. Beyond the above filtering criteria, we further clean up TRACE transaction records by eliminating when-issued bonds, locked-in bonds, and bonds with commission trading, special prices, or special sales conditions. We remove transaction records that are cancelled, and adjust records that are subsequently corrected or reversed. Bond trades with more than 2-day settlement are also removed from our sample.

Lastly, we focus on bonds which have 3 to 7.5 years remaining to maturity (time-to-maturity is measured each day during the sample period).[15] This criterion is due to the concern that in the CDS market which we introduce in the next subsection, the five-year CDS contracts often have the best liquidity. In order to match the five-year term, we limit the calculation of the CDS-bond basis only for bonds with time-to-maturity ranging from 3 to 7.5 years.

3.2 Credit Default Swap

We download single-name credit default swap data from Markit Inc. for U.S. firms. The prices are quoted in basis points per annum for a notional value of $10 million and are based on the standard ISDA contract for physical settlement. The original dataset provides daily market CDS prices in various currencies and different types of restructuring documentation clauses. Following a conventional rule, we choose the CDS price in US dollar and the documentation clause type as 'Modified Restructuring' (MR).[16]

There are several caveats in matching CDS with underlying reference entities. First, the

[15] The restructuring rule used in our CDS contract, 'Modified Restructuring', restricts the deliverable bonds to be within 30 months of the contract maturity. Therefore we consider a qualified deliverable bond have time-to-maturity up to 7.5 years, which is 30 months later than the 5-year CDS contract, the target maturity in this paper.

[16] Under the 2003 Credit Definitions by the International Swap and Derivative Association (ISDA), there are four types of restructuring clauses: Cumulative Restructuring (CR), Modified Restructuring (MR), Modified-Modified Restructuring (MM), and No Restructuring (XR). 'Modified Restructuring' is used by most broker-dealers in the U.S. market. This convention holds till April 8, 2009. Afterwards the U.S. market adopts the 'No Restructuring' convention. For consistency, we choose the MR documentation clause throughout our sample.

MR rule restricts the deliverable bonds to be within 30 months of the contract maturity. We check the bond maturity to follow the MR rule. Second, the underlying bond should be deliverable into the CDS contract. However, it is very difficult to identify whether bonds are deliverable into CDS contracts for a large sample of firms over a long time period, "since CDS conventions are often bilaterally defined in the over-the-counter market," as pointed out by Nashikkar, Subrahmanyam, and Mahanti (2011).

The original dataset provides a CDS spread term structure incorporating maturities of 1y, 2y, 3y, 4y, 5y, 7y, and 10y. We use all maturities in conjunction with matching interest rate swaps to calculate a term structure of default probability, which is an integral component in deriving the bond-implied CDS spread (PECDS) and hence the CDS-bond basis (see Appendix A). In the end we focus on the CDS-bond basis with a maturity of five years, because the 5-year CDS is by far the most liquid in the credit derivative market, and this is also the one mostly used in the literature.

The CDS data and corporate bond data are manually matched via company names. There are originally 1842 unique CDS underlying entities during 2006 - 2014. After matching to corporate bond data and applying the aforementioned filtering criteria, we finally have a total of 1.1 million daily observations during July 1st, 2006 to December 31st, 2014, representing 4415 pairs of CDS-bond basis for 679 unique CDS names. To reduce the influence of outliers in the cross-sectional regressions, we further winsorize the bases at the 0.5% and 99.5% level.[17]

3.3 Reference Rate

We use the U.S. dollar interest rates swaps as reference rates for the risk-free funding curve for computing credit spreads. An alternative choice might be to use government bond yields. However, as Blanco, Brennan and Marsh (2005) point out, "government bonds are no longer an ideal proxy for the unobservable risk-free rate" due to tax treatment, repo specialness, legal constraints, and other factors. Importantly, the Libor swap rate represents a better indicator

[17]Our empirical results in Section 5 are similar if winsorizing at the 1% and 99% level, or at the 0.25% and 99.75% level.

of the funding cost for financial intermediaries and typical basis swap traders than the Treasury curve. Therefore we use it as our benchmark funding curve for the basis calculations.[18]

As discussed in Section 2, we focus on the cross-sectional variation in the CDS-bond basis rather than its absolute level since we do not observe the true risk-free reference rate.

3.4 Summary Statistics

Table 1 presents summary statistics of the CDS-bond basis. The basis across all firms was slightly negative before the crisis, −10 bps on average between 7/1/2006 to 6/30/2007 (which is consistent with the evidence in Hull, Predescu, and White (2004)), but fell to −118 bps in the first phase of the financial crisis and to −324 bps after the bankruptcy of Lehman Brothers. The average basis remained negative around −137 bps after the financial crisis. Meanwhile the volatility of the basis kept increasing for all types of firms from an average of 59 bps before the crisis to 192 bps and further to 369 bps during the turmoil of the financial crisis. The volatility fell back to 152 bps after the crisis, still far away from the pre-crisis level. Firms with both the IG and the HY ratings share the same pattern as firms overall, whose bases became more and more negative and volatile as the financial crisis progressed. Moreover, the basis of HY firms is always more volatile than that of IG firms. Financial firms and non-financial firms share similar time-series patterns in their bases, except that financial firms have slightly higher volatility of basis during the peak of the financial crisis.

Table 1 also provides additional basis results across ratings from AAA/AA to CCC. Firms with lower credit rating tend to have more negative and more volatile bases during the crisis while the bases display a right-skewed 'smile' from AAA/AA to CCC before and after the crisis.

Figure 1 provides an illustration of the basis dynamics for IG and HY bonds. The solid blue line is the median value of the aggregated CDS-bond bases for bonds in each rating category, weighted by bond outstanding amounts. The dotted red lines are the 10th and 90th

[18]See also the swap-Treasury spread discussions in Hull, Predescu, and White (2004), Collin-Dufresne and Solnik (2001), and Longstaff, Mithal, and Neis (2005).

percentile of the aggregated bases. It is worth noting that the average CDS-bond basis for both IG and HY bonds after the financial crisis, though improved, are still far below their pre-crisis levels. Moreover, there exists a large dispersion of the basis values throughout the whole sample period.

We focus the analysis on the cross-sectional variation in the negative basis, because the basis sample became predominantly negative. As shown in the table below, 92.7% of the observations in the whole sample have negative bases. The proportion is even larger during the crisis period and the post-crisis period for about 94.4% and 93.9% respectively.

	Negative	Positive
Before Crisis	69.3%	30.7%
Crisis I	87.6%	12.4%
Crisis II	94.4%	5.6%
Post Crisis	93.9%	6.1%
Total	92.7%	7.3%

3.5 Preliminary Evidence on the Cross-sectional Variation in the Basis

Garleanu and Pedersen (2011) make the point that haircuts are typically around 25% for IG firms (and very similar across firms rated from AAA to BBB) and of the order of 55% for HY bonds (rated BB or lower). In their model, the basis differential between IG and HY bonds should be equal to the difference between haircut margins multiplied by the collateral funding spread (i.e., the difference between the collateralized and the uncollateralized funding rates). While indeed an important plausible determinant of the basis, our data suggests that there exist additional important factors. Indeed, as clearly shown in Table 1 there is tremendous amount of variation in the basis within a credit rating category, and certainly a lot of differences in the basis within the IG and the HY category.[19]

[19]Overtime, there are also a lot of variations in the basis in a way that cannot solely be explained by the variation in the collateral funding spread, and as we argue below, is unlikely to be explained solely by changes in haircuts.

Firm	Number of Days with Positive Basis		Rating	Industry
	Crisis I ($T = 295$)	Crisis II ($T = 271$)		
Newmont Mng Corp	286	250	BBB	Basic Materials
Berkshire Hathaway	127	244	AAA	Financials
Amern Tower Corp	237	226	BB	Technology
Emc Corp	259	188	BBB	Technology
MetLife Insurance Co	12	178	A	Financials
Boyd Gaming Corp	253	163	BB	Consumer Services
General Electric Co	89	154	AAA	Industrials
Windstream Corp	54	131	BB	Telecommunications
Penn Natl Gaming Inc	134	130	B	Consumer Services
Mylan Inc	204	122	BB	Health Care
AutoNation Inc	1	117	BB	Consumer Services
Las Vegas Sands Corp	108	106	B	Consumer Services

Note: Ratings are based on the values at the end of September 2008.

To illustrate this point even more dramatically, we present examples for twelve firms in our sample that have positive basis for more than 100 days during the second phase of the crisis (9/1/2008 - 9/30/2009, with 271 days) . These firms have diverse credit ratings ranging from B (Las Vegas Sands Corp and Penn Natl Gaming Inc) to AAA (Berkshire Hathaway, and GE), and belong to six separate industries. This is clearly at variance with a model that would have a single factor, such as haircuts or margins, to explain the basis.[20] Clearly, the haircut and margin requirement on Las Vegas Sands were much larger than for Berkshire bonds, and yet both display a positive basis (when most IG and HY bonds displayed strongly negative bases at the time). Some factors driving the individual basis are likely to be highly idiosyncratic. For example, it was suggested to us that the very positive basis on Berkshire was due to the large demand from CDS protection buyers by dealers who had non-mark-to-market in-the-money long-term volatility exposures to Berkshire. This would have driven the CDS on Berkshire up relative to the bond yield generating the positive basis.

[20]Indeed, the general model in Garleanu and Pedersen (2011) predicts that other factors (such as the covariance of the underlying cash-flows with aggregate consumption) in addition to the margin differential should predict the difference in the basis. It is only for the specific application to the CDS basis that they focus on the margin difference. Our data suggests it is important to look for additional factors.

We focus more systematically on the cross-sectional variation in the CDS-bond basis below.

4 Limit-to-Arbitrage Factors

The deviation between the same underlying bond's transaction price and its CDS spread, according to Section 2, could be driven by four types of factors: i) the collateral value of the underlying corporate bond, ii) the liquidity of bond trading, iii) the counterparty risk in buying a CDS contract, and iv) the funding constraints faced by investors who want to explore the arbitrage. In this section we construct various proxies for these factors and discuss their likely relationship to the cross-section of the CDS-bond bases.

4.1 Collateral Quality Proxy

In the basis trade, the purchase of a bond is funded via the repo market and two key variables determine the cost of the purchase: haircut and repo rate, which we label as the proxies of collateral quality. Haircuts of corporate bonds in the repo market often have narrow range in the investment-grade or non-investment-grade categories. That is, bonds have similar haircuts across ratings from AAA to BBB, or across ratings from BB and CCC. Garleanu and Pedersen (2011) make the point that haircuts are typically around 25% for most investment-grade firms and around 55% for most non-investment-grade firms. This rough haircut convention thus cannot help distinguish the wide range of the bases as shown in Figure 1. Given that haircut is dominantly driven by a bond's credit risk, we thus use credit rating as one proxy for collateral quality, noted as *Rat*. In detail, we assign numeric values for bond ratings ranging from 1 for CCC, 2 for CC−, all the way to 21 for AAA rated bonds, thus we have a granularity in capturing the collateral quality of bonds. The higher the rating indicator, the higher the expected collateral quality.

An alternative proxy for the collateral value is the repo rate, but the bond-level data on the repo rates are not publicly available. We thus rely on the bond lending fee as a second proxy of collateral value based on a unique corporate bond loan data. The data, provided by Markit, record daily corporate bond loan transactions where beneficial owners such as insurance companies, pension funds, and other institutional investors lend out bonds to end-users such as hedge funds and collect lending fees. We collect the real borrowing cost for each bond in our sample, *Fee*, as the transaction value-weighted average lending fees over all open transactions. The lending fee is typically negatively correlated with the repo rate of a bond and thus positively related to the bond's collateral quality. It is a source of additional return for bondholders. On the other hand, a high lending fee signals a higher shorting demand for a particular bond, thus it can also signal a negative view on the bond by market participants (see Bai (2018)).

To do the negative basis trade, an arbitrageur needs to buy bonds that are funded via the repo market using the same bonds as collateral. The haircut and the repo rate imposed on the repo transaction reduce the amount of leverage available to the arbitrageur. All else equal, we expect bonds with lower collateral quality, that is, a higher haircut or a lower lending fee to have a less profitable basis trade per unit use of expected risk capital. Thus, the lower the collateral quality the more negative the basis to equalize expected returns per unit of risk capital. So we expect a positive coefficient in cross-sectional regressions of the bases on collateral quality.

4.2 Bond Trading Liquidity

As the basis trade P&L analysis in Section 2.1 reveals, individual bond (and CDS) trading costs affect the profitability of the CDS-bond basis trade. Therefore, all else equal, arbitrageurs will seek basis trades with bonds that are more liquid and that have less (trading) liquidity risk in the sense that they are less likely to become illiquid when the basis trade further diverges. Recent studies, such as Bao, Pan, and Wang (2011), Dick-Nielsen, Feldhütter, and Lando (2012), find severe deterioration of liquidity in the corporate bond market during the financial crisis. In this section we construct three different measures of bond liquidity and liquidity risks, following the decomposition method proposed in Acharya and Pedersen (2005).

1. Bond illiquidity level (noted as Liq): measured by the Bao, Pan, and Wang (2011) gamma (L^{BPW}), which aims to extract the transitory component in the bond price.[21] Specifically, let $\Delta p_t = p_t - p_{t-1}$ be the daily log price change from $t-1$ to t. The L^{BPW} is defined as

$$L^{BPW} = -Cov_t(\Delta p_t, \Delta p_{t-1}). \tag{5}$$

   We compute L^{BPW} for each bond per month and assign the monthly value for weeks within that particular month. We expect bonds that are more illiquid to have more negative basis, that is a negative coefficient of the bases on the Liq measure in cross-sectional regressions.

2. Bond iliquidity beta (noted as β_{Liq}): the co-movement between bond i's illiquidity and the market illiquidity, where the market illiquidity is calculated as the average of individual bond illiquidity in the sample. As explained in Acharya and Pedersen (2005), investors want to be compensated for holding a security that becomes illiquid when the market in general becomes illiquid. Similarly, we expect basis trade arbitrageurs to prefer (all else equal) basis trades with bonds whose trading costs co-vary less with the overall bond market illiquidity. Hence, we expect a negative coefficient in the cross-sectional regression of bases on bond liquidity betas.

3. Bond liquidity market beta (noted as $\beta_{Liq,M}$): the co-movement between bond i's illiquidity and the market return, where the market return is measured by the CRSP value-weighted stock market return. As in Acharya and Pedersen (2005) we expect arbitrageurs to prefer at the margin negative basis trades for bonds that tend to have trading costs that co-vary more with market returns. Indeed, bonds with high trading

[21]There is no agreement on the best measure for corporate bond liquidity. Thus we also construct alternative liquidity measures such as the bid-ask price spread and the Amihud (2002) measure. These measures generate similar empirical results.

costs in market down turns potentially lead to lower profits upon an unwind precisely in bad states. Therefore we expect (all else equal) a less negative basis on bonds with larger bond liquidity market betas, that is, a positive regression coefficient.

The three measures provide a characterization of bond liquidity risk. Intuition suggests that the three measures should be correlated. In Table 2, we show that the correlation between liquidity (risk) measures is low during the non-crisis period, ranging from 2% to 8% in the absolute value, however the correlations increase to 27∼39% during the crisis period.

4.3 Counterparty Risk

Counterparty risk has become a primary concern facing participants in the financial markets during the 2007−2009 crisis. Counterparty risk is the risk that the protection seller, typically a broker dealer, cannot make good on its commitment to the protection buyer in case of default. Therefore counterparty risk should make the insurance less valuable and lower the CDS spread, possibly contributing to the negative basis. As explained previously, the higher the correlation between the default events of the underlying entity and the protection seller the bigger the expected counterparty risk.[22] The challenge is how to measure the correlation between the default risk of the underlying entity and the counterparty selling the CDS protection.[23]

The CDS market is over-the-counter and the exact nature of counterparties is not known. Furthermore, the process of netting makes it difficult to establish an aggregate measure of counterparty risk for individual reference entities.[24]

[22] That counterparty risk is not irrelevant, can be seen from the Lehman Brothers case. Suppose an investor had bought protection on Washington Mutual from Lehman Brothers. Washington Mutual defaulted only a few days after Lehman. Without marking to market, the investor would be a regular claimant in bankruptcy for the protection purchased from Lehman, leading to at best a partial loss. Of course, if ISDA agreements were well enforced, and provided the investor had negotiated full-two-way mark to market with Lehman, then the risk would be further mitigated. However, in practice, it is likely that most funds would have ended with at least some partial loss as a result of this double default.

[23] Arora, Gandhi, and Longstaff (2012) look for counterparty fixed effects using a proprietary data set of CDS transaction prices by 14 CDS dealers selling credit protection on one underlying firm to identify the counterparty risk component of CDS spreads.

[24] In September 2008 the bankruptcy of Lehman Brothers caused almost $400 billion to become payable to the buyers of CDS protection referenced against the insolvent bank. However the net amount that changed hands was around $7.2 billion. This difference is due to the process of "netting". Market participants

To establish a measure of the counterparty risk faced by an investor engaging into a specific basis trade, we notice that throughout our sample, the primary dealers on average trade more than 90% of total transaction dollar volume. It seems therefore reasonable to construct a counterparty risk measure for a representative CDS issuer using the list of primary dealers designated by the Federal Reserve Bank of New York.[25] These primary dealers are banks and security broker-dealers that trade in the U.S. government securities with the Federal Reserve System. To become qualified as a primary dealer, a firm must be in compliance with capital standards under the Basel Capital Accord, with at least $100 million of Tier I capital for a bank or above $50 million of regulatory capital for a broker-dealer. As trading partners of the central bank, these primary dealers often are the biggest and most competitive financial institutions who happen to be dominant issuers of credit default swap contracts. As of September 2008, there were 19 primary dealers such as Citigroup, J.P. Morgan Chase, and Goldman Sachs (the complete list is documented in Appendix B). The list changes over time since some primary dealers may fail to meet required capital standards. Accordingly, we update the components of the primary dealer index. For example, the index includes Lehman Brothers' Holdings before its bankruptcy on September 15, 2008, but exclude it afterwards and adds Nomura Securities International, Inc. starting from July 27, 2009.

For the primary dealer index, we calculate its CDS spread weighted by each constituent's market capitalization. As shown in Figure 2, the primary dealers have a strikingly increase in their default risk during the financial crisis of 2008−2009 and the peak of the European sovereign debt crisis of 2011−2012. A direct implication is that the insurance contracts sold by these dealers should be less valuable and the protection buyers face higher counterparty risk. We measure an underlying entity's counterparty risk as the beta coefficient of regressing

cooperated so that CDS sellers were allowed to deduct from their payouts the funds due to them from their hedging positions. Dealers generally attempt to remain risk-neutral so that their losses and gains after big events will on the whole offset each other.

[25]Data source: `http://www.newyorkfed.org/markets/pridealers_current.html`.

the change of firm-level CDS to the change of the primary dealer CDS:

$$\beta_{i,cp} = \frac{cov(\Delta CDS_i, \Delta CDS_{index})}{var(\Delta CDS_{index})}. \tag{6}$$

The higher the β_{cp}, the larger the likelihood of a joint default and the less valuable we expect the protection to be when purchased from that counterparty. So we expect a negative coefficient in the cross-sectional regression of the bases on counterparty betas.

4.4 Funding Liquidity Risk

For an arbitrageur entering a basis trade, the risk is that the basis becomes more negative at the same time as her funding costs widen. We thus proxy funding liquidity risk by the regression coefficient of the change in the basis on a measure of the change in funding costs or funding liquidity premium.

The literature has considered many proxies to measure the funding liquidity premium including the Libor-OIS spread, the TED spread (Libor-TBill), the Repo-TBill spread and the OIS-TBill spread. One concern with the Libor-indexed spreads is that they are contaminated by financial intermediary credit risk and hence might be correlated with counterparty risk, especially during the financial crisis. For this reason, we follow Nagel (2016) and use the three-month Repo-TBill spread. Bai, Krishnamurthy, and Weymuller (2018) propose an alternative measure of the OIS-TBill spread. Both the Repo-TBill spread and OIS-TBill spread have similar time-series patterns, and they have a correlation value of 0.90. All our empirical results remain unchanged in magnitude and significance if we use the OIS-TBill spread as the alternative proxy of funding liquidity premium.

Our estimate of the funding liquidity beta is therefore defined as:

$$\beta_{i,fl} = \frac{cov(\Delta Basis_i, \Delta \text{Repo-TBill})}{var(\Delta \text{Repo-TBill})}. \tag{7}$$

The lower the funding liquidity beta, the less aggressively an arbitrageur would invest in that basis trade, as the basis will become more negative when her funding cost increases. So we expect a positive coefficient in the cross-sectional regression of the bases on funding liquidity betas.

4.5 Summary

We summarize the expected signs of the cross-sectional determinants of the CDS-bond basis in the following table based on our previous discussion.

Factors	Expected cross-sectional correlation with negative basis
Lending fee	+
Credit rating	+
Bond illiquidity	−
Bond liquidity risk	−
Bond liquidity mkt risk	+
Counterparty Risk	−
Funding liquidity risk	+

Panel B of Table 2 presents the correlation values of limit-to-arbitrage factors. In the non-crisis period, the correlations among collateral quality proxies, bond trading liquidity, counterparty risk, and funding liquidity risk are trivial. For example, funding liquidity beta is almost uncorrelated with any other factors with the correlation values between -0.01 to 0.03; bond illiquidity or liquidity risk also have negligible correlation with other factors. In the crisis period, the correlation between bond illiquidity and bond liquidity risk increases, but the correlation values among other risk factors remain small. The low correlations relieve the concern of colinearity across limit-to-arbitrage risk factors.

5 Cross-sectional Determinants of the CDS-Bond Basis

We now investigate the cross-sectional variation in the CDS-bond basis using Fama and MacBeth (1973) regressions. To mitigate the impact of noise in the daily data, we adopt the weekly frequency throughout our empirical analysis, where the bond-level weekly basis is the average value of daily basis within each week for a particular bond. For limit-to-arbitrage factors except collateral quality, all beta coefficients are estimated for each bond at each week using a rolling window of past 60-week observations.[26,27] For collateral quality we calculate the weekly averages of the daily corporate bond loan lending fee and of its credit rating.

5.1 Bond-level Results

We report the Fama-MacBeth regression results for four subperiods: Before Crisis, Crisis I, Crisis II, and Post Crisis. In detail, weekly cross-sectional regressions are run for the following specification and nested versions thereof:

$$\begin{aligned} Basis_{i,t} &= \alpha_t + \gamma_{t,Rat}Rat_{i,t} + \gamma_{t,Fee}Fee_{i,t} + \gamma_{t,Liq}Liq_{i,t} + \gamma_{t,Bliq}\beta_{i,t,Liq} + \gamma_{t,BliqM}\beta_{i,t,LiqM} \\ &\quad + \gamma_{t,CP}\beta_{i,t,CP} + \gamma_{t,FL}\beta_{i,t,FL} + \varepsilon_{i,t}, \quad \forall t \end{aligned} \tag{8}$$

where $Basis_{i,t}$ is the CDS-bond basis on bond i in week t. Rat is the bond credit rating in numeric values, Fee is the bond lending fee, Liq is the bond trading illiquidity level, β_{Liq} denotes bond liquidity risk, β_{LiqM} denotes bond liquidity market risk, β_{CP} and β_{FL} denote the counterparty risk and funding liquidity risk, — all measures are defined in Section 4. We standardize explanatory variables cross-sectionally each week, therefore the coefficients are comparable across variables and over sample periods.

Table 3 reports the time-series average of the estimated coefficients γ_t's and the average adjusted R-squared values over the 466 weeks from July 2006 to December 2014 for the negative CDS-bond basis. The t-statistics based on Newy-West adjusted standard errors are given in squared brackets.

Focusing on the multi-variate regressions (since the results remain largely similar to the

[26] A bond is included in our sample if it has at least 24 weekly basis observations in the 60-week rolling window before the test week. Our data start from July 2001 and we report regression results since July 2006.

[27] Our results are also robust to different rolling windows in estimating limit-to-arbitrage risk factors, that is i) 36-week rolling window instead of 60 weeks using weekly data, or ii) 60-day rolling window using daily data.

univariate case) in the last column, we see that collateral quality, bond illiquidity, counterparty risk, and funding liquidity risk, all enter significantly especially during the crisis period, achieving an R^2 of 36.6%. The sign of the significant coefficients are consistent with our expectations summarized in Section 4.5. During the peak of the crisis from September 2008 to September 2009, the most important factors driving the cross-sectional differences in basis are credit rating with an estimated coefficient of 1.902 (t-stat=6.78), bond illiquidity with an estimated coefficient of -0.676 (t-stat=-3.16), funding liquidity risk with an estimated coefficient of 0.794 (t-stat=3.71), and counterparty risk with an estimated coefficient of -0.545 (t-stat=-5.22). In the first stage of the crisis from July 2007 to August 2008, the same set of factors have significant explanatory power, though to a relatively smaller degree. For example, the coefficient for credit rating drops from 1.902 in Crisis II to 0.570 in Crisis I, the coefficient for counterparty risk drops from -0.545 to -0.181, and that for funding liquidity risk drops from 0.794 to 0.132.

During the non-crisis period (both before and after the crisis) however, only credit rating remains consistently significant in explaining the cross-sectional variations of the basis, indicating that collateral quality is always relevant for a basis arbitrage trade. The estimated coefficients for counterparty risk are also statistically significant in the non-crisis period, but have much smaller magnitude (the coefficient is 0.059 before and 0.163 after the crisis). Also, the signs of the counterparty risk coefficient in the non-crisis period are positive, which is counter to our expectation; this is likely due to the strong colinearity of counterparty risk with credit rating and lending fee (the correlations are -0.31 and 0.21, respectively). Lastly, the bond liquidity factors (both illiquidity and liquidity risk) lose explanatory power in the non-crisis period.

There is one surprising finding, namely that lending fee has a significant but negative coefficient in the cross-sectional regressions throughout the crisis and non-crisis periods, which is at variance with our expectation. Our interpretation is that lending fee is also a proxy for the short interest in a bond. That is, a high lending fee not only arises when a bond is highly valued as a collateral and thus in high demand in the securities lending market (in this case, the bond has a low repo rate), but also may arise when a bond is being heavily shorted and thus is in scarce supply in the securities lending market. If the lending fee during the crisis mostly captures short interest, then the negative sign on lending fee could reflect the fact that arbitrageurs refrain from entering a basis trade on the related bond which is heavily shorted. In effect, arbitrageurs may worry about 'catching a falling knife' when entering such a basis trade, as the negative market view on the bond signaled by the high short interest suggests that the basis might diverge even further in the near future.

To conclude, the empirical model is reasonably successful in explaining the cross-sectional variation in the bases. All factors (except for lending fee) have the signs as expected, consistent with the hypothesis that the marginal investor being a leveraged hedge fund trade off risk and return when allocating scarce risk-capital to different basis investment opportunities.

5.2 Portfolio-level Results

To minimize the noise in the bond-level cross-sectional regression results, we also consider running the Fama and MacBeth (1973) regression at the portfolio level. We first examine the distribution of bond size throughout the sample and categorize bonds into four bins by the 25th, 50th, and 75th percentile. Within each bin of bond size, we further categorize bonds into another six bins by their credit ratings: AAA/AA, A, BBB, BB, B, CCC&Below. In so doing, we construct the time series of 24 corporate bond portfolios sorted by size and rating. We report in Panel A of Table 4 the average basis of the 24 portfolios. The basis is clearly monotonically increasing with bond credit quality and also monotonically decreasing with size as one might have expected.

In Panel B we report the portfolio-level cross-sectional regression result. We find a very high R^2 of around 80% for the bases of 24 portfolios. Most coefficients are statistically significant throughout the subsample periods. The signs on credit rating, bond illiquidity and liquidity risk, counterparty risk, and funding liquidity risk are also consistent with our expectation during the crisis. However, we note that outside the crisis period counterparty risk and bond illiquidity remain significant but surprisingly their signs flip likely due to the correlation with other risk factors.

6 Conclusion

We have analyzed the cross-sectional variation in the CDS-bond basis during the crisis. Focusing on the cross section of the CDS-bond basis is interesting as it provides a natural testing ground for the literature that models limits to arbitrage, and specifically the behavior of arbitrageurs with limited capital facing multiple arbitrage opportunities.

We find that, most notably during the post-Lehman crisis period, several limit-to-arbitrage measures such as collateral quality, bond illiquidity and liquidity risk, counterparty risk, and funding liquidity risk can 'explain' a significant fraction of the cross-sectional variation in the CDS-bond basis. After the crisis the explanatory power of risk measures decreases substantially, but collateral quality proxies such as credit rating and lending fee remain significant. Interestingly, throughout the sample bond lending fee remains highly negatively related to the bond basis. We interpret this as evidence for the fact that arbitrageurs refrain from entering the basis trade that have large trading frictions (i.e., that are costly to fund or to trade) associated with them and especially when there is a large amount of short interest in the bond (signaled by a high lending fee), which indicates that the basis might further diverge in the short run.

References

Acharya, V. V., Pedersen, L. H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77, 375–410.

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time series effects. Journal of Financial Markets 5, 31–56.

Arora, N., Gandhi, P., Longstaff, F. A., 2012. Counterparty credit risk and the credit default swap market. Journal of Financial Economics 103, 280–293.

Bai, J., 2018. What bond lending reveals? The role of informed demand in predicting credit spread changes, working paper, Georgetown University.

Bai, J., Krishnamurthy, A., Weymuller, C.-H., 2018. Measuring liquidity mismatch in the banking sector. Journal of Finance 73(1), 51–93.

Bao, J., Pan, J., Wang, J., 2011. Liquidity and corporate bonds. Journal of Finance 66, 911–946.

Blanco, R., Brennan, S., Marsh, I. W., 2005. An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. Journal of Finance 60, 2255–2281.

Coffey, N., Hrung, W. B., Sarkar, A., 2009. Capital constraints, counterparty risk, and deviations from covered interest rate parity, staff Reports 393, Federal Reserve Bank of New York.

Collin-Dufresne, P., Solnik, B., 2001. On the term structure of default premia in the swap and libor markets. Journal of Finance 56, 1095–1115.

D.E. Shaw Group, 2009. The basis monster that ate wall street. Market Insights March.

Dick-Nielsen, J., Feldhütter, P., Lando, D., 2012. Corporate bond liquidity before and after the onset of the subprime crisis. Journal of Financial Economics 103, 471–492.

Duffie, D., 1999. Credit swap valuation. Financial Analysts Journal 55, 73–87.

Duffie, D., Singleton, K., 2003. Credit Risk. Princeton University Press.

Elizalde, A., Doctor, S., Saltuk, Y., 2009. Bond-cds basis handbook. J.P. Morgan Credit Derivatives Research February 05.

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.

Feldhütter, P., Lando, D., 2008. Decomposing swap spreads. Journal of Financial Economics 88, 375–405.

Fleckenstein, M., Longstaff, F., Lustig, H., 2014. The tips-treasury bond puzzle. Journal of Finance 69(5), 2151–2197.

Fontana, A., 2011. The negative CDS-bond basis and convergence trading during the 2007/09 financial crisis, working paper 694, National Centre of Competence in Research Financial Valuation and Risk Management.

Garleanu, N., Pedersen, L. H., 2011. Margin-based asset pricing and deviations from the law of one price. Review of Financial Studies 24(6), 1980 – 2022.

Gromb, D., Vayanos, D., 2010. Limits of arbitrage: The state of the theory. Annual Review of Financial Economics 2, 251 – 275.

Hull, J., Predescu, M., White, A., 2004. The relationship between credit default swap spreads, bond yields, and credit rating announcements. Journal of Banking and Finance 28, 2789 – 2811.

Kim, G. H., Li, H., Zhang, W., 2016. CDS-bond basis and bond return predictability. Journal of Empirical Finance 38, 307–337.

Kim, G. H., Li, H., Zhang, W., 2017. The CDS-bond basis and the cross section of corporate bond returns. Journal of Futures Markets 17(8), 836–861.

Lando, D., 2004. Credit Risk Modeling Theory and Applications. Princeton University Press.

Longstaff, F. A., 2004. The flight-to-liquidity premium in u.s. treasury bond prices. Journal of Business 77, 511–526.

Longstaff, F. A., Mithal, S., Neis, E., 2005. Corporate yield spreads: Default risk or liquidity? new evidence from the credit default swap market. Journal of Finance 60, 2213–2253.

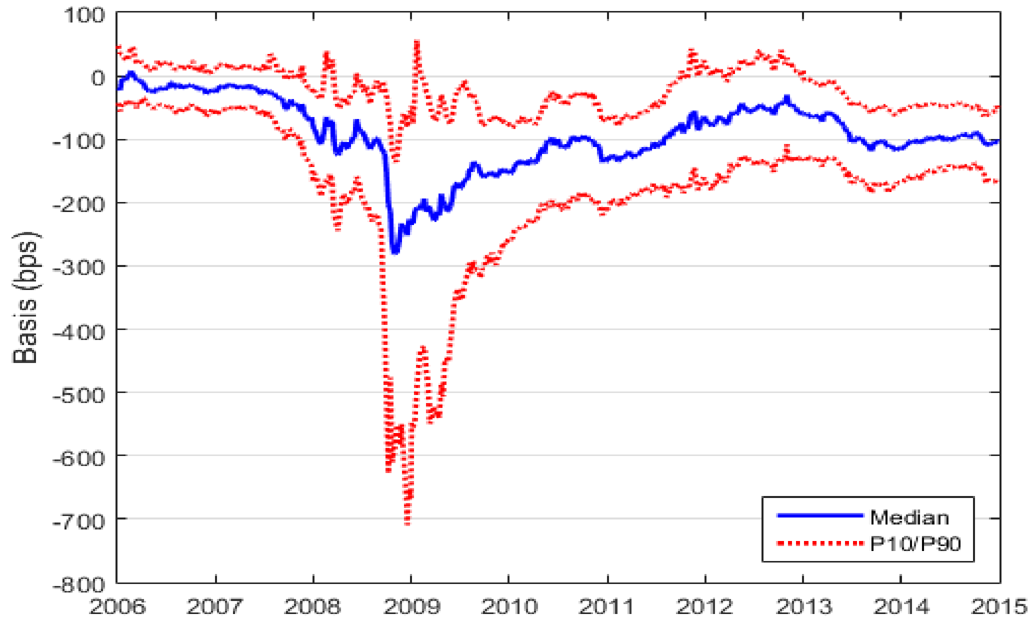
Mitchell, M. L., Pulvino, T. C., 2012. Arbitrage crashes and the speed of capital. Journal of Financial Economics 104(3), 469 – 490.

Nagel, S., 2016. The liquidity premium of near-money assets. Quarterly Journal of Economics 131, 1921–1971.

Nashikkar, A., Subrahmanyam, M. G., Mahanti, S., 2011. Liquidity and arbitrage in the market for credit risk. Journal of Financial and Quantitative Analysis 46, 627–656.

Shleifer, A., Vishny, R. W., 1997. The limits of arbitrage. Journal of Finance 88, 35–55.

A. Investment-grade bonds



B. High-yield bonds

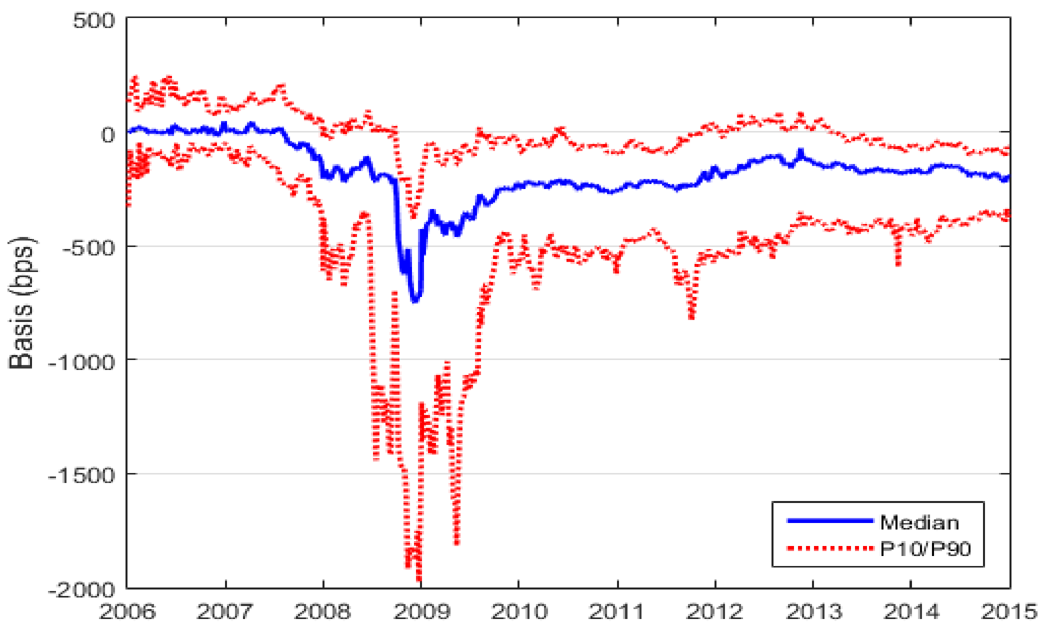


Figure 1: Dispersion of the CDS-Bond Basis by Credit Rating

Source: Markit Inc. and Federal Reserve Bank of New York.

Figure 2: The Average Five-Year CDS Spread for the Primary Dealers (bps)

Table 1: **Summary Statistics of Discrepancies in CDS and Cash Bond Spreads**

This table provides descriptive statistics for the average CDS-bond basis in four phases: Phase 1 is the period before the subprime credit crisis, named "Before Crisis" (July 2006 - June 2007), Phase 2 is the period between the subprime credit crisis and the bankruptcy of Lehman Brothers, called "Crisis I" (July 2007 - August 2008), Phase 3 is the period after Lehman Brothers' failure, "Crisis II" (September 2008 - September 2009), and Phase 4 is the period after the financial crisis, "Post Crisis" (October 2009 - December 2014). The basis is calculated as the difference between the CDS spread and the par-equivalent corporate bond spread, using the methodology in Appendix A. The summary statistics are reported for all bonds (ALL), investmeng-grade bonds (IG), high-yield bonds (HY), bonds issued separately by financial firms (F) and non-financial firms (NF), as well as across rating categires: AAA/AA, A, BBB, BB, B, and CCC. We calculate the cross-sectional mean, standard deviation, the 10th and the 90th percentile value of the bases across all bonds each day, and report the time-series average of these statistics. All entries are in basis points.

	Before Crisis				Crisis I				Crisis II				Post Crisis			
	Jul 2006 - Jun 2007				Jul 2007 - Aug 2008				Sep 2008 - Sep 2009				Oct 2009 - Dec 2014			
	Mean	SD	P10	P90	Mean	SD	P10	P90	Mean	SD	P10	P90	Mean	SD	P10	P90
ALL	-10	59	-57	45	-118	192	-273	14	-324	369	-667	-55	-137	152	-268	-32
IG	-17	30	-51	17	-83	108	-150	-10	-243	256	-451	-48	-101	71	-173	-32
HY	12	104	-107	142	-180	265	-486	57	-560	504	-1248	-114	-237	242	-477	-35
F	-24	26	-55	3	-117	161	-343	12	-351	450	-913	-4	-116	102	-205	-27
NF	-6	65	-56	59	-119	196	-254	15	-313	321	-625	-71	-145	165	-294	-34
AAA/AA	0	34	-31	51	-35	61	-90	32	-111	84	-218	-12	-68	40	-121	-23
A	-15	29	-49	17	-88	107	-154	-20	-187	180	-359	-32	-86	58	-145	-27
BBB	-23	29	-56	11	-103	68	-173	-37	-357	302	-607	-135	-122	77	-202	-49
BB	-7	40	-52	43	-107	126	-231	18	-493	313	-852	-182	-195	127	-363	-54
B	13	90	-104	125	-190	249	-499	62	-501	486	-1178	-72	-221	205	-445	-29
CCC	27	151	-128	167	-296	398	-807	120	-1113	568	-1747	-534	-413	418	-951	-8

Table 2: **Limit-to-Arbitrage Risk Factors**

This table shows the summary statistics of limit-to-arbitrage factors in Panel A, and their correlation values in Panel B, which is further divided into the crisis period (July 2007 - September 2008) and the non-crisis period (July 2006 - June 2007 and October 2009 - December 2014).

Panel A: Summary Statistics

Variable	Notation	Mean	Std	Med	P10	P90
Lending fee	Fee	0.12	0.40	0.09	0.04	0.16
Credit rating	Rat	12.64	3.53	13.50	7.00	16.50
Bond illiquidity (level)	Liq	0.42	1.68	0.11	0.00	0.88
Bond liquidity risk	β_{Liq}	0.62	2.72	0.21	-0.67	2.28
Bond liquidity mkt risk	β_{LiqM}	-0.03	0.36	-0.01	-0.19	0.13
Counterparty risk	β_{CP}	0.13	0.29	0.05	0.00	0.36
Funding liquidity risk	β_{FL}	-0.03	1.32	-0.03	-0.68	0.69

Panel B:

Correlation in the non-crisis period

	Fee	Rat	Liq	β_{Liq}	β_{LiqM}	β_{CP}	β_{FL}
Fee	1.00						
Rat	-0.29	1.00					
Liq	-0.01	-0.05	1.00				
β_{Liq}	0.00	-0.04	0.02	1.00			
β_{LiqM}	0.00	0.02	-0.07	-0.08	1.00		
β_{CP}	0.21	-0.31	0.03	0.00	0.00	1.00	
β_{FL}	-0.01	0.00	0.00	0.00	0.00	0.03	1.00

Correlation in the crisis period

	Fee	Rat	Liq	β_{Liq}	β_{LiqM}	β_{CP}	β_{FL}
Fee	1.00						
Rat	-0.20	1.00					
Liq	-0.04	-0.02	1.00				
β_{Liq}	0.00	-0.05	0.27	1.00			
β_{LiqM}	0.03	0.07	-0.38	-0.39	1.00		
β_{CP}	0.03	-0.03	0.03	0.02	0.02	1.00	
β_{FL}	0.00	0.05	-0.03	-0.01	0.08	-0.06	1.00

Table 3: **Multivariate Fama-MacBeth Regression of the *Negative* CDS-Bond Basis**

This table reports the average coefficients from the Fama-MacBeth cross-sectional regressions of the CDS-bond bases on risk factors including collateral quality proxies such as lending fee and credit rating, three proxies of bond trading liquidity and liquidity risk, counterparty risk, and funding liquidity risk. The sample is limited to negative basis only. The regression is conducted at the weekly frequency for 444 weeks from July 2006 to December 2014. Except collateral quality, all other factors are calculated for each bond at each week using a rolling window of past 60-week observations. The weekly lending fee or credit rating is calculated as the weekly average of daily observations. Credit rating is a series of numeric values with 1 referring to CCC, 2 to CC-,... and 21 to AAA. The t-statistics based on Newey-West adjusted standard errors are given in squared brackets. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Panel A. Before Crisis (Jul 2006 - Jun 2007)					Panel B. Crisis I (Jul 2007 - Aug 2008)				
Lending fee	0.000				0.000	-0.005***				-0.003***
	[-1.50]				[0.72]	[-4.61]				[-3.01]
Credit rating	0.070***				0.074***	0.622***				0.570***
	[7.25]				[3.86]	[6.79]				[7.25]
Bond illiquidity (level)		-0.000**			-0.020		-0.004			-0.236
		[-2.06]			[-1.45]		[-1.23]			[-0.85]
Bond liquidity risk		0.003			0.001		-0.345***			-0.465***
		[0.11]			[0.08]		[-2.86]			[-4.10]
Bond liquidity mkt risk		-0.004			-0.022		-0.822			-0.807*
		[-0.24]			[-1.17]		[-1.59]			[-1.92]
Counterparty risk			0.000***		0.059***			-0.003***		-0.181***
			[3.05]		[4.36]			[-4.51]		[-2.56]
Funding liquidity risk				0.000	0.001				0.001	0.132**
				[1.16]	[1.04]				[1.60]	[2.34]
Adj. R2	0.147	0.101	0.056	0.022	0.356	0.263	0.113	0.057	0.043	0.408
	Panel C. Crisis II (Sep 2008 - Sep 2009)					**Panel D. Post Crisis (Oct 2009 - Dec 2014)**				
Lending fee	-0.004***				-0.004***	-0.002***				-0.001***
	[-2.69]				[-2.58]	[-5.64]				[-5.01]
Credit rating	2.124***				1.902***	0.739***				0.779***
	[7.77]				[6.78]	[27.21]				[22.58]
Bond illiquidity (level)		-0.009***			-0.676***		-0.001***			0.002
		[-4.12]			[-3.16]		[-2.82]			[0.11]
Bond liquidity risk		0.538**			0.101		-0.001			0.006
		[2.34]			[0.39]		[-0.07]			[0.51]
Bond liquidity mkt risk		1.269***			0.517**		-0.017			0.033*
		[4.73]			[1.72]		[-1.49]			[2.17]
Counterparty risk			-0.007***		-0.545***			-0.002***		0.163***
			[-6.03]		[-5.22]			[-7.71]		[8.31]
Funding liquidity risk				0.010***	0.794***				0.000	0.039
				[5.30]	[3.71]				[0.75]	[0.77]
Adj. R2	0.215	0.075	0.037	0.080	0.366	0.287	0.022	0.025	0.145	0.442

Table 4: **The CDS-Bond Basis Portfolios Sorted by Size and Rating**

Each week we construct corporate bond portfolios based on bond outstanding amount and bond rating. Bond size is divided into four categories based on the 25th, 50th, 75th percentiles of the whole sample, and bond rating is categorized into six bins including AAA/AA, A, BBB, BB, B, and CCC&Below. There are a total of 24 portfolios. The sample is limited to negative basis only. Panel A reports the average basis value (bps) for each portfolio. Panel B reports the average coefficients from the Fama-MacBeth cross-sectional regressions of the CDS-bond bases for 24 portfolios on risk factors including collateral quality, three proxies of bond trading liquidity and liquidity risk, counterparty risk, and funding liquidity risk. The regression is conducted at the weekly frequency for 444 weeks from July 2006 to December 2014. Except collateral quality proxies, all other factors are calculated for each bond at each week using a rolling window of past 60-week observations. The weekly collateral quality is calculated as the weekly average of daily corporate bond loan lending fee. The t-statistics based on Newey-West adjusted standard errors are given in squared brackets. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Average basis (bps) for portfolios sorted by size and rating

	AAA/AA	A	BBB	BB	B	CCC&Below
Small	-126	-125	-201	-258	-318	-534
2	-79	-103	-146	-229	-283	-487
3	-75	-105	-132	-240	-285	-436
Large	-79	-98	-130	-238	-266	-631

Panel B. Fama-MacBeth Regression Results

	Before Crisis	Crisis I	Crisis II	Post Crisis
Lending fee	-0.002	-0.007***	-0.005***	-0.004***
	[-1.37]	[-4.93]	[-3.22]	[-6.43]
Credit rating	0.111***	0.542***	1.268***	0.702***
	[5.43]	[12.29]	[10.04]	[35.09]
Bond illiquidity (level)	-0.083***	0.013	-0.261	0.261***
	[-2.65]	[0.08]	[-1.12]	[5.83]
Bond liquidity risk	-0.053	-0.622***	-0.421	0.312***
	[-1.01]	[-5.80]	[-1.43]	[4.69]
Bond liquidity mkt risk	0.020	-0.503***	0.809***	0.002
	[0.52]	[-2.49]	[2.51]	[0.04]
Counterparty risk	0.145***	-0.281***	-0.800***	0.118***
	[5.60]	[-3.92]	[-6.43]	[2.45]
Funding liquidity risk	-0.011	0.111	0.720***	0.156***
	[-0.37]	[1.15]	[3.43]	[2.67]
Adj. R2	0.799	0.749	0.774	0.839

Appendix

A The Par-Equivalent CDS Methodology

We present the Par Equivalent CDS methodology developed by J.P. Morgan to calculate the CDS-bond basis in Section 2. This survival-based valuation approach provides an apple-to-apple measure across the cash-bond spread and the credit default swap spread.

The fair value of the coupon on a CDS is set so that the expected present value of the premium leg is equal to the expected present value of the contingent payment (see Duffie (1999)). Assuming that we have a zero-coupon discount curve $Z(t)$ extracted from swap spreads and assuming a constant intensity survival probability $S(t)$, the expected present value of the premium leg is given by:

$$\mathrm{PV}_{premium}(C) = \sum_{i=1}^{n} Z(t_i)S(t_i)\, C * dt + \sum_{i=1}^{n} Z(\frac{t_i + t_{i-1}}{2})[S(t_{i-1}) - S(t_i)]\, C * dt/2, \tag{9}$$

where the second component is the present value of the accrued interest upon default (assumed to occur halfway between t_{i-1} and t_i). The expected present value of the contingent leg is:

$$\mathrm{PV}_{contingent} = (1 - R) \sum_{i=1}^{n} Z(\frac{t_i + t_{i-1}}{2})[S(t_{i-1}) - S(t_i)], \tag{10}$$

where R stands for the recovery rate. The fair credit default swap spread is the number C that sets

$$\mathrm{PV}_{premium}(C) = \mathrm{PV}_{contingent} \tag{11}$$

The par-equivalent CDS uses the market price of a bond to calculate a spread based on CDS-implied default probabilities. First, we need to get a CDS-implied default probability curve $S_{CDS}(t_i)$ by sequentially plugging in CDS spread with maturity from 1-year to 10-year. Second, we need to get a bond-implied survival probability curve $S_{bond}(t_i)$. Using the CDS-implied survival probability as a prior, we calculate the bond-implied survival probability curve as the one that minimize the pricing error between the market price and derived bond price:

$$S_{bond}(t_i) = S_{CDS}(t_i) + \varepsilon. \tag{12}$$

$$\text{s.t.} \qquad \varepsilon = \arg\min \left(PV(S_{bond}) - \text{Market Price of Bond}\right)^2. \tag{13}$$

Then the bond-implied CDS spread term structure is defined by substituting the survival probability term structure fitted from bond prices, $S_{bond}(t)$, into the following equation for par equivalent CDS spreads, denoted as PECDS:

$$PECDS = \frac{(1-R)\sum_{i=1}^{n} Z(\frac{t_i+t_{i-1}}{2})[S_{bond}(t_{i-1}) - S_{bond}(t_i)]}{\sum_{i=1}^{n} \left[Z(t_i)S_{bond}(t_i) * dt + Z(\frac{t_i+t_{i-1}}{2})[S_{bond}(t_{i-1}) - S_{bond}(t_i)] * \frac{dt}{2}\right]} \tag{14}$$

B The Primary Dealers List

Below is the list of primary dealers as of July 27, 2009, reported on the website of the Federal Reserve Bank of New York.

BNP Paribas Securities Corp.
Banc of America Securities LLC
Barclays Capital Inc.
Cantor Fitzgerald & Co.
Citigroup Global Markets Inc.
Credit Suisse Securities (USA) LLC
Daiwa Securities America Inc.
Deutsche Bank Securities Inc.
Goldman, Sachs & Co.
HSBC Securities (USA) Inc.
Jefferies & Company, Inc.
J. P. Morgan Securities Inc.
Mizuho Securities USA Inc.

Morgan Stanley & Co. Incorporated

Nomura Securities International, Inc.

RBC Capital Markets Corporation

RBS Securities Inc.

UBS Securities LLC.