DOI: 10.1111/fima.12261

WILEY Financial Management

ORIGINAL ARTICLE

Price discovery and persistent arbitrage violations in credit markets

Hai Lin[1] | Kasing Man[2] | Junbo Wang[3] | Chunchi Wu[4]

[1]School of Economics and Finance, Victoria University of Wellington, Wellington, New Zealand

[2]College of Business and Technology, Western Illinois University, Macomb, IL

[3]School of Economics and Finance, City University of Hong Kong, Kowloon Tong, Hong Kong

[4]School of Management, State University of New York at Buffalo, Buffalo, NY

Correspondence
Chunchi Wu, School of Management, State University of New York at Buffalo, Buffalo, NY.
Email: chunchiw@buffalo.edu

Abstract

This paper investigates price violations in credit markets using a data sample spanning from 2002 to 2016. We find that price violations are highly persistent during the crisis period, particularly for speculative-grade bonds. There is evidence that price distortions and market disintegration are linked to market-wide and firm-level impediments to arbitrage and limited capital provision. Higher firm-level impediments to arbitrage lead to less market integration, and more severe and persistent pricing discrepancies. Moreover, we find that the negative CDS basis persists in the postcrisis period, which is attributable to dealers' lower capital commitment and deterioration in market-making quality.

1 | INTRODUCTION

The arbitrage-free condition has been a building block in the development of modern financial and asset pricing theories. For example, standard term structure models build on the assumption of no-arbitrage conditions, and asset pricing theory assumes that any temporary deviations of prices from efficient benchmarks can be arbitraged away quickly by rational traders. However, it has been shown that the arbitrage-free condition has often been violated in financial markets (see Kapadia & Pu, 2012) and that violations were especially severe during the subprime crisis (see Duffie, 2010; Mitchell & Pulvino, 2012).[1] In particular, acute price violations in credit markets during the crisis attracted considerable attention. The credit default swap (CDS) basis, or the difference between the CDS rate and yield spreads of a par bond with the same maturity as the CDS, widened to above 600 basis points after the collapse of Lehman Brothers. As Duffie pointed out in his 2010 American Finance Association presidential address, "The extreme negative CDS basis 'violations' … across broad portfolios of investment-grade bonds and high-yield bonds, respectively, is far too large to be realistically explained by CDS counterparty risk or by other minor technical details."

CDS and bond yield spreads both reflect a firm's credit risk premium. In a frictionless market, any discrepancy in these two variables will be eliminated quickly by arbitrage. This suggests that the CDS basis should be close to zero if arbitrage is perfect. However, arbitrage is rarely perfect, as cash flows of CDS and the reference obligation are typically

[1] For example, there were serious violations of covered interest rate parity, a negative spread between Treasury bond yields and London Inter-Bank Offered Rate (LIBOR) swap rates, and a breakdown of the capital structure arbitrage across equity and credit markets.

not perfectly aligned. There are also other complications; for instance, physically settled CDS prices contain the value of the cheapest-to-delivery (CTD) option, and arbitrage may require shorting the cash bond which can be costly or sometimes infeasible. Past studies prior to the subprime crisis have shown that the CDS basis was close to zero, or slightly positive, which can be attributed to the CTD option value or costly short selling. The huge negative CDS basis that occurred during the subprime crisis is thus more difficult to explain, and posts significant challenges to rational no-arbitrage pricing theory.

Violations of the arbitrage-based pricing relationship have profound implications for market efficiency and asset pricing. In particular, persistent price violations can cause asset price distortions and market disintegration. The literature has suggested various sources of asset mispricing. Price violations are commonly attributed to limits-to-arbitrage (see Brav, Heaton, & Li, 2010; Mitchell & Pulvino, 2012; Pontiff, 2006; Shleifer & Vishny, 1997). Arbitrage may not be feasible when transaction costs and the risk of a firm's security are excessively high. Funding constraints or limited arbitrage capital provision can prevent arbitrage activity (Brunnermeier & Pedersen, 2009; Mitchell & Pulvino, 2012) and cause serious mispricing in similar securities.

In this paper, we examine the issues related to the violation of arbitrage-free pricing relations in credit markets using a data sample spanning from 2002 to 2016, which covers both the crisis and normal periods. We employ two methods to assess the severity of pricing violations in credit markets. First, using a vector error-correction model (VECM), we examine the dynamics of CDS and bond spreads, and conduct a test on the equivalence of the prices of credit risk across the two markets over time via cointegration analysis. Second, we perform a nonparametric model-free test for market integration based on the concordance of price changes in the CDS and bond markets.

A unique character of the price distortion during the subprime crisis is that pricing discrepancies are not only large but also highly persistent. Besides studying the magnitude of CDS-bond mispricing, we adopt a long memory model to quantify the persistence in pricing discrepancies, an issue which is much less studied in the literature. The long memory model provides a generic time-series measure of persistence or long-range dependency (see Hosking, 1981; Lo, 1991; Zivot & Wang, 2006), which characterizes the slow-moving nature of arbitrage capital in times of stress very well. An advantage of this model is that the persistence in pricing discrepancies due to arbitrage capital shortage or other frictions can be nicely summarized by a parsimonious parameter that measures how slowly moving the arbitrage capital is and how persistent the resulting price violations are. After identifying the pricing discrepancies and their persistence in credit markets, we explore their relations to impediments to arbitrage at both the market and firm levels.

We document several interesting findings that contribute to the current literature. First, both cointegration and nonparametric tests show that pricing discrepancies across firms' CDS and bond markets are common and much more serious in the crisis period. The level and volatility of CDS-bond pricing discrepancies can be quite persistent. The persistence in price discrepancies is much stronger in the crisis period than during normal periods, and is higher for noninvestment-grade (IG) bonds than for IG bonds.

Second, we find that both the magnitude and persistence of price discrepancies are closely related to firm-level and market-wide impediments to arbitrage. Consistent with the theory of costly arbitrage (Shleifer & Vishny, 1997), firms with high risk, illiquidity, leverage, and return volatility experience more severe and persistent pricing discrepancies. Violations of price convergence are also more frequent, and price discrepancies are larger and more persistent in the period when funding constraints are high and market liquidity dries up. The relative role of firm-level impediments to systematic (market-wide) impediments depends on market conditions. We find that firm-level impediments to arbitrage play a nontrivial role in CDS-bond pricing violations during normal market periods, but that in times of stress, systematic impediments become more important factors.

Third, there is evidence that slow capital movement aggravates the persistence of price violations in credit markets during the subprime crisis. Using several proxies for arbitrage capital provision, we find that the sensitivity of firms to funding availability has high explanatory power for the persistence of price violations in credit markets in times of stress. In addition, the funding-related variables exhibit a strong persistence pattern during the crisis period which coincides with the persistence in CDS-bond pricing discrepancies. These findings lend support to the contention that persistence in the shortage of arbitrage capital causes persistence in pricing discrepancies in credit markets.

Finally, the negative CDS basis is not just a phenomenon during the subprime crisis. It continues to persist in the postcrisis period. We uncover evidence that the postcrisis negative CDS basis is attributable to dealers' lower capital commitment to market making, which reduces their ability and willingness to provide liquidity to bond markets and results in poorer market quality. Our results suggest that postcrisis regulation reforms have an unintended adverse effect on pricing efficiency in credit markets.

Understanding the cause of pricing discrepancies in bond and CDS markets is important for the development of financial theories. A number of papers have examined the issues of price discovery in credit markets and the CDS basis behavior (see, for example, Bai & Collin-Dufresne, 2018; Blanco, Brennan, & Marsh, 2005; Choi & Shachar, 2014; Das, Kalimipalli, & Nayak, 2014; Duffie, 2010; Fontana, 2011; Galil, Shapir, Amiram, & Ben-Zion, 2014; Li, Zhang, and Kim, 2014; Nashikkar, Subrahmanyam, & Mahanti, 2011). Our work differs from these studies in several key aspects. First, we propose quantitative measures of pricing violation and its persistence, explore the determinants of credit risk mispricing and duration, and investigate the role of impediments-to-arbitrage and limited capital provision in the CDS-bond pricing discrepancies. Second, we document evidence that persistence of price violations in credit markets is closely linked to persistence of slow capital movement. Third, we examine postcrisis CDS behavior and find that the negative basis also occurs in recent years and that this phenomenon is related to dealers' lower capital commitment to market making and deterioration in the quality of intermediacy provision.

The remainder of this paper is organized as follows. Section 2 discusses the data and variables used in empirical tests. Section 3 examines the (dis)integration between the CDS and bond markets and its link to impediments to arbitrage at both the firm and market levels, and limited arbitrage capital provision. Section 4 examines persistence in the level and volatility of pricing discrepancies and investigates the factors contributing to CDS basis persistence. Finally, Section 5 summarizes our major findings and concludes the paper.

2 | DATA

Our sample consists of CDS, bond, and stock data, compiled from multiple sources. Daily CDS data are provided by the Markit Group. Bond transaction data come from the Trade Reporting and Compliance Engine (TRACE), bond characteristic information from the Fixed Investment Securities Database (FISD), daily stock returns from the Center for Research in Security Prices (CRSP), and financial statement information from Compustat. The sample period runs from July 2002 to December 2016.

Markit reports the daily composite CDS spread by aggregating dealers' quotes. For each reference entity, a number of CDS contracts may exist with differences in maturity, bond seniority, currency denomination, and treatments of restructuring in credit event definition. We choose the US dollar-denominated five-year CDS contracts on senior unsecured debts with modified restructuring (MR) for our study and focus on the CDSs with US reference entities.

TRACE provides prices, yields, par value of transactions, and other trading information of corporate bonds. We filter out the observations with apparent recording errors, and use the data screening procedure recommended by Bessembinder, Kahle, Maxwell, and Xu (2009) to eliminate cancelled, corrected, and commission trades. The FISD database contains issue- and issuer-specific information, such as coupon rates and frequency, maturity date, issue amount, provisions, and credit rating history for all US corporate bonds maturing in 1990 or later.

Our data set only includes the firms for which CDS, stock, and bond data are available over the entire sample period. Bonds with floating rate coupons or embedded options (convertible, putable, callable, and sinkable, etc.) are excluded from the sample. We focus on straight bonds to avoid the potential confounding effects associated with embedded options.[2] To calculate the CDS-bond basis, we use the par equivalent CDS methodology developed by JP Morgan.

The CDS basis is measured by the difference between the quoted CDS spread and the par-equivalent CDS spread (*PECDS*) on the same reference entity:

[2] To be included in the sample, firms must also have data such as return volatility, leverage, number of CDS suppliers, and any other data required to calculate the Amihud illiquidity and expected default frequency. See Section 2.1 for a description of these data.

$$Basis_{i,t} = CDS_{i,t} - PECDS_{i,t},$$

where $CDS_{i,t}$ and $PECDS_{i,t}$ are the quoted CDS spread for the five-year contract and the par-equivalent CDS spread at time *t*, respectively. We follow the procedure in Nashikkar et al. (2011) and Bai and Collin-Dufresne (2018) to calculate the *PECDS*. Given the price information of bonds for a firm at time *t*, we calibrate the constant default intensity for the firm by minimizing the pricing errors of its corporate bonds. We use the bonds for each firm with a maturity between three and eight years in the calibration process. We then use the default intensity calibrated from bond prices to calculate the par-equivalent five-year CDS spread. The par-equivalent CDS spread is set equal to the coupon rate that equates the expected value of the premium leg to that of the contingent leg. Following the literature, the recovery rate is set at 40%. Using our matched sample, we are able to estimate par-equivalent spreads for 759 firms, which include 478 reference entities (firms) with an IG and 281 entities with a speculative grade (SG). In our baseline analysis, we use swap rates as reference rates for the risk-free funding curve to compute the spread. Besides swap rates, we use Treasury zero rates as an alternative measure of risk-free rates for our robustness check, which are downloaded from the Federal Reserve Board's Web site.[3]

Panel A of Table 1 provides a summary of the data for the whole sample period and for subperiods where the CDS basis is calculated using the swap rate. Figure 1 plots the CDS basis for the whole sample (1a) and by bond grade (1b). Over the whole sample period, the average CDS basis is negative (–0.22%), which is largely due to the reverse relation between CDS and bond spreads during the financial crisis and in the postcrisis period. We define the period from July 1, 2007 to June 30, 2009 as the crisis period (see also Dick-Nielsen, Feldhutter, & Lando, 2012; Friewald, Jankowitsch, & Subrahmanyam, 2012). In the precrisis period (July 2002 to June 2007), the average basis is positive (0.32%). The average basis then becomes quite negative (–1.06%) during the crisis period. In the postcrisis period (July 2009 to December 2016), the CDS basis remains negative with an average of –0.35%, though it is much milder than that in the crisis period. Volatility of the CDS basis also increases considerably during the crisis period. Both bond and CDS spreads were much more volatile during the crisis period, with the volatility of the former being substantially higher than the latter.

Turning to the results for subsamples, we find that the patterns of negative basis and high volatility are more pronounced for SG bonds than for IG bonds. During the crisis period, the average CDS basis is –0.77% for IG bonds and –1.77% for SG bonds. Volatility of the CDS basis is also much higher for SG bonds (2.56%) than for IG bonds (0.92%).

Panel C of Table 1 summarizes the data by rating. During the crisis period, the average CDS basis ranges from –0.60% (A bonds) to –1.77% (junk bonds). In the precrisis period, the basis is positive for all ratings, whereas in the postcrisis period, the basis is negative across all ratings.

Following the convention, we calculate the above CDS basis using swap rates. For robustness, we also calculate the CDS basis based on Treasury zero-coupon rates. Figure 2 plots the CDS basis based on Treasury zero rates for the whole sample (Figure 2a) and subsamples by bond grade (Figure 2b). As shown, the time-series pattern is similar to that in Figure 1. The CDS basis becomes more negative when using the Treasury zero rate of the same maturity as the risk-free rate. Panels B and D of Table 1 report the differences in the CDS basis based on the swap rate and Treasury zero rate. Results show that the measures of CDS basis are fairly sensitive to the choice of risk-free rates. The differences (in basis points) are larger during the crisis period and for SG bonds.

One of our main objectives is to examine the role of limited arbitrage in CDS-bond pricing discrepancies during the financial crisis. To accomplish this objective, we select a number of impediment-to-arbitrage variables at the market and firm levels. We discuss these variables below.

2.1 | Market-wide impediment variables

We consider a number of market-wide (systematic) impediment-to-arbitrage variables suggested by the literature. Past studies have suggested several important sources of arbitrage frictions: funding cost and constraints (Brunnermeier & Pedersen, 2009; Fontaine & Garcia, 2012), volatility (Pontiff, 2006) or market uncertainty (Shleifer & Vishny,

[3] See Gürkaynak, Sack, and Wright (2007) for the procedure of calculating the spot rates.

TABLE 1 Summary statistics

Panel A. Summary of the CDS Basis Using Swap Rates and Bond and CDS Spreads

		Basis			Bond			CDS		
	Period	Mean	Median	Std.	Mean	Median	Std.	Mean	Median	Std.
All bonds	Whole	−0.22	−0.14	0.74	2.07	1.78	1.48	1.85	1.61	0.88
	Precrisis	0.32	0.34	0.23	1.14	0.88	0.86	1.46	1.29	0.77
	Crisis	−1.06	−0.52	1.40	4.20	3.61	2.38	3.14	3.08	1.05
	Postcrisis	−0.35	−0.28	0.33	2.12	2.06	0.69	1.77	1.67	0.50
Investment grade	Whole	−0.18	−0.15	0.53	1.05	0.96	0.96	0.87	0.70	0.54
	Precrisis	0.27	0.27	0.13	0.41	0.19	0.57	0.67	0.46	0.58
	Crisis	−0.77	−0.37	0.92	2.20	1.46	1.69	1.43	1.18	0.80
	Postcrisis	−0.32	−0.27	0.25	1.18	1.12	0.39	0.85	0.84	0.24
Speculative grade	Whole	−0.35	−0.11	1.31	4.28	3.49	2.83	3.94	3.44	1.74
	Precrisis	0.35	0.45	0.74	2.80	2.11	2.02	3.14	2.85	1.57
	Crisis	−1.77	−0.75	2.56	8.11	6.96	4.53	6.34	6.19	2.16
	Postcrisis	−0.43	−0.33	0.59	4.25	4.05	1.39	3.82	3.71	1.02

Panel B. Differences in Basis Points between the CDS Basis Measures Using Swap Rates and Treasury Zero Rates

	Period	Mean	Median	Std.	Max	Min
All bonds	Whole	33.18	32.48	26.51	122.99	−21.00
	Precrisis	44.17	44.32	7.90	73.40	16.24
	Crisis	79.01	77.78	20.49	122.99	27.95
	Postcrisis	13.84	13.74	13.84	56.43	−21.00
Investment grade	Whole	32.03	31.45	25.51	121.25	−20.00
	Precrisis	43.36	43.40	7.71	71.07	16.20
	Crisis	74.81	73.81	20.62	121.25	22.86
	Postcrisis	13.22	13.41	13.35	54.74	−20.00
Speculative grade	Whole	35.68	35.13	28.57	126.06	−21.00
	Precrisis	46.00	45.90	8.62	81.75	16.35
	Crisis	87.08	85.00	20.26	126.06	38.90
	Postcrisis	15.17	14.59	14.95	71.00	−21.00

Panel C. Summary of the CDS Basis and Bond and CDS Spreads by Rating

		Basis			Bond			CDS		
	Period	Mean	Median	Std.	Mean	Median	Std.	Mean	Median	Std.
AAA/AA	Whole	−0.18	−0.09	0.66	0.91	0.66	1.20	0.74	0.43	0.63
	Precrisis	0.28	0.30	0.17	0.17	0.05	0.44	0.45	0.34	0.33
	Crisis	−0.93	−0.35	1.31	2.58	1.72	2.26	1.65	1.42	1.05
	Postcrisis	−0.28	−0.24	0.23	0.96	0.91	0.45	0.68	0.55	0.34
A	Whole	−0.12	−0.07	0.46	0.79	0.77	0.72	0.68	0.60	0.38
	Precrisis	0.28	0.28	0.12	0.28	0.10	0.48	0.56	0.39	0.49
	Crisis	−0.60	−0.25	0.77	1.58	1.01	1.25	0.98	0.83	0.51
	Postcrisis	−0.25	−0.18	0.25	0.92	0.87	0.29	0.68	0.66	0.13
BBB	Whole	−0.27	−0.23	0.66	1.54	1.33	1.32	1.26	1.03	0.91
	Precrisis	0.25	0.23	0.43	0.82	0.41	1.21	1.07	0.62	1.22

(Continues)

TABLE 1 (Continued)

Panel C. Summary of the CDS Basis and Bond and CDS Spreads by Rating										
		Basis			Bond			CDS		
	Period	Mean	Median	Std.	Mean	Median	Std.	Mean	Median	Std.
	Crisis	−0.94	−0.52	1.03	2.92	1.98	2.14	1.98	1.57	1.16
	Postcrisis	−0.44	−0.40	0.34	1.64	1.58	0.56	1.20	1.16	0.31
Below BBB	Whole	−0.35	−0.31	1.31	4.28	3.49	2.83	3.94	3.44	1.74
	Precrisis	0.35	0.45	0.74	2.80	2.11	2.02	3.14	2.85	1.57
	Crisis	−1.77	−0.75	2.56	8.11	6.96	4.53	6.34	6.19	2.16
	Postcrisis	−0.43	−0.33	0.59	4.25	4.05	1.39	3.82	3.71	1.02

Panel D. Differences between the CDS basis Measures Using Swap Rates and Treasury Zero Rates						
	Period	Mean	Median	Std.	Max	Min
AAA/AA	Whole	32.45	32.06	25.65	119.36	−19.00
	Precrisis	43.86	44.09	7.54	74.38	16.08
	Crisis	75.27	73.89	20.66	119.36	23.39
	Postcrisis	13.45	13.55	13.54	54.76	−19.00
A	Whole	31.83	31.16	25.25	120.66	−20.00
	Precrisis	43.31	43.48	7.61	68.70	16.40
	Crisis	73.77	73.25	20.32	120.66	22.99
	Postcrisis	13.04	13.36	13.21	54.09	−20.00
BBB	Whole	32.32	31.59	25.92	123.69	−21.00
	Precrisis	43.26	43.14	8.19	75.75	15.83
	Crisis	76.15	75.10	21.23	123.69	22.50
	Postcrisis	13.35	13.35	13.50	55.58	−21.00
Below BBB	Whole	35.68	35.13	28.57	126.06	−22.00
	Precrisis	46.00	45.90	8.62	81.75	16.35
	Crisis	87.08	85.00	20.26	126.06	38.90
	Postcrisis	15.17	14.59	14.95	71.00	−22.00

This table summarizes the bond yield, CDS spread, and CDS basis. The sample period spans from July 2002 to December 2016. The whole sample is further divided into three subperiods: the precrisis period is from July 2002 to June 2007, the crisis period is from July 2007 to June 2009, and the postcrisis period is from July 2009 to December 2016. Panel A provides a summary for all bonds, investment-grade, and speculative-grade bonds, where the CDS basis (in percentage) is calculated using swap rates. Panel B show the differences in the CDS basis (in bps) using the swap rate and Treasury zero rates. Panels C and D provide the corresponding summary statistics by rating category. There are 478 firms with an investment grade and 281 firms with a speculative grade.

1997), illiquidity (Hu, Pan, & Wang, 2013; Kapadia & Pu, 2012), and counterparty risk (Gorton and Metrick, 2012; Mitchell & Pulvino, 2012). We select variables related to these arbitrage impediments to perform empirical tests.

The Libor-OIS spread is widely used as a measure for funding cost (see Fontaine & Garcia, 2012). We measure the Libor-OIS spread (*LOIS*) by the difference between the three-month Libor rate and the overnight index swap (*OIS*) rate, both collected from the Bloomberg system. In addition, we consider several other variables related to funding constraints and costs: (1) changes in primary dealers' position in long-term corporate securities (*DPDPL*), which captures the effect of funding liquidity shocks that force dealers to liquidate their assets; (2) net convertible hedge fund inflow (*CF*) and net total hedge fund inflow (*HF*), which measure the funding availability of convertible hedge funds and the entire hedge fund industry, respectively; and (3) total repo volume (*RV*) in the repo funding market.[4]

[4] We also construct major indicators for funding costs in the repo market, such as the Libor-repo spread (*LR*), the spread between MBS and Treasury repo rates (*MBST*), and the spread between the general collateralized repo rate and Treasury bill rate (*GCT*). Our results are robust to these measures.

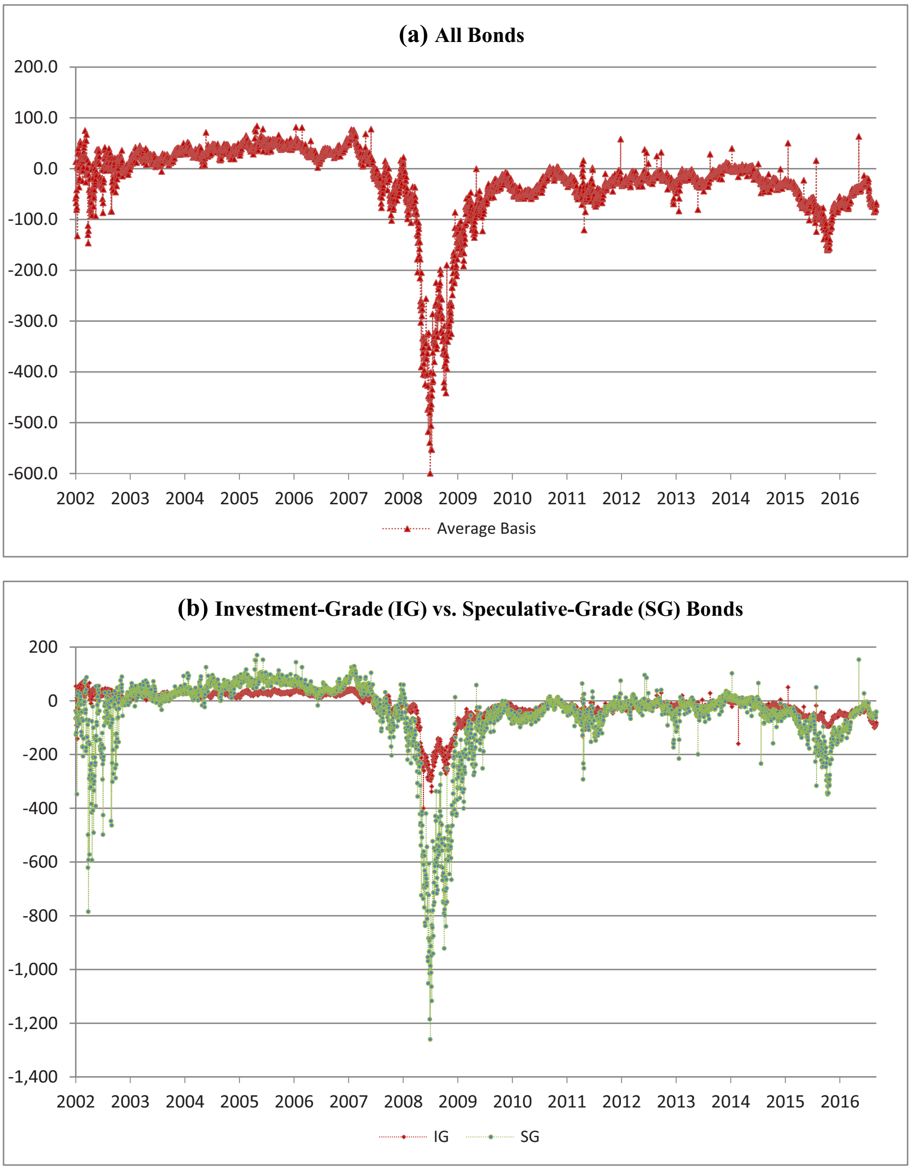

FIGURE 1 The Time Series of Average CDS Basis Using Swap Rates. This figure plots the mean observed basis (in basis points) over the sample period. Panel A plots the basis averaged across all bonds. Panel B shows the results of investment- versus speculative-grade bonds. The red line (square) represents the observed data for investment-grade bonds and the green line (circle) represents the observed data for speculative-grade bonds

Mitchell and Pulvino (2012) suggest that funding shortages in hedge funds contribute to arbitrage crash. The variables of net convertible hedge fund inflow and total hedge fund flow are used to capture this effect. Deleveraging (*DPDPL*) and RV data are collected from the Federal Reserve Bank (FRB) of New York, and net convertible hedge fund inflow and total hedge fund flow data are obtained from the TASS database, aggregated over all hedge funds in this database.

In addition to the above funding cost/constraint variables, we measure the cash shortfall (*CSF*) of primary dealers. Primary dealers play an important role in providing liquidity and arbitrage capital. A severe cash shortfall will hamper their ability to perform these functions. We estimate the firm's cash flow shortfall as investment + dividends − available

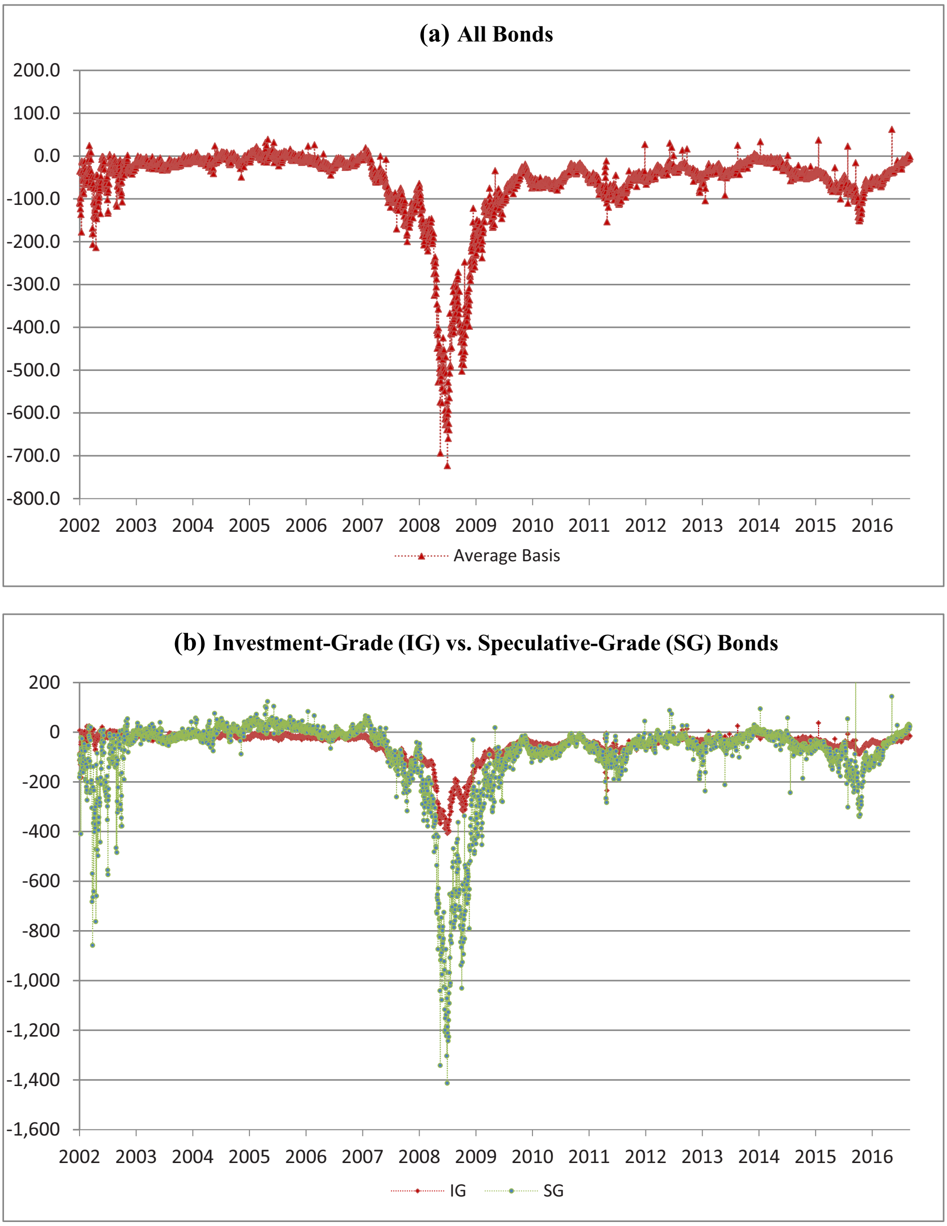

FIGURE 2 The Time Series of Average CDS Basis Using Treasury Zero Rates. This figure plots the mean observed basis (in basis points) over the sample period. Panel A plots the basis averaged across all bonds, and panel B shows the results of investment- versus speculative-grade bonds. In panel B, the red line (square) represents the observed data for investment-grade bonds and the green line (circle) represents the observed data for speculative-grade bonds

cash flow, where available cash flow equals cash flow from operations – preferred dividends. This formula for cash shortfall is similar to that suggested by Daniel, Denis, and Naveen (2017) and adapted to financial firms. We use the quarterly data from Compustat to calculate the cash shortfall for primary dealers.[5]

[5] Although the Compustat data are quarterly, different firms report the data in different months. In the end, we obtained a series of monthly averages for primary dealers.

For market-wide liquidity, we consider the Pastor-Stambaugh (PS) and the Hu-Pan-Wang (2013, HPW) illiquidity index (*NOISE*).[6] The PS bond market liquidity index (*PSB*) is constructed using the same method as in Pastor and Stambaugh (2003) by requiring all bonds to have at least ten transactions per month. The PS bond market liquidity index is calculated using the transaction data of TRACE. The HPW illiquidity index (*NOISE*) is downloaded from Jun Pan's Web site.

For market volatility, we use VIX (*VIX*), provided by the Chicago Board Options Exchange (CBOE), as a measure for market uncertainty. For the counterparty risk measure, we calculate the average CDS spread for primary dealers (*PDCDS*) to capture the credit risk of financial intermediaries in fixed-income markets.

2.2 | Firm-level impediment variables

As to the firm-level impediments to arbitrage, we consider several variables. Shleifer and Vishny (1997) and Mitchell, Pulvino, and Stafford (2002) suggest that uncertainty of returns discourages arbitrage due to lack of diversification in the arbitrageur's portfolio and uncertainty over the distribution of arbitrage returns. We use the volatility of stock returns (*Volatility*) for each firm as a measure of this uncertainty. The literature has also suggested that arbitrage is riskier for firms with higher risk. We use leverage (*Leverage*) and expected default frequency (*EDF*) as measures of firm risk. Leverage is the ratio of the book value of debt to the sum of book value of debt and market value of equity. We calculate the expected default frequency to measure a given firm's probability of default. We use Merton's (1974) model to estimate the likelihood of default, and estimate the distance-to-default and expected default frequency using the method suggested by Bharath and Shumway (2008).[7] To capture the effect of liquidity in the CDS market, we use the number of CDS suppliers (*Supply*) as a proxy for liquidity provision.

Moreover, arbitraging illiquid bonds is harder, as it is more difficult to locate counterparties for these bonds (Duffie, Garleanu, & Pedersen, 2007). Illiquid bonds also tend to have higher trading costs that impede arbitrage as it is more costly to trade them (Dick-Nielsen et al., 2012; Friewald et al., 2012; Lin, Wang, & Wu 2011; Nashikkar et al., 2011). We choose the Amihud individual bond illiquidity measure (*Amihud*) as a variable of firm-specific illiquidity. The Amihud firm-specific illiquidity measure is constructed from the transaction data of the firm's bonds using the Amihud (2002) method.

3 | DISINTEGRATION IN CREDIT MARKETS

We begin the analysis by examining the (dis)integration between the CDS and corporate bond market using different tests. We then link the violation of equilibrium pricing relations to impediments to arbitrage in order to assess the role of limited arbitrage in market disintegration.

3.1 | Cointegration tests

A formal test of the equilibrium relation or the equivalence of bond and CDS spreads in credit markets is the cointegration test (see Blanco et al., 2005). We can think of the observed price as a combination of the efficient price of credit risk and noise. If CDS and bond markets exhibit an equilibrium relation, their prices should be cointegrated, and any transitory deviations (noise) from the equilibrium will be corrected quickly if there are no arbitrage frictions.

To perform the cointegration test, we first estimate the VECM of CDS and bond spreads. Denote $Y_t = (p_t, q_t)'$, where p_t and q_t are CDS spreads and bond yield spreads measured by PECDS. The VECM model can be written as:

[6] We also consider the Amihud bond illiquidity measure using the method of Lin et al. (2011), as well as the on-/off-the-run spread, which is the difference between five-year on- and off-the-run Treasury yields (see Longstaff, Mithal, & Neis, 2005) from the FRB Web site. Our results are robust to these liquidity measures.

[7] The distance to default is calculated using the iterated estimate of volatility of firm value. We calculate the monthly distance to default for each firm using items 45 and 51 in the quarterly Compustat data file, the risk-free rate from the FRB, and the number of shares outstanding and share price from the CRSP.

TABLE 2 Cointegration tests

Period	Percentage of firms with cointegration
Precrisis	77.14%
Crisis	62.92%
Postcrisis	84.62%

This table summarizes cointegration test results based on both observed CDS and bond spreads for the normal and crisis periods. The precrisis period is from July 2002 to June 2007, the crisis period is from July 2007 to June 2009, and the postcrisis period is from July 2009 to December 2016. Reported figures are the percentage of firms which do not reject the hypothesis of cointegration with one cointegrating vector at the 5% significance level.

$$\Delta Y_t = c_0 + \delta\omega' Y_{t-1} + \sum_{i=1}^{m} \Psi_i \Delta Y_{t-i} + u_t, \tag{1}$$

where c_0 is the constant, u_t's are serially uncorrelated innovations with mean zero and covariance matrix $\mathrm{Var}(u_t)$with diagonal elements σ_1^2 and σ_2^2, and off-diagonal elements $\rho\sigma_1\sigma_2$, $\omega = (1, \zeta)'$ is the cointegration vector with the first element normalized to one, $\delta = (\delta_1, \delta_2)'$ is the vector of responses for the error-correction term, $\Delta = 1 - B$ is the difference operator with B as the back-shift operator, and m is the lag order. We apply Johansen's (1995) trace test to the VECM. Details of the test procedure are described in Appendix A.

Table 2 reports the results of cointegration tests at the 5% significance level for both normal and crisis periods. Results show that a higher proportion of firms support the cointegration hypothesis in normal periods (including pre- and postcrisis periods). This result is consistent with the finding of Blanco et al. (2005) based on the data before the subprime crisis. However, the situation worsens during the crisis period, with only 63% of the firms supporting the cointegration hypothesis. The CDS and corporate bond markets are much less integrated during the crisis period.

An important question is what may have caused the deterioration in the bilateral credit market relation. A plausible reason is that it became more difficult to arbitrage to restore the equilibrium relation during the crisis period. To investigate the role of limited arbitrage in market disintegration, we link the violation of integration to variables related to the impediments of arbitrage at both firm and market levels. A finding of a significant relationship between the violation of market integration and variables related to impediments to arbitrage will be supportive of the limited arbitrage hypothesis.

We calculate the mean value of each impediment variable for firms that reject or accept the cointegration test, and conduct two-sample mean difference tests for these two groups. Panel A of Table 3 reports these tests for different subperiods. Results show that firm-level impediment differences are quite significant for the crisis period, suggesting that limited arbitrage plays a more important role in times of stress.[8] Firms which reject the cointegration hypothesis tend to have high stock return volatility, default risk (EDF), and leverage, and low liquidity during the crisis period. In contrast, for the normal period, these issues are less important.

In the preceding analysis, we use firm-level impediment variables as measures for arbitrage frictions. Besides firm-level impediments, the literature has suggested that funding cost and constraints, counterparty risk, uncertainty and illiquidity, and lack of arbitrage capital at the market level can deter arbitrage (see Brunnermeier & Pedersen, 2009; Garleanu & Pedersen, 2011). To understand the role of these systematic impediments to arbitrage during the crisis, we examine the differences between market-level impediment variables in the normal and crisis periods. We employ variables associated with four dimensions of limits to arbitrage: funding cost and constraints (*LOIS*, *DPDPL*, *CF*, *CSF*), liquidity (*NOISE*, *PSB*), volatility (*VIX*), and counterparty risk (*PDCDS*).[9] In addition, we use net total hedge fund flow (*HF*) and *RV*, which are regarded as good proxies for the provision of arbitrage capital (see Mitchell & Pulvino, 2012).

Panel B of Table 3 reports mean values of market-wide impediment variables for the normal and crisis periods, where the normal period includes both pre- and postcrisis periods. Results show that funding cost and constraints,

[8] This is also consistent with Mitchell and Pulvino (2012), who find that the importance of firm-level impediments is conditional on the market conditions.

[9] Unreported results show a similar pattern for other market-level impediment variables. To avoid cluttering the table, we do not report these results here (available upon request).

TABLE 3 Impediments to arbitrage and market integration

Panel A. Mean Difference Tests Based on the Cross-Section of Firm Characteristics					
Period	**Variable**	**Noncointegrated**	**Cointegrated**	**Difference**	***t*-Statistic**
Precrisis	*Supply*	1.05	1.13	−0.07	−1.07
	Amihud	0.01	0.01	0.00	−0.32
	Leverage	0.39	0.46	−0.06	−0.70
	EDF	0.09	0.07	0.02	0.37
	Volatility	0.03	0.03	0.00	0.26
Crisis	*Supply*	1.09	1.08	0.02	0.79
	Amihud	0.04	0.02	0.02***	2.90
	Leverage	0.51	0.44	0.07*	1.81
	EDF	0.22	0.15	0.08*	1.77
	Volatility	0.04	0.03	0.01***	2.92
Postcrisis	*Supply*	1.00	1.11	−0.11**	−2.01
	Amihud	0.02	0.01	0.01	0.63
	Leverage	0.39	0.44	−0.05	−0.51
	EDF	0.03	0.09	−0.06	−1.58
	Volatility	0.02	0.03	−0.01	−0.95

Panel B. Mean Difference Tests of Market-level Impediments across Periods				
Variable	**Normal**	**Crisis**	**Difference**	***t*-statistic**
NOISE	0.21	0.76	0.54***	53.71
VIX	0.18	0.32	0.14***	36.06
PSB	0.01	−0.07	−0.07***	−13.31
PDCDS	0.79	1.28	0.50***	15.65
LOIS	0.21	0.81	0.60***	43.56
DPDPL	0.28	−0.36	−0.64***	−11.92
CF	−0.12	−0.74	−0.62***	−4.85
CSF	0.11	0.41	0.30***	3.10
HF	−3.30	−12.91	−16.21***	−3.09
RV	2.11	1.68	−0.43***	−3.69

Panel A reports mean difference tests on firm impediments that reject or accept the cointegration test. *Volatility* is monthly stock return volatility, *Amihud* is the Amihud daily individual bond illiquidity measure, *EDF* is the expected default rate calculated by the Merton model, *Leverage* is the book value of debt to the sum of book value of debt and market value of equity, and *Supply* is the number of CDS providers. Panel B reports tests of differences in the means of market-level impediment variables between the normal and crisis periods, where the normal period covers the whole sample period except the crisis period from July 2007 to June 2009. *NOISE* is the illiquidity index of Hu et al. (2013); *VIX* is the volatility index; *PSB* is Pastor-Stambaugh bond liquidity index; *PDCDS* is the average rate of the CDS against primary dealers; *LOIS* is the yield spread between Libor and OIS rates; *DPDPL* is the change in net long-term security holdings by primary dealers; and *CF* is the net convertible fund flow. *CSF* measures the cash shortfall of primary dealers using the method in Daniel et al. (2017). In addition, we perform mean difference tests for net total hedge fund inflows (*HF*) and repo volume (*RV*). The signs *, **, and *** indicate the 10%, 5%, and 1% levels of significance.

illiquidity, volatility, and counterparty risk are all significantly higher in the crisis period, confirming that the market-level impediments to arbitrage heighten during the crisis period. Consistent with Duffie (2010) and Mitchell and Pulvino (2012), we find substantial convertible hedge fund (*CF*) and hedge fund outflows (*HF*) during the crisis period, indicating that arbitrage capital is significantly reduced ($t = -4.85$ and -3.09, respectively) during the crisis period. Additionally, primary dealers' cash shortfalls (*CSF*) are substantially higher ($t = 3.10$) and *RV* drops dramatically during the crisis (t = –3.69). These results are consistent with the hypothesis that lower arbitrage capital provision leads to more severe violations of the arbitrage-free pricing relation during the crisis period.

3.2 | Nonparametric tests

Besides the cointegration tests, we conduct a model-free nonparametric test for market integration based on the concordance of price changes in the CDS and bond markets. There are advantages of using the nonparametric test to assess market integration. As this test method is not conditional on a particular specification for price processes, it is free from the model specification bias. Additionally, the measure accounts for all possible pairs across all price observations, and it is independent of the horizon over which spreads are observed.

When the CDS basis is significantly negative, arbitrageurs can exploit profitable opportunities by buying a bond and a CDS simultaneously. Thus, when there is an arbitrage opportunity, a pair of CDS and reference bond spreads should move in the opposite direction. For example, when there is a negative basis, the arbitrageur's activity will push the bond yield spread (*BYS*) down and drive the CDS spread up, and eventually close the gap to eliminate the profitable opportunity. If this is not the case and profitable opportunities remain, the no-arbitrage condition is violated. Thus, we can assess market (dis)integration based on the frequency of arbitrage opportunities (or spread changes in the same direction).[10] For a given period with T observations, we can define γ_i for firm I as[11]:

$$\gamma_i = \sum_{\tau=1}^{T-1} \sum_{k=1}^{T-\tau} 1_{[\Delta CDS^{\tau}_{i,k} \Delta BYS^{\tau}_{i,k} > 0]}, \tag{2}$$

where $\Delta CDS^{\tau}_{i,k} = CDS_i(k+\tau) - CDS_i(k)$and $\Delta BYS^{\tau}_{i,k} = BYS_i(k+\tau) - BYS_i(k)$, $1 \leq k \leq T - \tau$, $1 \leq \tau \leq T - 1$, τ indicates the time interval to calculate price changes, and bond yield spreads *BYS* are par-equivalent CDS spreads (*PECDS*). Given T, the bond and CDS markets are more integrated for firm I than for firm j, if $\gamma_i < \gamma_j$. γ is related to the Kendall correlation κ in that $\kappa = \frac{4\gamma}{T(T-1)} - 1$. When there is no mispricing, $\kappa = -1$. More generally, a larger κ implies a less integrated market. In our empirical investigation, we calculate γ_i for nonoverlapping intervals τ = 5, 10, 25, and 50 (or weekly, biweekly, monthly, and bimonthly, respectively).[12]

Table 4 reports the results of nonparametric tests for pricing discrepancies over different time horizons using daily data. At the bottom of each panel, we report t-values for the differences in pricing discrepancies (or convergence) between 5- and 50-day intervals. For example, for the analysis based on the whole sample in the precrisis period (the second set of results in the top panel), the proportion of pricing divergences to total observations is 47.59% at the 5-day interval and 40.16% at the 50-day interval; the difference is –7.42%, which is significant at the 1% level. Over the same period, the proportion of comovements of spreads converging in the right direction is 46.48% at the 5-day interval and 59.31% at the 50-day interval; the difference is 12.83%, which is significant at the 1% level. These results suggest that prices converge in the right direction. However, pricing discrepancies appear to be quite persistent as there is still a high percentage of comovements representing arbitrage opportunities at the 50-day interval.

For the crisis period, we find that pricing discrepancies are much more serious. For example, the proportions of pricing divergences ($\Delta CDS^* \Delta BYS > 0$) are 51.27% at the 5-day interval and 46.05% at the 50-day interval, which are all higher than those for the precrisis periods. The difference in pricing divergences between 5 and 50 days is –5.22%, which is significant at the 1% level.

An interesting finding is that the pricing discrepancies are also quite persistent in the postcrisis period. The proportion of pricing divergences is 47.53% at the 5-day interval and 47.07% at the 50-day interval. Turning to the convergence column, we find that the proportion of price movements in the right direction is 49.45% at the 5-day interval

[10] Bakshi, Cao, and Chen (2000) and Kapadia and Pu (2012) use a similar approach to study market relations.

[11] See Kapadia and Pu (2012) for a similar definition. In the present case, we examine the integration between CDS and bond yield spreads, instead of CDS spreads and stock prices.

[12] That is, for each τ, we calculate $\sum_{k=1}^{T-\tau} 1_{[\Delta CDS^{\tau}_{i,k} \Delta BYS^{\tau}_{i,k} > 0]}$ to assess the extent of market disintegration. We then report the proportion of comovements in CDS and bond yield spreads converging in the right direction or not.

TABLE 4 Pricing discrepancies

	Interval (Days)	The full sample period: Fraction (%) $\Delta CDS^* \Delta BYS < 0$	The full sample period: Fraction (%) $\Delta CDS^* \Delta BYS > 0$	The full sample period: Fraction (%) $\Delta CDS^* \Delta BYS = 0$	The precrisis period: Fraction (%) $\Delta CDS^* \Delta BYS < 0$	The precrisis period: Fraction (%) $\Delta CDS^* \Delta BYS > 0$	The precrisis period: Fraction (%) $\Delta CDS^* \Delta BYS = 0$	The crisis period: Fraction (%) $\Delta CDS^* \Delta BYS < 0$	The crisis period: Fraction (%) $\Delta CDS^* \Delta BYS > 0$	The crisis period: Fraction (%) $\Delta CDS^* \Delta BYS = 0$	The postcrisis period: Fraction (%) $\Delta CDS^* \Delta BYS < 0$	The postcrisis period: Fraction (%) $\Delta CDS^* \Delta BYS > 0$	The postcrisis period: Fraction (%) $\Delta CDS^* \Delta BYS = 0$
All	5	44.75	50.39	4.86	46.48	47.59	5.93	43.60	51.27	5.13	49.45	47.53	3.02
	10	45.12	51.90	2.98	49.22	47.75	3.03	44.15	52.56	3.28	49.31	48.46	2.23
	25	46.27	52.28	1.45	54.15	44.69	1.17	47.09	51.45	1.46	50.16	48.54	1.30
	50	47.65	51.58	0.77	59.31	40.16	0.53	53.34	46.05	0.61	52.17	47.07	0.76
	50–5	2.90***	1.18**	−4.09***	12.83***	−7.42***	−5.41***	9.74***	−5.22***	−4.53***	2.71***	−0.46	−2.26***
	t	(4.49)	(2.04)	(−11.40)	(16.49)	(−10.44)	(−12.26)	(9.43)	(−5.51)	(−8.60)	(4.07)	(−0.80)	(−6.61)
IG	5	46.41	49.54	4.05	47.87	46.90	5.24	45.65	50.51	3.85	50.29	47.49	2.22
	10	46.93	50.65	2.42	50.86	46.49	2.65	46.05	51.84	2.12	50.31	48.06	1.63
	25	47.79	51.06	1.15	55.59	43.35	1.06	48.00	51.31	0.68	51.05	47.96	0.99
	50	48.92	50.42	0.66	61.13	38.34	0.53	53.74	46.03	0.23	52.61	46.70	0.69
	50–5	2.51***	0.88	−3.39***	13.27***	−8.56***	−4.71***	8.09***	−4.47***	−3.62***	2.32***	−0.78	−1.54***
	t	(3.29)	(1.28)	(−8.70)	(14.75)	(−10.20)	(−9.67)	(6.12)	(−3.68)	(−6.27)	(3.07)	(−1.13)	(−5.48)
SG	5	41.99	51.80	6.20	44.20	48.72	7.08	40.28	52.50	7.21	47.87	47.60	4.53
	10	42.14	53.96	3.90	46.53	49.81	3.66	41.10	53.74	5.17	47.43	49.21	3.36
	25	43.75	54.30	1.94	51.77	46.89	1.34	45.62	51.66	2.72	48.48	49.65	1.88
	50	45.56	53.49	0.95	56.31	43.16	0.52	52.70	46.09	1.22	51.33	47.76	0.91
	50–5	3.57***	1.69	−5.25***	12.11***	−5.56***	−6.56***	12.42***	−6.42***	−6.00***	3.46***	0.16	−3.61***
	t	(3.05)	(1.62)	(−7.54)	(8.46)	(−4.38)	(−7.76)	(7.55)	(−4.25)	(−5.95)	(2.66)	(0.16)	(−4.38)

This table reports the comovements (γ_i) of CDS and corporate bond spreads as a proportion of total observations. The results are reported for the whole sample (All), and for investment-grade (IG) and speculative-grade (SG) categories. We calculate the fraction of comovement of CDS and bond yield spreads in each direction, using all pairs of observations over nonoverlapping intervals of 5, 10, 25, and 50 days, respectively. The sample period is from July 2002 to December 2016. Comovements are pricing discrepancies that represent arbitrage opportunities if $\Delta CDS^* \Delta BYS > 0$. The precrisis period is from July 2002 to June 2007, the crisis period is from July 2007 to June 2009, and the postcrisis period is from July 2009 to December 2016. The t value for the test of the difference between the proportions at the 5- and 50-day intervals is reported at the bottom row for each sample. The signs *, **, and *** indicate the 10%, 5%, and 1% levels of significance.

and 52.17% at the 50-day interval; the difference is 2.71%, which is significant at the 1% level. Again, bond and CDS spreads are converging but at a lower speed in the postcrisis period than in the precrisis period. This finding is consistent with the pattern in Figure 1 that the CDS basis remains negative in the postcrisis period, though the magnitude is much milder than in the crisis period.

When the sample is divided into IG and SG groups, the pricing discrepancies are more serious for riskier bonds. In addition, the proportion of comovements of CDS and bond spreads in the right direction is much lower for SG bonds. These results show that pricing discrepancies are more persistent for SG bonds. The higher risk of low-grade bonds can make arbitrage riskier and result in higher funding constraints (e.g., haircuts) that reduce the capital available for undertaking arbitrage to eliminate pricing discrepancies. This may explain why there are more frequent and persistent pricing discrepancies for junk bonds.

The proportion of price divergences increases during the crisis period and is much higher for the SG bonds. Consistent with the cointegration test, results show that markets are less integrated during the crisis period. Arbitrage opportunities are more frequent and pricing discrepancies are more persistent for riskier bonds. These findings support the predictions of limited-arbitrage theory (Shleifer & Vishny, 1997), which suggest that arbitrage is riskier in an uncertain market and for risky securities, and ineffective arbitrage leads to more frequent violations and persistent price discrepancies.

To establish a link between pricing discrepancies and arbitrage frictions, we examine firm characteristics for each CDS/bond pair. If limited arbitrage plays a significant role, firms with high impediments to arbitrage will more likely experience pricing discrepancies. Table 5 reports the results of regressions of the nonparametric pricing divergence measure, or Kendall's κ, against each impediment variable over different intervals τ. For the whole sample period, firm-level impediment variables are generally significant and their coefficients are of predicted signs. The average adjusted R^2 is around 15%. Results show that the extent of market disintegration increases with firm default risk (*EDF*), stock return volatility, leverage, and bond illiquidity (*Amihud*), and decreases with CDS liquidity (*Supply*). When dividing the sample into subperiods, we find a similar pattern for each subperiod and stronger results for the crisis period. This finding is consistent with cointegration tests, and confirms that impediments to arbitrage are important factors contributing to unusual price discrepancies in the crisis period. Several impediment variables are also significant in the normal periods, indicating that limited arbitrage has a real effect not just in times of stress.

In summary, we find that firms with high impediments to arbitrage are more likely to experience pricing discrepancies. Moreover, pricing discrepancies are more severe and markets are less integrated when market-level arbitrage frictions are elevated during the crisis. The nonparametric method shows a clearer pattern of frequent short-term price violations in credit markets. Collectively, there is compelling evidence that pricing discrepancies and violations of market integrations are closely linked to market-wide and firm-level impediments to arbitrage.

4 | PERSISTENCE IN PRICING VIOLATIONS

Misalignment in credit markets is marked not only by extreme price discrepancies but also by their unusually high persistence. Investigating the causes of persistence in pricing discrepancies will thus shed more light on the sources of mispricing. For example, slow movement in investment capital to trading opportunities can contribute to persistence in pricing discrepancies (Duffie, 2010). Besides funding constraints, illiquidity and market uncertainty (Pontiff, 2006) can increase the cost of undertaking arbitrage and prolong price violations. In this section, we propose a persistence measure based on the long memory model and explore the link between persistence in pricing discrepancies and the proxy variables for slow-moving capital and arbitrage impediments that potentially contribute to persistent basis violations.

4.1 | Persistence in CDS-bond pricing discrepancies

The persistence behavior of the CDS basis can be captured by long-range dependency or long memory. When a financial series exhibits long memory, the autocorrelation between two observations k periods apart decays in a slower

TABLE 5 Regressions for the Kendall ratio

Period	Horizons	*Supply*	*Amihud*	*Leverage*	*EDF*	*Volatility*
Whole	$\tau = 5$	-0.04^{***}	0.17	0.09^{***}	0.13^{***}	1.34^{***}
		(−3.39)	(0.63)	(6.02)	(4.90)	(5.87)
	$\tau = 10$	-0.03^{*}	0.39	0.14^{***}	0.19^{***}	1.94^{***}
		(−1.94)	(1.29)	(8.37)	(6.37)	(7.63)
	$\tau = 25$	-0.03^{*}	0.81^{**}	0.19^{***}	0.22^{***}	2.27^{***}
		(−1.79)	(2.06)	(8.54)	(5.62)	(6.79)
	$\tau = 50$	−0.04	0.95^{*}	0.22^{***}	0.25^{***}	2.66^{***}
		(−1.58)	(1.90)	(7.88)	(4.97)	(6.17)
Precrisis	$\tau = 5$	-0.07^{***}	−0.08	0.08^{***}	0.19^{***}	2.02^{***}
		(−2.93)	(−0.19)	(2.87)	(4.01)	(3.18)
	$\tau = 10$	-0.05^{*}	−0.18	0.11^{***}	0.25^{***}	2.97^{***}
		(−1.94)	(−0.39)	(4.12)	(4.91)	(4.51)
	$\tau = 25$	−0.03	0.09	0.15^{***}	0.35^{***}	4.46^{***}
		(−0.81)	(0.16)	(4.46)	(5.75)	(5.72)
	$\tau = 50$	0.00	0.29	0.19^{***}	0.37^{***}	4.92^{***}
		(−0.12)	(0.45)	(4.99)	(5.34)	(5.45)
Crisis	$\tau = 5$	-0.08^{***}	0.65^{**}	0.07^{**}	0.07^{***}	0.39^{**}
		(−2.87)	(2.17)	(2.54)	(2.64)	(2.27)
	$\tau = 10$	-0.07^{***}	0.64^{**}	0.10^{***}	0.09^{***}	0.46^{**}
		(−2.70)	(2.20)	(3.56)	(3.39)	(2.53)
	$\tau = 25$	-0.04^{**}	0.42^{**}	0.11^{***}	0.08^{**}	0.24^{*}
		(−2.38)	(2.22)	(3.38)	(2.54)	(1.70)
	$\tau = 50$	-0.07^{*}	0.22^{**}	0.11^{***}	0.10^{***}	0.51^{**}
		(−1.76)	(2.52)	(2.92)	(2.81)	(2.22)
Postcrisis	$\tau = 5$	-0.05^{***}	−0.53	0.14^{***}	0.39^{***}	2.16^{***}
		(−3.20)	(−1.30)	(7.00)	(5.50)	(5.21)
	$\tau = 10$	-0.05^{***}	−0.58	0.18^{***}	0.47^{***}	2.80^{***}
		(−2.95)	(−1.21)	(7.76)	(5.78)	(5.82)
	$\tau = 25$	-0.05^{**}	0.02	0.22^{***}	0.55^{***}	3.76^{***}
		(−2.05)	(0.04)	(7.72)	(5.33)	(6.37)
	$\tau = 50$	-0.04^{**}	0.53	0.28^{***}	0.63^{***}	4.76^{***}
		(−2.40)	(0.71)	(7.59)	(4.81)	(6.40)

This table reports coefficient estimates of the regression of Kendall κ over different intervals τ against each explanatory variable for the whole sample period, as well as for the normal and crisis periods. The precrisis period is from July 2002 to June 2007, the crisis period is from July 2007 to June 2009, and the postcrisis period is from July 2009 to December 2016. The explanatory variables include various firm-level impediments to arbitrage. *Supply* is the number of CDS suppliers, *Amihud* is the Amihud individual bond illiquidity measure, *Volatility* is the firm's stock return volatility calculated based on the daily stock return of the month, *EDF* is the expected default probability calculated by the Merton model, and *Leverage* is the ratio of the book value of debt to the sum of book value of debt and market value of equity. The *t* values are in parentheses, and *, **, and *** indicate the 10%, 5%, and 1% levels of significance, respectively.

hyperbolic rate of k^{2d-1}, where d is the long memory parameter, as opposed to the faster exponential rate of a^k for a short memory time series.[13] A persistent time series does not have summable autocorrelations, and it is the slow declining rate of autocorrelations that generates the behavior of persistence.[14]

There are a number of ways to estimate the long memory parameter d. The choice of an estimation method depends on the time-series property of a financial series. We find that the time series of the CDS basis is near-nonstationary. For this type of series, it is more appropriate to use a fractionally integrated ARFIMA (p, d, q) process to estimate d. Appendix B provides the details of the estimation procedure.

4.2 | Persistence in CDS basis volatility

Arbitrage frictions can not only cause persistence in the level, but also the change in the CDS basis (or volatility), as basis changes tend to move in the same direction and cluster. A simple volatility measure is the absolute basis change, and for this volatility variable, we can estimate its persistence by ARFIMA. A more desirable method to analyze volatility persistence utilizes the GARCH framework, as basis changes exhibit pronounced volatility clustering. We employ the frictionally integrated GARCH model (FIGARCH) proposed by Baillie, Bollerslev, and Mikkelsen (1996) to measure volatility persistence through the long memory parameter d. Consider an MA(1)-GARCH(1,1) model for the basis change:

$$\Delta basis_t = c_1 + e_t - \theta_1 e_{t-1}, \tag{3}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{4}$$

where the conditional variance of the error term e_t is σ_t^2.This model captures the first-order autocorrelation and persistence in volatility and can be easily extended to higher orders of dependence.

Let $v_t = e_t^2 - \sigma_t^2$; then we can rewrite the above conditional variance equation as:

$$[1 - (\alpha_1 + \beta_1)B]e_t^2 = \alpha_0 + (1 - \beta_1 B)v_t,$$

where B is the back-shift operator. The FIGARCH model introduces a long memory parameter d into the squared residuals equation:[15]

$$[1 - (\alpha_1 + \beta_1)B](1 - B)^d e_t^2 = \alpha_0 + (1 - \beta_1 B)v_t, \tag{5}$$

or,

$$\beta(B)\sigma_t^2 = \alpha_0 + [\beta(B) - \phi(B)(1 - B)^d]e_t^2, \tag{6}$$

where φ(B)=1 − (α_1+β_1)B, $\beta(B) = 1 - \beta_1 B$, and parameter d associated with the shock give rise to persistent conditional volatility.

---

[13] The parameter a is a constant describing the form that the autocorrelation function (ACF) decays at an exponential rate of a number to the power k.

[14] Consider a simple long memory time series model $(1 - B)^d x_t = e_t$, where B is the back-shift operator, e_t is a stationary ARMA process, and x is the time series, where d is a real number between −0.5 and 0.5. This fractional difference series exhibits persistent behavior that can be shown as follows. Taking a binomial expansion of $(1 - B)^d x_t$ gives: $\sum_{j=0}^{\infty} \varphi_j x_{t-j} = e_t$, where $\varphi_0 = 1$ and $\varphi_j = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}$ is the binomial coefficient, and $\varphi_j \approx \frac{j^{-d-1}}{\Gamma(-d)}$ for large j. Thus, the model can be considered an AR model of an infinite order with slow decaying coefficients. Such dependency of the series on its infinite past is the basis for its persistence behavior. Furthermore, it can be shown that the lag k autocorrelation of x_t decays in a hyperbolic rate, $\rho_k = O(k^{2d-1})$, and that the autocorrelations are not summable (see Hosking, 1981). For a time series which follows a fractionally integrated ARFIMA (p, d, q) model, the time series is stationary when d is between −0.5 and 0.5.

[15] Though using the same notation, we note that this d is in a different model form and hence has different property from that of the ARFIMA model. For instance, Baillie et al. (1996) show that the model is not weakly stationary but is strictly stationary and ergodic when d is between 0 and 1.

TABLE 6 Persistence in the CDS basis and funding-related variables

Panel A. Persistence in CDS Basis

	Precrisis	Crisis	Postcrisis
|ΔBasis|			
Investment grade	0.06	0.35	0.10
Speculative grade	0.10	0.46	0.12
Basis			
Investment grade	0.06	0.40	0.08
Speculative grade	0.08	0.41	0.10
|Basis|			
Investment grade	0.06	0.44	0.09
Speculative grade	0.08	0.46	0.11

Panel B. FIGARCH Estimates of Volatility Persistence in CDS Basis

Bonds	Period	c_1	θ_1	α_0	β_1		
Investment grade	Precrisis	0.03	–0.60***	3.89	0.39*	0.23	0.15**
		(0.27)	(0.00)	(0.12)	(0.09)	(0.18)	(0.02)
	Crisis	–0.25	–0.38***	7.55**	0.64***	0.41***	0.40***
		(0.22)	(0.00)	(0.04)	(0.00)	(0.01)	(0.00)
	Postcrisis	0.01	–0.45***	5.46***	0.34***	0.45***	0.11***
		(0.45)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Speculative grade	Precrisis	0.25**	–0.69***	13.75**	0.51***	0.41***	0.28***
		(0.04)	(0.00)	(0.02)	(0.00)	(0.01)	(0.00)
	Crisis	–0.41	–0.31***	13.90*	0.86***	0.17	0.91***
		(0.30)	(0.00)	(0.10)	(0.00)	(0.23)	(0.01)
	Postcrisis	0.12	–0.49***	33.32***	–0.43***	–0.33**	0.25***
		(0.17)	(0.00)	(0.00)	(0.01)	(0.04)	(0.00)

Panel C. Persistence in Funding-Related Variables

	Precrisis	Crisis	Postcrisis
Dealers' financing	0.34	0.42	0.32
Failures-to-deliver by dealers	0.16	0.30	0.29

Panel A reports the estimates of persistence (d) for the CDS basis, absolute basis, and absolute basis changes. Parameters are estimated using the ARFIMA method. The estimation procedure is described in Appendix B. Panel B reports the MA(1)-FIGARCH(1,1) estimates for $\Delta basis$of investment-grade (IG) bonds and speculative-grade (SG) bonds. The model is $\Delta basis_t = c_1 + e_t - \theta_1 e_{t-1}$, where $(1-(\alpha_1+\beta_1)B)(1-B)^d e_t^2 = \alpha_0 + (1-\beta_1 B)v_t$,where$v_t = e_t^2 - \sigma_t^2$and α_0 is the constant term in the conditional variance Equation (6). The p-values are given in parentheses. The signs *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Panel C reports the estimates of persistence (d) for primary dealers' financing and failures-to-deliver. All estimates are obtained for three subperiods. The precrisis period is from July 2002 to June 2007, the crisis period is from July 2007 to June 2009, and the postcrisis period is from July 2009 to December 2016.

4.3 | Estimation of basis level and volatility persistence

We estimate persistence parameter d for the CDS basis, absolute value of the basis (|basis|), and absolute value of basis changes (|Δbasis|) first using the daily CDS basis series averaged across all firms in the categories of IG and SG, respectively, to come up with measures to capture the market-wide persistence in the basis. Panel A of Table 6 reports the estimates of persistence parameter d by bond grade and period (normal vs. crisis) using the ARFIMA model. The first set of results provides estimates for the absolute basis changes of the CDS basis index. Results show that basis changes are much more persistent during the crisis period, and the persistence is much stronger for SG bonds

($d = 0.46$) than for IG bonds ($d = 0.35$). The second set of results includes d estimates for the CDS basis level. Results again show that the CDS basis is highly persistent during the crisis period and is more persistent for SG bonds. The third set of results in Panel A includes persistence estimates for the absolute basis, which show a similar pattern of persistence.

An intuitive way to understand the magnitude of d is to translate its value into the number of days that the basis will persist using the impulse response function (IRF). The higher the value of d, the longer the basis will persist. Appendix B provides examples for translating the d value into the number of days that the basis persists. As shown in these examples, a relatively small difference in the d value can result in a fairly large difference in the half-life, in terms of the number of days.

The above results show that the CDS basis is persistent in both level and changes, and is highly persistent during the crisis period. The higher persistence in both level and changes of the CDS basis coincides with the period of greater market-level impediments to arbitrage and shortage of arbitrage capital. As shown in Panel B of Table 3, market-level impediments are significantly higher during the crisis period, suggesting that systematic impediments play a role in the CDS basis persistence. Another interesting finding is that although persistence is substantially weakened after the subprime crisis, the persistence of the CDS basis in the postcrisis period is higher than that in the precrisis period.

Panel B of Table 6 reports estimates of the MA(1)-FIGARCH(1,1) model for basis changes of IG and SG bonds for both the normal and crisis periods. Again, we use the daily basis averaged across all firms in each rating group in this estimation. Results show that in normal times, the CDS basis for both IG and SG bonds do not show high volatility persistence as reflected in small d estimates. In contrast, during the crisis period, the CDS basis becomes quite persistent, particularly for SG bonds ($d = 0.91$). Results strongly suggest that changes in the basis become highly persistent during the crisis period. The higher persistence in basis changes coincides with higher market-wide impediments to arbitrage during the crisis period, again suggesting a close link between persistent pricing discrepancies and limits to arbitrage.

In summary, price discrepancies are more persistent in times of stress and for risky firms. These findings are consistent with the hypothesis that greater impediments to arbitrage for risky firms (Shleifer & Vishny, 1997) and in turbulent times (Duffie, 2010) lead to persistent pricing discrepancies over an extended period. To substantiate this hypothesis, we estimate the persistence parameter d for each firm and link the persistent basis behavior to variables of impediments to arbitrage.

4.4 | Persistence in pricing violations and firm-level impediments to arbitrage

To investigate the role of limited arbitrage in persistent pricing violations, we run the cross-sectional regression of the persistence measure (d) for individual firms against firm-level impediment variables for the whole sample period and different subperiods. The parameter d is estimated using weekly CDS basis data.[16] According to the invariance principle of persistence (see Man & Tiao, 2006; Tsai & Chan, 2005), d estimates are independent of time aggregation; that is, the weekly d estimate should theoretically be identical to the daily estimate.[17] Thus, using weekly data should preserve the information in the basis persistence measure d.

Table 7 reports the results of cross-sectional regressions of persistence (d) in the CDS basis level and absolute changes. The first five columns report the regression estimates against each firm-level impediment-to-arbitrage variable. For the whole sample period, most impediment variables are significant at the conventional level and their coefficients are of the expected sign, indicating that CDS basis level and change are more persistent for firms with high return volatility, illiquidity, default risk, and leverage. When dividing the sample period into subperiods, we find that the relation is much stronger during the crisis period. Overall, the results strongly suggest that the persistence in pricing discrepancies is closely related to the variables of firm-level impediments to arbitrage.

[16] Using weekly data mitigates the effects of infrequent trading for some corporate bonds.

[17] It is well-known that the long memory parameter d is invariant to time aggregation (Tsai & Chan, 2005). Theoretically, estimates based on daily and weekly data should be identical. Using weekly intervals enables us to increase the sample size and the power of tests.

TABLE 7 Cross-sectional basis persistence on firm-level variables

d	Period	*Supply*	*Amihud*	*Leverage*	*EDF*	*Volatility*	$Beta_{DF}$	$Beta_{Fails}$	$Beta_{HF}$
|ΔBasis|	Whole	0.02	1.98***	0.12***	0.31**	2.63***	0.05***	0.08	0.09***
		(0.71)	(3.97)	(3.57)	(2.10)	(2.86)	(3.62)	(1.31)	(5.21)
	Precrisis	0.00	1.92**	0.04	0.57	−4.81	0.04	0.06	0.01
		(0.02)	(2.24)	(0.55)	(0.34)	(−1.16)	(1.48)	(0.73)	(0.20)
	Crisis	0.01	1.15***	0.13***	0.17***	2.35***	0.09***	0.06**	0.07***
		(0.15)	(4.14)	(2.59)	(3.00)	(3.15)	(2.61)	(2.48)	(3.30)
	Postcrisis	0.04	0.10	0.09**	0.46**	1.95*	0.04*	0.03	0.10*
		(1.07)	(0.13)	(2.42)	(2.14)	(1.67)	(1.81)	(0.60)	(1.67)
Basis	Whole	0.01	0.46**	0.01	0.14**	0.04	0.03**	0.04	0.06***
		(0.85)	(2.16)	(0.84)	(2.25)	(0.10)	(2.11)	(0.81)	(4.27)
	Precrisis	−0.02	0.12	0.00	−0.01	−0.11	0.00	0.16*	0.05
		(−0.91)	(1.41)	(0.13)	(−0.35)	(−0.47)	(0.13)	(1.72)	(1.02)
	Crisis	−0.01	0.32*	0.07***	0.38**	1.23*	0.10***	0.04**	0.07***
		(−0.52)	(1.69)	(2.96)	(2.56)	(1.79)	(3.06)	(2.07)	(3.69)
	Postcrisis	0.03	0.23	0.01	0.05	0.33	0.04**	−0.12	−0.12**
		(1.54)	(0.69)	(0.36)	(0.60)	(0.61)	(1.98)	(−1.61)	(−2.06)

This table reports the results of the cross-sectional regressions of the CDS basis persistence measure (*d*). The dependent variables of cross-sectional regressions are the persistence measures of CDS basis and absolute basis changes (|ΔBasis|), respectively. The first five columns report regressions on five firm-level variables of impediments to arbitrage: *Supply*, *Amihud*, *Leverage*, *EDF*, and *Volatility*, while the last three columns report estimates of regressions on the betas of CDS basis associated with dealers' financing (*DF*), failures to deliver (*Fails*), and hedge fund flow (*HF*). Results are reported for the whole sample period and three subperiods. The precrisis period is from July 2002 to June 2007, the crisis period is from July 2007 to June 2009, and the postcrisis period is from July 2009 to December 2016. The *t*-values are in parentheses, and *, **, and *** indicate the 10%, 5%, and 1% levels of significance, respectively.

4.5 | Limited arbitrage capital provision and persistent pricing discrepancies

One unique feature of the subprime crisis was the drastic deterioration of liquidity and credit quality in financial markets. Many large financial institutions incurred substantial portfolio losses and liquidity suddenly dried up. Failure to regain capital impaired the ability of financial intermediaries to absorb supply shocks. A number of studies have suggested that slow-moving capital has aggravated price distortions in credit markets (see Duffie, 2010; Mitchell & Pulvino, 2012; Mitchell, Pedersen, & Pulvino, 2007). Theoretical models (see, for example, Duffie, 2010; Garleanu & Pedersen, 2011) have been developed to show that funding constraints and limited provision of arbitrage capital lead to large and persistent pricing violations in credit markets. However, there are relatively few empirical tests on this theoretical implication. In this section, we explore the role of limited arbitrage capital provision in persistent pricing discrepancies.

We approach this issue from two angles. First, if limited arbitrage capital provision is the primary cause for persistent pricing discrepancies, we should observe a pattern of persistence in this variable, as in the CDS basis. A finding of a similar pattern of persistence in the shortage of arbitrage capital provision will therefore provide plausible evidence that persistent pricing violations are linked to slow-moving capital. Second, if limited capital provision is the driving force for pricing discrepancies, the CDS-bond price misalignment should be more persistent for firms that are more susceptible to the shortage of arbitrage capital.

Using several proxy variables (e.g., net convertible fund, total hedge fund flow, and primary dealers' cash shortfall), we have already shown that arbitrage capital provision is significantly lower during the financial crisis (see Panel B of Table 3). This finding lends support to the argument that the shortage of arbitrage capital contributes to pricing discrepancies. The persistence in pricing discrepancies should therefore be linked to persistence in the shortage of arbitrage capital. Put differently, persistence in arbitrage capital shortage leads to persistence in pricing discrepancies. Moreover, in reality, the susceptibility of CDS-bond pricing discrepancies to arbitrage frictions is likely to vary across firms. Firms that are more sensitive to arbitrage frictions should experience greater persistence in pricing discrepancies when the shortage of arbitrage capital is persistent.[18] Thus, we can perform a cross-sectional test on the effect of limited arbitrage capital provision on the persistence of firm-level pricing discrepancies.

To estimate the long memory parameter *d*, we need higher-frequency data for the proxy of arbitrage capital provision. Variables such as cash shortfall and net hedge fund flow that we used earlier are not desirable, as they are only available at quarterly or monthly intervals. These lower-frequency data are not suitable to estimate the persistence parameter in the long memory model.[19] To overcome this difficulty, we use dealers' funding and failures to deliver as alternative proxies for capital provision, which are available at weekly intervals and have higher frequency. Dealers' financing includes the total amount of funds obtained by primary dealers through repo and other financing channels. This variable should be a good proxy for capital provision to dealers. Failure-to-deliver (*Fails*) is the amount of failures to deliver collateral in the repo market. Higher failures to deliver collateral signify a shortage of dealers' capital. The literature has shown that persistent settlement failures increase the risk in the funding market, which reduces the willingness of lenders to supply capital (see Liu & Wu, 2017). Dealers' financing and delivery failure data come from the FRB of New York. Figure 3 displays the time series of these variables.

We use the same method to estimate the persistence of these funding-related variables as in the CDS basis investigation. Panel C of Table 6 reports the estimates of persistence parameter *d* for the funding-related variables. Results show that these variables exhibit high persistence during the crisis period. For example, the *d* value of dealers' financing is 0.42, indicating high persistence in the shortage of capital provision to primary dealers during the crisis period. A similar pattern is found for failures to delivery: the persistence of this variable is also much higher in the crisis period than in the pre- and postcrisis periods. Importantly, the pattern of these persistence parameter estimates is similar to that for the CDS basis reported in Panel A of Table 6. The results suggest that persistence in the CDS-bond pricing discrepancies is tied to persistence in the shortage of arbitrage capital during the financial crisis.

To perform cross-sectional regression tests, we estimate the sensitivity (beta) of a firm's CDS basis to the funding-related variables. As mentioned above, the CDS basis should be more persistent when the firm is more sensitive or vulnerable to shortage in arbitrage capital provision. We estimate the betas associated with three key funding variables: dealers' financing, failures to delivery, and hedge fund flow (net), and use them to explain the persistence in pricing discrepancies for individual firms. The last three columns of Table 7 show the results of regressing the firm's CDS basis persistence parameter (*d*) against each beta. As shown, the coefficients of betas are quite significant during the crisis period. Firms with higher betas of funding-related variables have higher *d* value or higher persistence in pricing discrepancies. These results reinforce the finding of persistence in funding-related variables above, and suggest that the persistence in firm-level CDS-bond pricing discrepancies during the financial crisis is linked to persistence in limited arbitrage capital provision. While systematic impediments are important factors contributing to persistent pricing discrepancies in times of stress, they are a lesser concern during the normal period. As shown in Table 7, the coefficients of betas associated with aggregate capital provision are less significant and smaller during the normal period. In contrast, firm-level impediment variables such as illiquidity, volatility, and default risk (EDF) are often significant and have larger coefficients than betas in the pre- or postcrisis periods, suggesting that they play a bigger role in normal times.[20]

[18] Funding provision is a good variable to assess the impact of limited arbitrage capital on CDS basis persistence. However, this variable is not available at the firm level. To the extent that funding is important for closing the CDS-bond pricing discrepancy, this effect can be captured by the sensitivity of the CDS basis to the capital provision variables.

[19] One difficulty is that the number of observations is low. For example, if we use hedge fund flow, there are only 24 monthly observations available over the course of the crisis period. Using cash shortfall is even worse, as there are only 8 quarterly observations, which is not feasible for estimating *d* in the long memory model.

[20] The R^2 values show a similar pattern, indicating that firm-level impediments explain about twice as large the variation in *d* as betas.

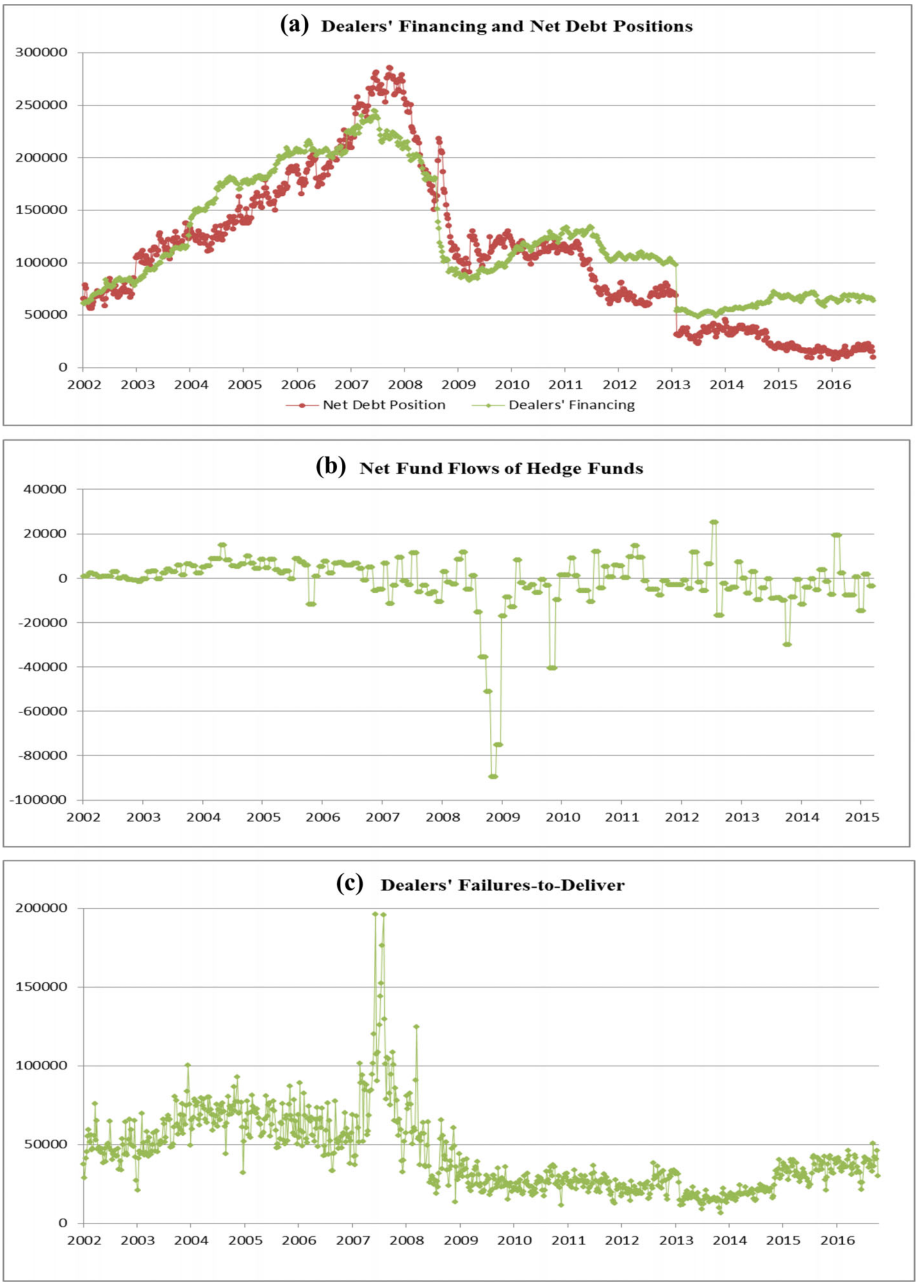

FIGURE 3 The Time Series of Funding-Related Variables. Panel A shows primary dealers' financing (square) and net debt or inventory positions (circle), panel B shows net fund flow of hedge funds, and panel C displays dealers' failures-to-deliver. All units are in millions

4.6 | The CDS basis in the postcrisis period

As shown earlier, the CDS basis remained negative after the subprime crisis. An important question is what may have caused the negative CDS basis in the postcrisis era when the arbitrage capital constraint seems unbinding.

After the financial crisis, there were several reforms in bank regulations. The postcrisis reforms in financial market regulations, such as the Dodd-Frank Act, the Volcker Rule, and Basel III requirements, have affected overall capital

TABLE 8 Determinants of the CDS-bond basis

	Precrisis		Crisis		Postcrisis	
	(1)	(2)	(1)	(2)	(1)	(2)
Dealers' net positions	0.07	0.10***	−0.11	−0.02	0.36***	0.63***
	(1.53)	(2.61)	(−0.72)	(−0.16)	(4.81)	(3.75)
Failures-to-delivery		0.07**		0.01		−0.59***
		(2.51)		(0.45)		(−3.66)
Dealers' financing		0.31***		1.64**		−0.22
		(3.80)		(2.49)		(−1.26)
Hedge fund flow	0.01	0.01	0.36***	0.26***	−0.02	0.03
	(0.07)	(0.37)	(4.21)	(3.23)	(−0.92)	(1.00)
Bond return volatility	−0.21*	−0.18	−0.68***	−0.65***	−0.30***	−0.29***
	(−1.85)	(−1.59)	(−6.05)	(−4.45)	(−7.49)	(−7.50)
Amihud illiquidity	−0.03	−0.02	0.08	0.01	0.06	−0.01
	(−1.57)	(−1.05)	(1.57)	(0.20)	(1.14)	(−0.08)
R^2	0.33	0.56	0.86	0.90	0.34	0.47

This table reports the results of weekly regressions for aggregate CDS basis based on swap rates. The regressors include primary bond dealers' net debt positions (inventories), failures-to-delivery and financing, hedge fund flow, bond return volatility, and Amihud illiquidity measure. The net debt positions, failures-to-delivery, and financing by primary dealers are scaled by total corporate bond supply. Hedge fund flow is scaled by hedge fund size. Bond return volatility is the mean of aggregate bond volatility calculated using sample data from the past six months. All independent variables are normalized to have zero mean and unit variance. The precrisis period is from July 2002 to June 2007, the crisis period is from July 2007 to June 2009, and the postcrisis period is from July 2009 to December 2016. The Newey-West (1987) robust standard errors are used to calculate the *t*-values in parentheses. The signs *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

commitment and quality of market making by bond dealers (see Bessembinder, Jacobsen, Maxwell, & Venkataraman, 2018). Due to more stringent regulations that raised capital requirements for banks, the cost of running a bond trading business increased significantly.[21] The postcrisis regulation reforms lower the incentive for dealers to commit capital to market making, and as a consequence, reduce the ability of dealers to provide liquidity to absorb order imbalances and price volatility, adversely affecting market quality.

Higher volatility and lower liquidity due to a decrease in dealers' capital commitment can impede arbitrage and widen pricing discrepancies. This may contribute to the negative basis in the postcrisis era. To investigate this possibility, we examine the role of dealers' capital commitment in the CDS-bond pricing relationship. In light of literature, we use dealers' inventories (net debt positions) as a measure of their capital commitment (see Bessembinder et al., 2018). Higher capital commitment enables dealers to absorb order imbalances into their inventories, leading to a positive relation between their capital commitment and inventories. The data on inventories (net debt positions) held by primary bond dealers are available from the FRB of New York (see Figure 3a).

To understand the role of dealers' liquidity provision relative to funding constraints in postcrisis CDS basis behavior, we run the regression of aggregate CDS basis against dealers' inventory (net positions) and hedge fund flow (net), with controls for the effects of bond return volatility, illiquidity, and other variables.[22] If the lack of dealers' capital commitment is primarily responsible for the negative basis in the postcrisis period, we should observe a greater effect of dealers' inventory than that of hedge fund flow on the CDS basis, as the funding constraint is apparently not binding in this period.

Table 8 reports the results of regressions for the three subperiods. For brevity, we only report the results based on the weekly CDS basis calculated by using swap rates, but our results are robust to the use of Treasury zero rates to

[21] See the TABB Group report in February 2016. This may also explain why the number of bond market makers has declined in recent years.

[22] Dealers' net positions are divided by trading volume to account for rises in corporate bond issuance and volume.

construct the CDS basis and to the use of daily data. All explanatory variables in the regression are standardized to facilitate a consistent comparison for the size of coefficients.[23] In Model (1), we include bond dealers' net debt positions (inventories) and net hedge fund flow, with controls for bond volatility and liquidity. An interesting finding is that dealers' inventory is highly significant in the post-crisis period, whereas hedge fund flow is not. Results show that dealers' inventory or capital commitment is a more important determinant of the negative CDS basis than the funding constraint variable in the postcrisis period. This finding suggests that the decrease in dealers' capital commitment in the postcrisis period contributes to the negative CDS basis. Higher bond return volatility in this period also has a significant negative effect on the CDS basis. In contrast, for the crisis period, the effect of hedge fund flow (net) is highly significant and has a much greater effect on the CDS basis than dealers' inventory does, suggesting that the binding funding constraints are mainly responsible for pricing discrepancies in the crisis period. For the precrisis period, neither variable is significant.

When we add the variables of failures-to-deliver and dealers' financing in Model (2), the relative importance of dealers' inventory to hedge fund flow remains unchanged. Dealers' capital commitment (inventory) is more important, in terms of the size of coefficient and *t*-value, than the funding constraint variables (e.g., net hedge fund flow and dealers' financing) in the postcrisis period, confirming that funding constraint is not an issue after the subprime crisis.

Overall, there is evidence that the negative CDS basis in the postcrisis period is attributable to the lower capital commitment and poorer quality of market making of bond dealers. The results suggest that bank regulation reforms after the financial crisis have an unintended consequence for the CDS-bond pricing inefficiency in credit markets.

5 | CONCLUSIONS

How can two similar securities have large pricing discrepancies? This is an extremely important issue that bears on asset pricing and the efficiency of resource allocation in the economy. Our paper addresses this issue using the data of credit markets, which exhibited an unusually negative CDS basis during the subprime crisis. We examine the persistence of pricing discrepancies and explore the potential factors behind it. To substantiate our hypothesis, we conduct extensive tests to examine the role of impediments to arbitrage and limited capital provision in CDS-bond mispricing.

We find that CDS-bond pricing violations are quite common and closely related to variables of impediments-to-arbitrage and limited capital provision. Firms with high leverage and risk and low liquidity are more likely to experience disintegration in credit markets. In addition, violations of market integration are more severe when arbitrage impediments at the market level are higher, as in the crisis period. Moreover, firms with higher impediments to arbitrage have greater persistence in pricing discrepancies. Limited arbitrage in credit markets is important not only in the crisis period, but also in normal times.

The funding-related variables exhibit a persistence pattern resembling that of the CDS pricing discrepancies. Firms that are more sensitive to the shortage of arbitrage capital exhibit higher persistence in pricing discrepancies. Results suggest that persistence in the shortage of arbitrage capital contributes to persistence in CDS-bond pricing discrepancies.

Finally, we find that the negative CDS basis continues to persist in the postcrisis period. This phenomenon is attributable to dealers' lower capital commitment to market making, which reduces their ability and willingness to provide liquidity to bond markets, and hence, adversely affects market quality and dealers' intermediacy.

ACKNOWLEDGMENTS

We thank Bing Han (Editor) and an anonymous referee for their extremely helpful suggestions and comments. We thank Jack Bao, Efraim Benmelech, Sheng-Syan Chen, San-Lin Chung, Toby Daglish, Jefferson Duarte, Thierry Foucault, Marcin Kacperczyk, Dongmei Li, Kai Li, Gustavo Manso, and Jack Robles for their insightful critiques. We also

[23] Standardized variables have a mean equal to zero and a standard deviation equal to one.

thank the participants at the 2014 Conference of the European Financial Management Association, the International Conference on Corporate Finance and Financial Markets at City University of Hong Kong, seminars at Victoria University of Wellington and National Taiwan University, and colleagues at SUNY Buffalo for helpful comments. Junbo Wang acknowledges financial support from City University of Hong Kong (Project 7004979 and 9220087) and a research grant from the National Science Foundation of China (No. 71720107002).

REFERENCES

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

Bai, J., & Collin-Dufresne, P. (2018). The CDS-bond basis. *Financial Management*, forthcoming.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*, 3–30.

Bakshi, G., Cao, C., & Chen, Z. (2000). Do call prices and the underlying stock always move in the same direction? *Review of Financial Studies*, *13*, 549–584.

Bessembinder, H., Jacobsen, S., Maxwell, W., & Venkataraman, K. (2018). Capital commitment and illiquidity in corporate bonds. *Journal of Finance*, *73*, 1615–1661.

Bessembinder, H., Kahle, K., Maxwell, W., & Xu, D. (2009). Measuring abnormal bond performance. *Review of Financial Studies*, *22*, 4219–4258.

Bharath, S., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies*, *21*, 1339–1369.

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.

Brav, A., Heaton, J. B., & Li, S. (2010). The limits of the limits of arbitrage. *Review of Finance*, *14*, 157–187.

Brunnermeier, M., & Pedersen, L. H. (2009). Market liquidity and funding liquidity. *Review of Financial Studies*, *22*, 2201–2238.

Choi, J., & Shachar, O. (2014). Did Liquidity Providers Become Liquidity Seekers? Evidence from the CDS-bond Basis during the 2008 Financial Crisis. University of Illinois-Urbana Working Paper.

Daniel, N., Denis, D., & Naveen, L. (2017). How do Firms Use Their Financial Flexibility? University of Pittsburgh Working Paper.

Das, S., Kalimipalli, M., & Nayak, S. (2014). Did CDS trading improve the market for corporate bonds? *Journal of Financial Economics*, *111*, 495–525.

Dick-Nielsen, J., Feldhutter, 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. (2010). Presidential address: Asset price dynamics with slow-moving capital. *Journal of Finance*, *4*, 1237–1267.

Duffie, D., Garleanu, N., & Pedersen, L. H. (2007). Valuation in over-the-counter markets. *Review of Financial Studies*, *20*, 1865–1900.

Fontaine, J.-S., & Garcia, R. (2012). Bond liquidity premia. *Review of Financial Studies*, *25*, 1207–1254.

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

Friewald, N., Jankowitsch, R., & Subrahmanyam, M. (2012). Illiquidity or credit deterioration: A study of liquidity in the us corporate bond market during financial crises. *Journal of Financial Economics*, *105*, 18–36.

Galil, K., Shapir, O. M., Amiram, D., & Ben-Zion, U. (2014). The determinants of CDS spreads. *Journal of Banking and Finance*, *41*, 271–282.

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

Gorton, G., & Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics*, *104*, 425–451.

Gürkaynak, R. S., Sack, B., & Wright, J. H. (2007). The US Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, *54*, 2291–2304.

Hassler, U., & Kokoszka, P. (2010). Impulse responses of fractionally integrated processes with long memory. *Econometric Theory*, *26*, 1885–1861.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, *68*, 165–176.

Hu, G. X., Pan, J., & Wang, J. (2013). Noise as information for illiquidity. *Journal of Finance*, *68*, 2223–2272.

Hurst, H. E. (1951). Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, *116*, 770–799.

Johansen, S. (1995). *Likelihood Based Inference in Cointegrated Vector Error Correction Models*. Oxford: Oxford University Press.

Kapadia, N., & Pu, X. (2012). Limited arbitrage between equity and credit markets. *Journal of Financial Economics*, *105*, 542–564.

Li, H., Zhang, W., & Kim, G. H. (2014). The CDS-bond basis and the cross-section of corporate bond returns. National University of Singapore Working Paper.

Lin, H., Wang, J., & Wu, C. (2011). Liquidity risk and expected corporate bond returns. *Journal of Financial Economics*, *99*, 628–650.

Liu, S., & Wu, C. (2017). Repo counterparty risk and on-/off-the-run treasury spreads. *Review of Asset Pricing Studies*, *7*, 81–143.

Lo, A. W. (1991). Long term memory in stock market prices. *Econometrica*, *59*, 1279–1313.

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.

Man, K. S., & Tiao, G. C. (2006). Aggregation effect and forecasting temporal aggregates of long memory processes. *International Journal of Forecasting*, *22*, 267–281.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, *29*, 449–470.

Mitchell, M., Pedersen, L. H., & Pulvino, T. (2007). Slow moving capital. *American Economic Review*, *97*, 215–220.

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

Mitchell, M., Pulvino, T., & Stafford, E. (2002). Limited arbitrage in equity markets. *Journal of Finance*, *57*, 551–584.

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

Pastor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*, 642–685.

Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics*, *42*, 35–52.

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *Journal of Finance*, *52*, 35–55.

Tsai, H., & Chan, K. S. (2005). Temporal aggregation of stationary and nonstationary discrete-time processes. *Journal of Time Series Analysis*, *26*, 613–624.

Zivot, E., & Wang, J. (2006). *Modeling Financial Time Series with S-Plus*, 2nd Ed. Seattle, WA: Insightful Corporation.

How to cite this article: Lin H, Man K, Wang J, Wu C. Price discovery and persistent arbitrage violations in credit markets. *Financial Management*. 2020;49:207–233. https://doi.org/10.1111/fima.12261

APPENDIX A: COINTEGRATION ANALYSIS

This appendix describes the procedure of cointegration tests. The cointegration analysis is conducted at the firm level. Two markets are cointegrated if they exhibit a long-term equilibrium relation. For the cointegration analysis to be meaningful, we need to check if the two spread series are each nonstationary I(1). Thus, we first conduct an augmented Dickey-Fuller (ADF) test on each series, with the null hypothesis that each series has a unit root. If both series accept the null hypothesis of unit root, we move forward to the cointegration analysis; otherwise, the case will be left out. After confirming that a unit root exists in each spread series, we set up a VAR model and perform the Johansen (1995) cointegration test. We describe the test procedure below.

Let $Y_t = (p_t, q_t)'$, where p_t and q_t are CDS and bond yield spreads (PECDS), respectively. We consider a VECM model of the following form[24]:

$$\Delta Y_t = c_0 + \Pi Y_{t-1} + \sum_{i=1}^{m} \Psi_i \Delta Y_{t-i} + u_t. \tag{A1}$$

[24] The diagnosis shows that an AR order of 3 is adequate for the VECM model.

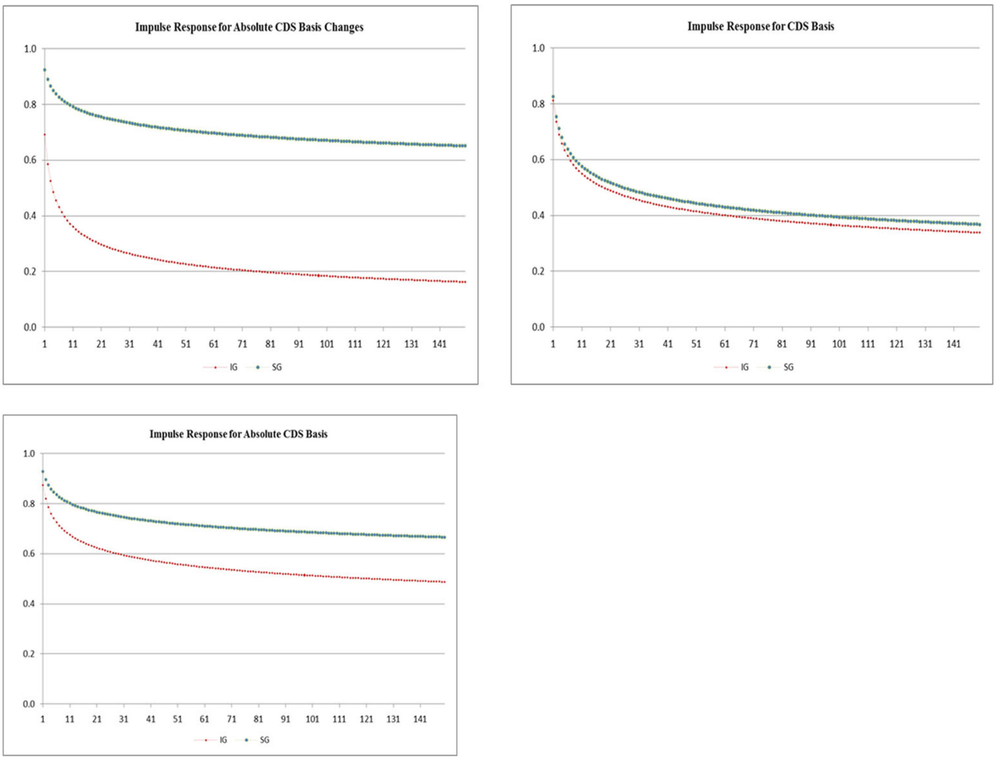

FIGURE A1 Impulse Response Functions. This figure plots the impulse response of absolute CDS basis changes, CDS basis, and absolute CDS basis for investment- and speculative-grade bonds

Based on the rank of the impact matrix Π, the number of cointegration vectors is determined. Specifically, there are three possibilities: 1) the case of no cointegration vector with the rank of $\Pi = 0$, in which both series are nonstationary I(1) but they are not cointegrated since a cointegration vector cannot be found; 2) the case of two cointegration vectors with the rank of $\Pi = 2$, in which both CDS and BYS are stationary series; and 3) the relevant case of one cointegration vector with the rank of $\Pi = 1$. In the last case, both series are nonstationary I(1) but their linear combination is stationary; therefore, $\Pi = \delta\omega'$,and $\omega = (1, \zeta)'$ is the normalized cointegration vector.

We use Johansen's (1995) trace test to draw the inference, which is a sequential test procedure. Step 1 tests the null hypothesis Ho: no cointegration vector against the alternative hypothesis Ha: at least 1 cointegration vector. If Ho is not rejected, then no cointegration vector is concluded. If Ho is rejected, step 2 further tests if there are one or two cointegration vectors; that is, Ho: one cointegration vector versus Ha: two cointegration vectors. Thus, to conclude one cointegration vector, the trace test statistic has to be significantly large in step 1 and is significantly small in step 2. The test procedure and critical values are discussed in the “urca” package in R.

APPENDIX B: ESTIMATION AND INTERPRETATION OF THE LONG MEMORY PARAMETER

A number of standard methods can be used to estimate the long memory parameter *d*, for example, the rescaled range R/S method, periodogram methods, or a parametric model of ARFIMA (*p*, *d*, *q*) with maximum likelihood estimation, to name just a few (see Hurst, 1951; Lo, 1991; Zivot & Wang, 2006; and the references therein). The choice of a suitable estimation method depends on the time-series property of the variable investigated. During the financial crisis period, the CDS basis exhibits a pronounced V shape (see Figures 1 and 2) and a nearly nonstationary behavior. We need to take care of this time-series property in the basis series; otherwise, estimation of the persistence parameter will be affected by time-series structures. For example, the persistence parameter (d) estimate of the rescaled R/S method could simply reflect the nonstationary behavior when the time series is near a random walk.

To overcome this problem, we adopt a two-step prefilter approach to estimate the persistence parameter based on an ARFIMA (p, d, q) framework. In the first step, we filter out other time series-dynamics with an aim of better revealing the persistence structure of long memory. That is, we fit an ARMA (p, q) model to the basis series first. In the second step, we use the residuals (the filtered series) from the ARMA fitting to estimate the d value of the ARFIMA. Using this two-step procedure, we find that the ACF of the filtered series exhibits a well-behaved slow decaying pattern and d estimates are quite stable. As an example, the ARFIMA ($\underline{1}$, d, 1) is in the form $(1-\phi_1 L)(1-L)^d\, y_t = (1{-}_1 L)e_t$, where d is the long memory parameter and the ARMA(1,1) component captures any remaining short-term dynamics of the basis series. The process is stationary when d is in the range of –0.5 to 0.5. Its ACF is slowly decaying at a hyperbolic rate of k^{2d-1}, where k is the lag order. We estimate the parameters of the model by maximum likelihood, using the "fracdiff" package in R.[25] For comparison, we also estimate the persistence parameter using the rescaled-range R/S method and find that our two-step procedure produces much more stable d estimates.

The best way to understand the magnitude of d is to translate the d value into the number of days that the basis will persist by the impulse response function (IRF). To get a sense of the practical meaning of the value of d, we compute the impulse response (IR) for the estimates of persistence parameters in Panel A of Table 6, using the ir.arfima function in the "afmtools" package of R.[26] We compute the impulse response up to 150 lags (see figure A1 below). For example, the d values for |Δbasis| during the crisis are 0.35 and 0.46 for investment grade (IG) and speculative grade (SG) bonds, respectively. For IG (lower curve), the impulse response R_1 is 0.69 and $R_{150} = 0.16$, and it takes 13 days (half-life) to decrease to $R_{13} = 0.35$. For SG (upper curve), R_1 is 0.92 and $R_{150} = 0.65$, and the half-life is longer than 150 days. Clearly, the basis changes of SG are much more persistent. For the CDS basis, the d values during the crisis are 0.40 and 0.41 for IG and SG, respectively. For IG (lower curve), the impulse response R_1 is 0.81 and $R_{150} = 0.34$, and it takes 57 days (half-life) to decrease to $R_{57} = 0.40$. For SG (upper curve), R_1 is 0.83 and $R_{150} = 0.37$, and the half-life is 78 days at $R_{78} = 0.41$. As indicated, even though the difference in d values seems relatively small, it translates into a difference in the half-life of 21 days. Finally, for the absolute basis, |Basis|, the d values during the crisis are 0.44 and 0.46 for IG and SG, respectively. For IG (lower curve), the impulse response R_1 is 0.87 and $R_{150} = 0.49$. For SG (upper curve), R_1 is 0.93 and $R_{150} = 0.67$. Both half-lives are longer than 150 days.

[25] Direct estimation of the ARFIMA(p, d, q) model that accounts for both d and time series dynamics does not work out well. We find the likelihood function does not have a clear peak as the model itself cannot distinguish whether the nonstationary behavior comes from the basis persistence or from the unit root.

[26] As an example, a stationary ARFIMA (1, d, 1) model $(1-\phi L)(1-L)^d x_t = (1+\theta L)\, a_t$ has an equivalent MA(∞) expression with the MA coefficient $\psi_j \sim \frac{1-\theta L}{1-\phi L} j^{d-1}/\Gamma(d)$ for large j (see Hassler & Kokoszka, 2010). The IR is computed as $R_j = \sum_{i=0}^{j} \psi_i \eta_{j-i}$, where $\eta_t = \Gamma(t+d)/[\Gamma(t+1)\Gamma(d)]$.