The Empirical Relationship between CDS Prices and Credit Spreads in Long Time Series:
An Application to Large U.S. Banks

Jordan Trudel

Department of Finance
Hanken School of Economics
Helsinki
2017
Title of Thesis:

The Empirical relationship between CDS prices and Credit Spreads in Long Time Series

Abstract:

The article tests for cointegration between the CDS price and credit spread in long time series for six large American banks. The data span from 22 February 2010 to 31 December 2016. The results show that there is evidence of cointegration for most of the series when using the Johansen Trace test, but only for one when the Bayesian information criterion is used. Further results show that the series of two banks price credit risk equally up to a constant, but only one series does not reject exact parity. Tests on the dynamics between the two markets show that the CDS market contributes to most of price discovery. Further tests on the residuals show that the model is correctly specified and support the evidence for cointegration, which suggest that the BIC is too restrictive in this case.

Keywords: CDS, Credit Spread, Cointegration, Long Time Series, Price Discovery
# CONTENTS

1 INTRODUCTION ........................................................................................................... 1
   1.1. Pricing credit risk ............................................................................................... 2
   1.2. Overview of the credit markets ........................................................................... 3
   1.3. Arbitrage relationship and basis ......................................................................... 4
   1.4. Determinants of short-run relationship .............................................................. 7
   1.5. Unit root and integration ..................................................................................... 9
   1.6. Cointegration ...................................................................................................... 12

2 ARTICLE ......................................................................................................................... 14
   2.1. Introduction ......................................................................................................... 14
   2.2. CDS price and credit spread .............................................................................. 16
   2.3. Empirical Model .................................................................................................. 18
       2.3.1. Cointegration .............................................................................................. 18
       2.3.2. Vector error correction model and price discovery .................................... 20
   2.4. Data ..................................................................................................................... 21
   2.5. Results ................................................................................................................. 25
       2.5.1. Price discovery ............................................................................................ 30
       2.5.2. Tests ........................................................................................................... 32
   2.6. Conclusion .......................................................................................................... 37

REFERENCES ................................................................................................................. 39
1 INTRODUCTION

There are numerous ways for companies to access capital such as bank loans, emitting new shares on the stock market and emitting debt on the fixed income market. The thesis focuses on the latter one. Companies are subject to different degrees of riskiness and face external events that can have a negative impact on their ability to pay their debt. As such, investors require a compensation for facing this risk. The rate required by investors to face the bond’s risk is known as the credit spread and represents a measure of the excess required return over the risk-free rate.

In the beginning of the 1990’s, new types of instruments were created as derivatives of credit risk. Among these credit derivatives, the most common is the credit default swap (CDS), which acts as an insurance against the probability of default of a company. It is directly linked to the fixed income instruments of the company and makes them virtually risk-free. In other words, it represents the same unit of credit risk as the credit spread priced in a different market. Therefore, the two measures should be tightly linked. If the two markets do not price the risk equally, arbitrage opportunities will arise.

The discrepancies between the two credit prices give rise to something called the basis in the approximate non-arbitrage relationship derived by Duffie (1999). It is simply the measure, in basis points, of the difference between the credit default swap spread and the credit spread:

\[ \text{basis}_t = P_{CDS,t} - P_{CS,t}, \]

where \( P_{CDS,t} \) is the CDS price at time \( t \) and \( P_{CS,t} \) is the credit spread at time \( t \). If the markets were perfect, the basis would be a stationary series centered on 0. However, because of different imperfections and factors affecting the two markets, this relation might not hold in practice.

Although the price of the credit default swap and credit spread can differ substantially in the short-run, there generally exists a long-term equilibrium between the two because they represent the same unit of risk priced in two different markets, and a large difference would lead to arbitrage opportunities. This long-run relationship can be empirically tested through cointegration. Simply put, cointegration is when there exists a linear combination between two time series integrated of order 1 so that this combination is integrated of order 0. In other words, the relationship is stationary even though the two
separate series have a unit root. Since the two spreads should be equal as theory posits, the cointegration vector representing them should be \([1, -1, 0]\). However, since credit risk prices can differ due to imperfections in the markets and that these differences can persist through long periods of time, this perfect relationship might be too restrictive. There can exist, for example, a constant basis which would make the cointegration vector \([1, -1, c]\). The thesis aims to test this relationship between the two units of risk. The thesis is organised in an article format and is presented as follows. The present chapter outlines the background information about CDS price, credit spread and cointegration. The second chapter consists of the article and is much more to the point and direct than the usual thesis.

### 1.1. Pricing credit risk

Although this thesis does not concentrate on the pricing of credit risk per say, but on the relationship between two series of credit spreads and CDS prices, a short introduction to the pricing models help the comprehension on how credit events are taken into account in the price of the credit instruments and the dynamics between the two markets. Two particular types of models are most commonly used to price credit risk: structural models and reduced-form models. The structural models (Collin-Dufresne et al. (2001)) are largely based on Merton (1974) and the option pricing theory, and consider the risk default as an endogenous process. In these models, it is assumed that there is complete information about the firm’s value and that the bond will default when the firm’s value goes under the debt value. This can be perceived as a put option on the assets of the firm with a strike price equal to the debt value. Credit risk is simply the probability that the underlying asset falls under the strike price at maturity or, in other words, that the value of the firm’s assets is lower than the value of its debt. The default probability is thus a function of the firm’s financial structure: its leverage, assets return volatility, time to maturity, loss given default and the risk-free rate. Alternatively, reduced-form models (e.g. Duffie and Singleton (1999)) assume that default is exogenous and can be represented by an intensity process. Default happens at a random stopping time and the arrival intensity is stochastic. In these models, default is a probabilistic event and credit spread is the risk-neutral valuation when there is no possibility of arbitrage.

Although complete information about the firm value might be an unrealistic assumption, the structural models have been preferred by practitioners because of the weak economic rationales of the stochastic arrival intensity of default in the reduced-form models.
1.2. Overview of the credit markets

The first credit derivative instruments were created in the early 1990s, with the creation of the modern credit default swap contract being attributed to J.P. Morgan in 1994 (Lanchester (2009)). An organised market started in 1996 with many banks looking for new ways to manage their credit exposure in a similar manner as those used for hedging and diversifying interest and currency risks. The market experienced a sharp growth thereafter with many other financial institutions starting to trade. Banks, although they are still the biggest participant in the market on both the seller and buyer side, have seen their role decline. On the other hand, hedge funds, who only had a marginal participation in the market in the beginning, have now increased their buyer and seller role to the second biggest participant in the credit derivatives market. Other participants include insurance companies, pension funds, mutual funds and corporations. The role of individual participants is marginal in credit default swaps trading. According to the British Bankers Association (BBA), credit derivatives have grown from $US 180 billion in 1997 to over $US 20 trillion in 2006. The International Swaps and Derivatives Association (ISDA) has an even bigger estimate for 2006 at over $US 34 trillion. In 2007, the market was at his highest point with a notional value of over $US 62 173 billion. However, since the financial crisis, the market has shrunk considerably and lost 38% of its value from 2007 to 2008 alone. The downfall of the market has continued since then. The Bank for International Settlement (BIS) estimates the market value for the CDS market during the first half of 2016 to be $US 11 777 billion.

Credit default swap derivatives remain however a very important market to transfer risk in an easy manner because of the liquidity of these instruments compared to the bond market. The lack of liquidity of the bond market comes from the fact that most of the investors hold their securities until maturity, which makes the number of transactions very low compared to other markets such as the stock market or the money market (Alexander et al. (1998)). Shorting securities on the bond market is also difficult and the repurchase agreements on risky bonds are usually for short terms. These drawbacks on the flexibility of the bond market make the derivative market attractive as investors can short credit risk for long periods of time at low costs and enjoy greater liquidity. By comparison to the above figures, the total debt securities market around the world was $US 21 701 billion in the first semester of 2016 (BIS). With a value of more than half the debt securities market, the credit default swap market is of uttermost importance in the management of risk for many investors. The participants in the bond market are quite alike the ones for the CDS market with banks, pension funds and mutual funds as the
most active. Other participants include governments, financial institutions and individuals.

In order to centralise the credit derivatives market and rule out some disagreements and legal concerns, the ISDA issued in 1999 the first version of its set of definitions for standard credit derivatives. This was later amended and updated in 2003 and in 2014, which is now the current version. The definitions describe the basic framework used in the credit derivative transactions and the events which constitute a default event. Default events include bankruptcy, failure to pay, obligation default or acceleration, and repudiation or moratorium (for sovereign entities). Bond restructuration was amended from the early definitions although it still constitutes a credit event under certain conditions. These conditions include a reduction or delay in the interest rate or principal paid at maturity, a decrease in the bond payment priority, or a change in currency or composition in any payments. The controversy concerning restructuration originally occurred since the deliverable assets do not necessarily come due and payable in case of such event. In the case where not all debt is due and payable when restructuration occurs and when there exist bonds traded at discount, the price of the CDS includes a cheapest-to-deliver (CTD) option and it does not represent a pure measure of risk anymore. In 2014, ISDA also included government initiated bail-in as a credit event for European corporate CDS. Moreover, they produce the most commonly used standardised CDS contract, the Master Agreement.

1.3. Arbitrage relationship and basis

The price of credit default swap and credit spread represent the same unit of risk, i.e. the risk of default of a risky fixed income asset, priced in two different markets. As such, one would expect the two units of risk to be priced the same because of the price parity principle. If they are not, arbitrage opportunities arise.

To illustrate this, take a simple example. An investor that buys a par-valued bond with a maturity of $t$-years and a yield-to-maturity of $y$ eliminates most of his risk of default when he enters in a $t$-year CDS contract, but decreases his return by the price of the CDS, $p$. His final return on his now near default-free bond, $y−p$, should be the very close to the $t$-year risk-free rate $r$, and thus his payment $p$ should be the same as $y−r$, i.e. the credit spread. If $y−p$ is lower than $r$, the investor can short the bond, sell a CDS contract and take a long position in the risk-free asset for a risk-free arbitrage profit. In the same
manner, if \( y - p \) is higher than \( r \), the investor can buy the bond, enter in a buyer position on a CDS contract and short the risk-free asset for an arbitrage profit.

However, this arbitrage is valid only in a perfect situation which can rarely be found in the market. In order for the arbitrage to be perfect, the credit spread has to be calculated from a risky and risk-free floating-rate note (Duffie (1999)), but these instruments are rare. When using fixed-coupons bonds, the payment dates of the reference instrument and the CDS have to match exactly and the recovery on default has to be a constant fraction of the face value (Houweling and Vorst (2002)). Other requirements to make the arbitrage relationship perfect are that the risk-free curve is flat, the interest rate is constant and the CDS payment in case of default is the principal plus accrued interests on the reference risky par yield bond times one minus the recovery rate in case of default (Hull and White (2000a)). These conditions only happen in rare instances in the market, but Hull and White (2000a) also show that the arbitrage relationship holds quite well in periods of low interest rate and flat yield curves for assets trading close to par. Using bonds that trade close to par is important when testing the arbitrage relationship between credit spreads and CDS prices because of the payment structure of the swap in case of default. Since the protection seller pays the difference between the par value of the bond and its post default value, an investor that bought the bond below par value will only lose the difference between the discounted price of the bond and its post default value and will thus make a supplementary gain over his investment of par minus discounted price. This will tend to push the CDS spread higher than the credit spread causing departure from arbitrage (Mengle (2007)).

As mentioned above, the credit default swap market is more liquid than the bond market. Since there are no cash constraints in the CDS market, i.e. there is no initial cash exchange when entering in the CDS contract, this market answers more quickly to a change in the underlying credit risk than the bond market where the cash constraints can be substantial. The difference in response time between the markets leads to price discrepancies in the short-run (Zhu (2006)). Another factor that can cause the arbitrage relationship not to hold is when there is a physical settlement upon default. The protection buyer can then choose within a pool of pre-specified assets leading to a cheapest-to-deliver option. This increases the CDS spread compared to the credit spread and it is impossible to value this option analytically since there is no benchmark on the post default value of deliverable bonds (Blanco et al. (2005)). Also, since the short sale of bonds is difficult in practice, it creates an asymmetry in the ability to capitalise on
arbitrage opportunity when the price of the CDS is higher than the bond spread. The CDS price will thus tend to remain higher when there are significant repo costs in shorting the bond (Duffie (1999)). The true credit spread in this situation has to include the repo costs and is higher than the estimated measure (bond yield minus risk-free rate). Both the option and the repo costs are difficult to quantify as there rarely exists reliable repo costs data and there is no valuation model for the CTD option, but their values are both bounded by zero which makes the CDS spread an upper limit of the price of the credit risk and the credit spread a lower limit (Blanco et al. (2005)).

Transaction costs are another factor to take into account when analysing departures from equilibrium. If these costs are substantial, it decreases the profit from small arbitrage opportunity allowing them to persist through time. Furthermore, Collin-Dufresne et al. (2001) show that movements in the liquidity premium can explain a large portion of the movements in the total variation of the credit spreads. The lack of liquidity in the bond market can push the investors towards the CDS market, creating disparities in the demand and supply causing price movements that are unrelated to default expectations in the short-run (Blanco et al. (2005)). Another consideration to take into account for the arbitrage relationship to hold is the maturity of the bonds used to calculate the credit spread. CDS prices are quoted on a constant maturity of 5 years, thus using a bond with a decreasing time to maturity will disrupt the parity relationship of the two units of credit risk. In order to maintain comparable risk structure between the two units of risk, credit spreads have to be modified to be expressed in the same manner.

The short-run discrepancies between the two credit prices are called the basis. It is simply the measure, in basis points, of the difference between the credit default swap spread and the credit spread:

\[
basis_t = P_{CDS,t} - P_{CS,t} = P_{CDS,t} - (y_t - r_t),
\]

where \(P_{CDS,t}\) is the CDS price at time \(t\) and \(P_{CS,t}\) is the credit spread at time \(t\). Furthermore, \(y_t\) is the yield-to-maturity at time \(t\) and \(r_t\) is the risk-free rate at time \(t\), so that \(P_{CS,t} = y_t - r_t\). If the markets were perfect, the basis would be 0. However, as noted above, this relation might not hold in practice because of the different factors affecting the two markets. Blanco et al. (2005) find that the mean average basis for their 33 reference entities is 6 basis points when using swap rates for risk-free proxies and the mean average absolute basis is 15 basis points. To put this number into perspective, they also find that the average bid-ask spread is 12 basis points for the CDS and 9.5 basis points for the
bonds. When using government bonds as risk-free proxies, they find that the mean average basis decreases to –41 basis points and that the mean average absolute basis increases to 46 basis points. In a similar manner, Zhu (2006) finds that the mean average basis is 15 basis points and that the mean average absolute basis is 29 basis points when the credit spread is calculated with swap rates. When Treasuries are used as a proxy, these numbers change to –52 and 64 basis points, respectively. The large differences in the basis using the different proxies prove that the government bonds are no longer valid proxies for the risk-free rate. As noted briefly above, this is due to the taxation treatment, repo specials, scarcity premium and their benchmark status. Elton et al. (2001) point out that Treasuries are exempt of state taxes in the United-States and that adjusting the rate using

\[ r_f = \frac{r_{\text{treasury}}}{1-(1-t_g)t_s}, \]

where \( t_g \) and \( t_s \) are the federal and state tax rates respectively, greatly reduces the basis although the price differentials are still higher than when using swap rates as proxies. Also, Reinhart and Sack (2002) show that the Treasury yields and the risk-free rates have become increasingly dissociated due to an increase in the volatility of idiosyncratic Treasury shocks. Moreover, Liu et al. (2006) demonstrate that while the swap and CDS spread are mainly affected by changes in the expected default risk, the government bonds rates are mostly driven by liquidity factors which causes departure from equilibrium when these rates are used to proxy the risk-free rates. Although the arbitrage relation in these studies seems to hold well on average, there can be large price differentials for short periods of time and these differences vary from entity to entity.

1.4. **Determinants of short-run relationship**

In order to understand the short-term dynamics of credit spreads and CDS prices, and how the basis varies over time, it is important to look into the variables that can affect the probability of default and the changes in the expected recovery value of the bond. Collin-Dufresne et al. (2001) propose several different variables for their structural model to explain the changes in the credit spread. First, they consider the changes in the spot interest rate since higher spot rates have a static effect that increases the risk-neutral drift of the firm valuation process and which in turn reduces the risk-neutral probability of default and thus decreases the credit spread (Longstaff and Schwartz (1995)). The second variable that they use is the change in the slope of the yield curve. The slope of the term structure has a role in determining the future spot rate and as such can influence the probability of default in the long term. For example, an upward term structure signals that the future spot rates will be on the rise as the mean short-term rates revert around
the long-term rates. An increase in the short-term rate signals an improvement in the general condition of the economy which would lead to a decrease in credit spreads. Third, they consider the change in leverage of the reference entity since default happens when the assets have a lower value than the liabilities. An increase in the leverage thus brings more risk to the company and would increase the credit spread. The fourth variable is the change in market volatility. According to the contingent-claim theory, a debt claim has similar features to selling a put option. Relating these two concepts, the credit spread should increase with volatility since the value of the option also increases. In other words, the higher the volatility, the higher the risk of default and the higher the credit spread. The fifth variable is the change in the probability or the magnitude of a downward jump in the firm value and relates to the implied volatility in option pricing theory. A higher probability of a downward jump in the firm value increases the probability of default and decreases the expected recovery. Thus, it increases the credit spread. A similar situation occurs when the magnitude of the jump is larger. Last, they consider the change of business climate since this variable will affect the expected recovery rate even if it does not necessarily decrease the probability of default. Their results show that the theoretical variables that should determine the changes in the credit spread, although correctly signed, have low explanatory power and that the changes are driven by an unidentified common factor.

Blanco et al. (2005) extend the model of Collin-Dufresne et al. (2001) and also research how the CDS spread and credit spread are differently affected by the factors. They use mainly the same variables although they use the change in equity prices instead of the change in leverage as a proxy to measure the firm financial health over a short horizon. They also include a liquidity factor since a constriction in the liquidity of the bond increases the risk for the holder which will affect both credit measures, and a lagged measure of the basis to consider the adjustment of the spreads to a disturbance in the equilibrium. They find that in the short-run firm-specific factors have a larger effect on the CDS prices than on the credit spread while the macroeconomic variables affect the credit spread the most. The lagged basis is highly significant and correctly signed for both markets, and it is also greater in absolute magnitude for the credit spread. This adjustment of the credit spread market to the CDS market through the lagged basis also implies that both markets will react equally to firm-specific factors in the long-run because of the arbitrage-based equivalences; the CDS reacts to changes in firm-specific variables and the credit spread reacts to lagged changes in CDS prices, thus, in the long-run, credit spreads are equally affected by the firm-specific factors. Although the
presence of a lagged basis increases the explanatory power of their models (though only marginally for the CDS regressions), they also find that the residuals are highly cross-correlated which suggest, as in Collin-Dufresne et al. (2001), the presence of an unidentified common factor.

Another way to look at the short-term dynamics between the two markets is with price discovery. The method, which will be explained in more details in the article, measures which market adapts the most to a change in the other market. Previous research generally finds that the CDS market leads in term of price discovery. Blanco et al. (2005) find that the estimated adjustment coefficient for the credit spread market, $\hat{\alpha}_2$, is positive and statistically significant for 25 out of 27 entities and that the estimated adjustment coefficient for the CDS market, $\hat{\alpha}_1$, is significantly negative in only 8 cases. They find that on average, CDS contributes to around 80% of price discovery according to the Hasbrouck and Gonzalo-Granger measures. Moreover, they find that CDS prices Granger-cause 4 out of the 6 entities where there was no evidence for cointegration. Zhu (2006) finds similar results. In 17 out of 23 cases, CDS prices lead credit spreads in price discovery. He also finds that CDS prices contribute to around 63% of price discovery on average using the Gonzalo-Granger measure. He argues that this can explain the increase in basis in 2002 when many companies saw their credit rate deteriorate; since CDS spread moves faster than the bond spread, it will cause a pricing error between the two series and increase the basis until the bond spread converges to the CDS spread. Norden and Weber (2009) also find that CDS tend to lead the bond market in price discovery, but in addition find that the stock market leads the other two.

1.5. Unit root and integration

A common trait of many financial time series, which is also at the center of the cointegration concept, is the unit root. A unit root is when the characteristic equation of a time series process has a root on the unit circle. The characteristic equation comes from the linear and homogenous difference equation that relates the present value of a variable to its past values. It is the $p^{th}$ order polynomial that equates the coefficient of the equation to 0. Take the autoregressive (AR) process of order $p$

$$X_t = \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + Z_t,$$

where $Z_t$ has a stationary distribution, e.g. $Z_t \sim IID(0, \sigma^2)$. The process can be rewritten as
\[ X_t - \varphi_1 X_{t-1} - \cdots - \varphi_p X_{t-p} = Z_t. \]

Using the lag operator, the AR(p) process can be written as

\[ (1 - \varphi_1 L - \cdots - \varphi_p L^p)X_t = \varphi(L)X_t = Z_t, \]

where \( \varphi(L) = (1 - \varphi_1 L - \cdots - \varphi_p L^p) \). For arbitrary \( z \),

\[ \varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p. \]

The stationary condition of the previous model is that all the roots of the characteristic polynomial

\[ 1 - \varphi_1 z - \cdots - \varphi_p z^p = 0 \]

lie outside the unit circle, i.e. are greater than 1 in absolute value. As a simple example, let \( \{X_t\} \) be a time series following the AR(1) model

\[ X_t = \varphi_1 X_{t-1} + Z_t, \]

where \( Z_t \) has a stationary distribution, e.g. \( Z_t \sim IID(0, \sigma^2) \). The characteristic polynomial for this autoregressive model is

\[ 1 - \varphi_1 z = 0. \]

The root of the AR(1) is \( z = \frac{1}{\varphi_1} \) and the model is stationary when \( |\varphi_1| < 1 \). The process has a unit root when \( \varphi_1 \) is equal to 1 and becomes a random walk, which is non-stationary since the mean and variance depend on time and will tend to infinity as time goes to infinity. It is easily seen by iterating the AR(1)

\[ X_t = \varphi(\varphi X_{t-2} + Z_{t-1}) + Z_t \]

\[ X_t = \varphi(\varphi(\varphi X_{t-3} + Z_{t-2}) + Z_{t-1}) + Z_t. \]

As the iteration goes to infinity and assuming that \( X_0 = 0, X_t \) can be expressed as

\[ X_t = \sum_{j=0}^{\infty} \varphi^j Z_{t-j}, \]
which does not converge if $|\varphi_1| = 1$. Therefore, no interference can be made from the random walk model as it would lead to a spurious regression.

The above autoregressive model of order 1 is said to be difference stationary and integrated of order 1, I(1), which means that it has to be differentiated one time in order to be stationary

$$\Delta X_t = X_t - \varphi_1 X_{t-1} = Z_t.$$ 

The integrated process can be extended to higher orders, I(d), d=1, 2, 3... where $d$ represents the number of time that a series $\{X_t\}$ has to be differentiated in order to become stationary and I(0).

When two series are integrated and unrelated, a regression of one series on the other will lead to spurious regression. For example, take the regression of two random walk $\{Y_t\}$~I(1) and $\{X_t\}$~I(1), where

$$Y_t = Y_{t-1} + Z_{1,t},$$

$$X_t = X_{t-1} + Z_{2,t},$$

with $Z_{1,t}$~$IID(0, \sigma_1^2)$ independent of $Z_{2,t}$~$IID(0, \sigma_2^2)$ using a linear model

$$Y_t = \beta X_t + \epsilon_t,$$

where $\epsilon_t$~$IID(0, \sigma_2^2)$. The least square estimator for beta is

$$\hat{\beta} = \frac{\sum_{t=1}^{n} Y_t X_t}{\sum_{t=1}^{n} X_t^2}.$$ 

Since the two series are not related in any way, the true value of beta is 0 and the estimated coefficient should be equal to 0 too. However, Phillips (1986) shows that this is not the case even as the number of observations tends to infinity. Instead, $\hat{\beta}$ tends to a random variable that can be expressed as a function of bivariate standard Brownian motions and it is not consistent. The $t$-statistic on the estimate diverges to infinity as the number of observations tends to infinity making the probability of rejecting the true null hypothesis ($\beta = 0$) equal to one. Therefore, any inference made on the estimate of beta would be false as the test would systematically give a false outcome even as the number of observation tends to infinity.
A test for the presence of a unit root for the AR(1) model is the Dickey-Fuller test (Dickey-Fuller (1979)) which tests the hypothesis $H_0: \varphi_1 = 1$ against $H_a: |\varphi_1| < 1$ with the test statistic

$$\tau = \frac{\hat{\varphi}_1 - 1}{(S^2 / \sum_{t=1}^{n} X_{t-1}^2)^{1/2}},$$

where

$$S^2 = \frac{1}{n} \sum_{t=1}^{n} (X_t - \hat{\varphi}_1 X_{t-1})^2.$$

The statistic, while looking similar to the $t$-test, is non-standard and its asymptotic distribution was derived by Dickey-Fuller (1979) as $n$ tends to infinity. The unit root test can also be generalised to include autoregressive models of higher order and is known as the Augmented Dickey-Fuller test (Dickey-Fuller (1981)). The difference between the two models is that the augmented version test for the roots of the characteristic polynomial rather than one parameter alone. The test statistic has the same asymptotic distribution as the original Dickey-Fuller test. For both statistics, the critical values also depend on the presence of a constant and a trend.

### 1.6. Cointegration

Cointegration is when there exists a linear combination between two series integrated of order 1 such as this combination is integrated of order 0. In other words, there is no unit root in the relationship even though the two series contain a unit root. Suppose that the random walk $\{X_t\}$-I(1) in the previous subsection is used to generate the linear process $Y_t = \beta X_t + \epsilon_t$ where $\epsilon_t \sim IID(0, \sigma^2_\epsilon)$. Then $\{Y_t\}$ is a linear combination of $\{X_t\}$ and its first difference is

$$\Delta Y_t = Y_t - Y_{t-1} = \beta (X_t - X_{t-1}) + \epsilon_t - \epsilon_{t-1},$$

which can be rewritten as

$$Y_t = Y_{t-1} + \beta Z_t + \epsilon_t - \epsilon_{t-1} = Y_{t-1} + W_t,$$

where $W_t = \beta Z_t + \epsilon_t - \epsilon_{t-1}$ and is a linear combination of independent and identically distributed error terms that are independent of each other. The process $\{Y_t\}$ is also integrated of order one and contains a unit root, but the case is very different from the
spurious regression. The two series are said to be cointegrated with cointegration regression $Y_t = \beta X_t + \epsilon_t$ and cointegration vector $(1, -\beta)'$ since $Y_t - \beta X_t = \epsilon_t$ is I(0).

Most previous studies find that CDS prices and credit spreads are cointegrated and that there is a long-term relationship between the two series. Blanco et al. (2005) find that cointegration holds for all 16 American reference entities and for 10 out of 17 European entities. They suppose that the results for the European entities are due to the presence of a CTD option within the CDS prices, the absence of repo cost data, and the wide bid-ask spread that made the two series move in seemingly unrelated way. Furthermore, they find that when they impose restriction on the cointegration vector, the two markets price risk equally up to a constant at the 1% significance level for all the American reference entities and for all the European reference entities where cointegration was found. However, when they restrict the constant to be 0, the number of American entities diminishes to 11 out of 16 and the number of European entities to 5 out of 10. Similarly, Zhu (2006) finds evidence for cointegration for all 23 reference entities\(^1\) and does not reject the $[1, -1, 0]$ restriction on the cointegration vector for 14 entities. Additionally, De Wit (2006) finds evidence of cointegration in 87 cases out of 144 reference entities while Norden and Weber (2009) find evidence of cointegration for 36 out of 58 reference entities.

\(^1\) Zhu finds that bond spread series of Bank of America is integrated of order 0 and thus does not use the series for this entity in the cointegration test.
2 ARTICLE

2.1. Introduction

Many researches have been done on the long-term equilibrium between different instruments of the financial market. Two of these instruments are credit default swaps (CDS) and bonds. A credit default swap is an instrument which protects the buyer against the default of the underlying bond. In other words, it makes the bond virtually risk-free, and consequently, the price of the CDS should be equal to the credit risk of the bond, i.e. its credit spread. If not, arbitrage opportunities would arise. However, investors are not always able to take advantage of these arbitrage opportunities because of various market imperfections, and CDS prices and credit spreads can differ substantially in the short-run. Nevertheless, there generally exists a long-term equilibrium between the two units of risk which can be empirically tested through cointegration.

Several researchers use short series to study this relationship. For example, Blanco et al. (2005) use data covering only 18 months. They argue that, since they look at an arbitrage relation, it should revert rapidly to equilibrium and that the average half-life of the deviations within their sample is only six days and as such, their data are sufficiently long for the analysis. This justification is based on Hakkio and Rush (1991) who find that the ratio of the length of the series to the half-life of deviations is more important in cointegration analysis than the length of the series itself. However, there might be persistent shocks that push the two series apart which, if not taken into account, could increase the half-life of the deviations and make the analysis over a short period of time less pertinent. Since cointegration is a long-run relationship, a longer data frame should be preferred.

Figure 1, retrieved from Ahlgren and Catani (2014), represents the cointegration relation between the CDS price and credit spread of Bank of America. The data are from 1 January 2009 to 31 January 2012 and are daily observations. As the figure shows, the two series can depart from equilibrium for long periods of time (over a year) which could bias the results and lead to no evidence of cointegration while the two series actually have a long-run equilibrium. This shows that research using long time series is very relevant to the analysis of the equilibrium between the two measures of risk. Furthermore, Ahlgren and Catani (2017b) show that the power of tests of cointegration is high after 1000 observations or around four years of data.
Figure 1  Cointegration relationship for Bank of America

The purpose of the article is to empirically analyse the theoretical relationship between the CDS price and credit spread of large U.S. banks. More specifically, the research will test for the presence of cointegration between the two units of credit risk using long time series. Furthermore, the article investigates the dynamics of the short-term relation between the CDS price and credit spread by testing for price discovery. This article does not aim at finding new determinants of the short-term relationship between the two series of credit risk, nor to find new ways of pricing it. The test for the short-term relationship is limited to price discovery and does not look deeper into the different short-term determinants of credit risk. It also does not investigate the pricing of the credit risk itself, but uses the data as input to test the relationship. It is based on previous literature and compares past findings with those using long time series.

The data include CDS price and credit spread series for six large American banks: Bank of America, Citigroup, Goldman Sachs, JP Morgan, Morgan Stanley, and Wells Fargo. The series span from the beginning of 2010 to the end of 2016. The results show that there is cointegration between CDS prices and credit spreads for most of the banks when using Johansen trace test (Johansen (1996)), but there is cointegration for only one of them when using the Bayesian information criterion (BIC) (Cavaliere et al. (2015)). Out of the series for which there is evidence of cointegration with the Johansen trace test, two accept the long-run parity relationship, but only one prices credit risk equally in the two markets. The results also show that the CDS market contributes the most to price discovery and that the bond market adjusts to disruptions in the long-run equilibrium from the past period for most of the banks under observation. Tests on the residuals show that the cointegrated VAR model is correctly specified and support the results for
cointegration with the Johansen Trace test. This shows that the BIC is too conservative and will falsely accept rank 0 even if there is evidence for cointegration.

The remainder of the paper is organised as follows. Section 2.2 describes CDS price and credit spread. Section 2.3 presents the empirical model. Section 2.4 discusses the data used for the analysis. Section 2.5 presents the results. Section 2.6 concludes.

2.2. CDS price and credit spread

A credit derivative is an agreement between two parties that shifts the credit risk from one party to the other. The value of the derivative is directly linked to the value of another asset, the underlying. In the case of a credit default swap, it is an over-the-counter (OTC) contract where an investor takes a long position in the risk associated with the probability of default on the bonds of a company or government. This investor, the protection seller, agrees to a settlement with another investor, the protection buyer, in case the bond defaults. The settlements can either be physical or cash. In a physical settlement, the protection seller pays the par value of the bond to the buyer in exchange for the physical delivery of the reference asset. In the cash settlement, there is only one cash flow where the protection seller pays the notional amount minus the post default value of the reference asset. The post default value of the bond is most often determined by an auction on the bond after default.

In exchange for this protection, the protection buyer pays a periodic fee to the seller. This fee is usually quarterly and the annualised fee, expressed as a percentage, is referred to as the CDS price or CDS spread. The standard periodic payment dates are March 20, June 20, September 20 and December 20 and also serve as maturity dates. When a CDS is issued prior to those dates, it is subject to a stub period where the fees are accrued until the next standard date and continues according to the usual schedule afterwards. In case of default, the buyer is compensated for his loss, but if there is no default, no compensation is paid and the seller receives the periodic protection fee of the credit default swap until maturity. Although the protection buyer enters into the CDS contract in order to remove his credit default risk, he still faces counterparty risk concerning the default of the protection seller and the replacement of the contract. Moreover, he faces basis risk when the protection instrument doesn’t match directly the hedged asset, e.g. different maturities between the hedged bond and the credit default swap contract. On the other hand, the protection seller takes a long position in the default risk of the reference entity. He also faces counterparty risk if the buyer becomes unable to pay the
protection fee. A credit default swap can be seen as an insurance contract with the only difference being that the protection buyer doesn’t have to hold the underlying asset in order to enter into the contract; investors are free to buy or short the protection and speculate about the probability of a credit event. This feature is what gives rise to the cheapest-to-deliver (CTD) option in physical settlements. In standard CDS contracts, there are usually many different assets from the reference entity that are accepted as a trade upon default, thus the protection buyer who didn’t hold a bond for the reference entity can now choose from a pool of different bonds trading at different prices to deliver to the protection seller.

Credit spread, on the other hand, is simply the difference between the interest rate of a risky bond and the risk-free rate. It is a measure of the risk an investor takes for investing in the bond market. Although this measure appears quite simple, finding the right risk-free proxy to calculate the credit spread can be challenging. The obvious choice would be to use long-term government bonds, but these are known to be unreliable because of the different factors that affect them such as the taxation treatment, repo specials, scarcity premium and their benchmark status. Another more reliable proxy proposed by McCauley (2002) would be the swap rates in the reference entity’s currency, e.g. the USD for American companies and the Euro for European companies. Swaps have the advantage of being available in practically unlimited quantities as they are synthetic and are also quoted on a constant maturity basis. However, since one leg of the swap is linked to the London Interbank Offered Rate (LIBOR), which is not a perfect risk-free measure, the swap rates contain some credit risk premium. Nonetheless, a n-year swap rate can be considered as the interest rate on a n-year bond where the emitting entity has a 0 probability of change in its credit rating at the beginning of each accrual period, which is 6 months for American plain vanilla swap, but can go up to 12 months in other markets (Collin-Dufresne and Solnik (2001)). Since the one-year probability of default for investment-grade companies is practically 0, it makes the swap rates virtually risk-free and as such a good proxy for that rate. Duffie (1999) and Houweling and Vorst (2002) also propose the use of general collateral or repo rates rather than swap rates. Although these might represent a better proxy for the risk-free rate, they are only available up to one year and thus cannot be used to calculate the credit spread for bonds of longer maturity.
2.3. Empirical Model

2.3.1. Cointegration

Cointegration can be tested in a vector autoregressive (VAR) framework. The model has the form

\[ X_t = \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + D_{lt} + \mu + Z_t, \]

where \(\mu\) is a vector of restricted constants, \(D_{lt}\) is a matrix of dummy variables and \(Z_t \sim IID(0, \Sigma)\). The model can be rewritten in the error correction form as

\[ \Delta X_t = \Phi X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + D_{lt} + \mu + Z_t, \]

where \(\Phi = \sum_{i=1}^{p} \Phi_i - I\) and \(\Gamma_i = -\sum_{j=i+1}^{p} \Phi_j\). The dummy variables are unrestricted so as they do not enter the cointegration vectors and will not affect the asymptotic distribution (Juselius (2006)). The process has unit roots when the determinant of the characteristic polynomial, \(\Phi(z) = I - \Phi_1 z - \cdots - \Phi_p z^p\), has roots at one, i.e. when \(\det(\Phi(z)) = 0\) has roots at one. When there are fewer unit roots than series, some of them have roots in common and the process \(\{X_t\}\) is cointegrated. In other words, when the number of linearly independent cointegration vectors is less than the number of series, i.e. when the matrix \(\Phi\) has reduced rank, the system is cointegrated. Cointegration can furthermore formally be stated under Johansen’s version of Granger’s representation theorem (Johansen (1996)). Making the definition \(\Gamma = I - \sum_{i=1}^{p-1} \Gamma_i\), the previous system \(\{X_t\}\) with roots at either \(|z| > 1\) or \(|z| = 1\) from \(\det(\Phi(z)) = 0\) is cointegrated if and only if \(\Phi = \alpha \beta'\) where \(\alpha\) and \(\beta\) are \(m \times r\) matrixes of full rank \(r < m\) and \(\alpha' \Gamma \beta_\perp\) has full rank \(m - r\). \(\alpha_\perp\) and \(\beta_\perp\) are the orthogonal complements of \(\alpha\) and \(\beta\) such as \(\alpha' \alpha_\perp = 0\) and \(\beta' \beta_\perp = 0\).

Cointegration is used to test if two integrated series are affected by the same stochastic trend, which can be removed by taking a suitable linear combination of the process, thereby making it stationary (Johansen (1996)). As discussed above, in the case of CDS prices and credit spreads, both series should be equal and thus the cointegration vector should be \([1, -1, 0]'\). However, a restricted constant is added to the vector to account for the factors affecting the series and making the basis persistently different from 0. A restricted constant is preferred in this case because an unrestricted constant would allow for a linear trend in each of the variables.
The cointegration rank can be determined by a likelihood ratio trace test (Johansen, 1996a) with test statistic:

\[ Q(r) = -n \sum_{i=r+1}^{m} \log(1 - \hat{\lambda}_i), \]

where the eigenvalues, \( 1 > \hat{\lambda}_1 > \cdots > \hat{\lambda}_m \), are the squared partial canonical correlation between \( \Delta X_t \) and \( X_t \), and \( n \) is the size of the sample. The hypotheses tested are \( H_0: \text{rank}(\Phi) = r \) against \( H_a: \text{rank}(\Phi) = m \) where \( r = 0, 1, 2, \ldots, m-1 \) and \( m \) is the full rank of the matrix \( \Phi \). If \( r = m \), the matrix \( \Phi \) is of full rank and thus \( \{X_t\} \) is stationary and I(0). This is equivalent to \( \det(\Phi(z)) \neq 0 \) for \( |z| = 1 \) with the characteristic polynomial. When \( r = 0 \), there are \( m \) unit roots and the process \( \{X_t\} \) is I(1) and there are no cointegration relationships. There can only be cointegration when \( 0 < r < m \) and the matrix \( \Phi \) has reduced rank. The distribution of the trace statistic is non-standard and was derived by Johansen (1996) with its critical values.

Cointegration tests like the Johansen trace test used in small finite samples have been shown to display an upward bias, i.e. the estimated rank is larger than the true rank, in presence of conditional heteroskedasticity (Cavaliere et al. (2010a)) and non-stationary heteroskedasticity (Cavaliere et al. (2010b)), most particularly in the presence of permanent change in the error variance. Furthermore, Ahlgren and Catani (2017b) show that the trace test has low power in sample under 1000 observations because of the strong persistence and very high persistence in volatility present in cointegrated systems of CDS prices and bond spreads which may lead to the non-rejection of the null hypothesis of no cointegration. Other possible methods to test for cointegration are the bootstrap and wild bootstrap tests, which are extensions of the Johansen model. Ahlgren and Catani (2017b) also show that the asymptotic trace test and bootstrap test are unreliable when the distribution of the errors is heavy-tailed with finite variance and infinite fourth moment, and that CDS and credit spread series usually have these characteristics. The wild bootstrap test is thus preferred in these conditions. In their paper, Cavaliere et al. (2015) show that the wild bootstrap and BIC tests give virtually the same results for large samples when testing for cointegration.

If there are evidences suggesting that there is conditional or non-stationary heteroskedasticity in the residuals, the information criteria can be used as an alternative to determine the presence of cointegration. As shown by Cavaliere et al. (2015), the
Bayesian information criterion (BIC) is more reliable than the likelihood ratio trace for finite samples in these conditions. However, since the sample used in this article is large, ARCH effects are not a problem and the cointegration rank keeps its asymptotic distribution (Cavaliere et al. (2010a)). The information criterion is:

\[
BIC (r) = T \log |\widehat{\Sigma}^{(r)}| + \log(T) \pi(r)
\]

where \(\widehat{\Sigma}^{(r)}\) is the residual covariance matrix estimated from the VAR model and \(\pi(r)\) is a function of the number of parameters. The rank that minimise the information criterion is thus the rank of cointegration between the two time series. In the article, the BIC test is also used to determine cointegration rank between the series rather than the bootstrap and wild bootstrap, and the results are compared with the asymptotic trace test.

It is also worth to consider the addition of control variables in the VAR in order to exclude the effect of certain factors on the cointegration relationship. However, it is shown by Engle et al. (1983) that the analysis of such models is only valid under the assumption of weak exogeneity and even then, Harbo et al. (1998) show that the deterministic term makes it very difficult to determine the cointegration rank without modelling the full system. Thus, this is beyond the scope of the article and the VAR model used to determine the cointegration rank in this paper will be similar to those of previous studies.

### 2.3.2. Vector error correction model and price discovery

Theory predicts that there is a long-term relationship between the CDS prices and credit spreads, but the way that the two series interact with each other in the short run is also important to the comprehension of credit risk. Price discovery is a measure of the dynamic short-term behaviour of the two series with a focus on the lead-lag relation. It is the efficient and timely incorporation of the implicit information in investor trading in the market prices (Lehmann (2002)) and it measures the degree to which the prices in a particular market adjust to eliminate the pricing errors from the long-run equilibrium relation in the previous period (Davidson et al. (1978)).

Although price discovery is a short-term relationship, it relies on Johansen’s version of Granger’s representation theorem (Johansen (1996)) for the vector error correction model (VECM) and thus there needs to be cointegration in order to get the measurement. In case of non-cointegration, the simpler measure of the Granger causality can be used to test for causality between the series. According to Blanco et al. (2005), price discovery
should occur in the market with the highest number of transactions by informed traders. Because the synthetic nature of the credit derivatives market allows for a lot more flexibility and liquidity than the cash market, participants in the CDS market can take both long and short positions in larger contracts and transact more than on the bond market. As such, theory would suggest that the CDS prices would lead the credit spreads.

When there is cointegration, the VAR model in the previous section can be rewritten as

$$\Delta X_t = \alpha \beta' X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + D_{lt} + \alpha \rho + Z_t,$$

where $\beta$ are the cointegration vectors and $\beta' X_{t-1}$ are the stationary and I(0) cointegration relations. The matrix $\beta' X_{t-1}$ represent the long-term equilibrium of the I(1) variables $\{X_t\}$ and can be express as $\beta' X_{t-1} = c$ where $c$ is a constant. The parameters $\alpha$ are the adjustment at time $t$ for a disruption in the equilibrium $\beta' X_{t-1} - c$ at time $t-1$ and $\rho$ is the restricted constant. The model also allows to isolate the common trend as $\alpha' \sum_{i=1} Z_i$ where $Z_t \sim IID(0, \Sigma)$. When $\alpha_i$ is significant, it means that $\{X_i\}$ will adjust to the other variables when there is a disruption of equilibrium in the previous period. In other words, the other variables contribute more to price discovery than $\{X_i\}$. Price discovery can be tested for all the variables to determine their relationship with the cointegration relations.

When there is no evidence of cointegration between the CDS prices and credit spreads, the VECM cannot be used as $\alpha \beta'$ is not defined; Johansen’s version of Granger’s representation theorem is only valid when $0 < r < m$. The simpler measure of Granger causality in VAR-in-difference has to be used to test for price discovery. The VAR-in-difference is used rather than the vector autoregressive model since the variables are not themselves stationary and the test require two stationary and I(0) series. A variable Granger cause another one if its past values significantly help to explain the present value of the other variable. The significance of the coefficients can be tested using a $t$-test for any particular values or using an $F$-test to test for the joint significance.

### 2.4. Data

The data include six large American banks: Bank of America, Citigroup, Goldman Sachs, JP Morgan, Morgan Stanley, and Wells Fargo. The data are collected from Thomson Reuters Eikon and range from 22 February 2010 to 30 December 2016. The bonds used
have fixed coupons and exclude high yield instruments and instruments with integrated features such as convertible bonds, callable bonds or index linked bonds. The swap rate is used as a risk-free proxy to calculate the credit spread. Since CDS prices are quoted on a constant maturity of 5 years, the bonds rate have to be interpolated in order for the risk structures to be comparable since risk increases with maturity, all else being equal. The interpolated rate of return is then

\[ YTM_{low} - \frac{(YTM_{high} - YTM_{low})}{(TTM_{high} - TTM_{low})} \times (TTM_{low} - \text{maturity}) = \text{interpolated rate}, \]

where \( YTM_{low} \) is the yield-to-maturity of the bond with maturity lower than 5 years, \( YTM_{high} \) is the yield-to-maturity of the bond with maturity higher than 5 years, \( TTM_{low} \) is the time-to-maturity of the bond with the lower maturity, \( TTM_{high} \) is the yield-to-maturity of the bond with the higher maturity and \( \text{maturity} \) is the required maturity.

To compute the interpolated rate, a bond with maturity of 5 years and another one with maturity of 6.5 years from the interpolation starting point are used. The interpolation period is 1.5 year so that the bonds are not too close to maturity. When the lower maturity bond reaches a time to maturity of 3.5 years, a new bond is used for the longer maturity of 6.5 years to this point and the previous long maturity bond, with now 5 years to maturity, is used as the shorter maturity bond. This procedure is repeated until the rates are interpolated for the whole sample. A summary of the data is presented in Table 1.

### Table 1  Data summary

<table>
<thead>
<tr>
<th></th>
<th>Number of observations</th>
<th>Mean basis</th>
<th>Mean absolute basis</th>
<th>Standard deviation</th>
<th>Standard deviation 2011-2012</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of America</td>
<td>1724</td>
<td>-69.39</td>
<td>79.26</td>
<td>51.51</td>
<td>75.14</td>
<td>-188.22</td>
<td>183.54</td>
</tr>
<tr>
<td>Citigroup</td>
<td>1709</td>
<td>-20.95</td>
<td>38.47</td>
<td>51.31</td>
<td>84.56</td>
<td>-304.01</td>
<td>165.57</td>
</tr>
<tr>
<td>Goldman Sachs</td>
<td>1716</td>
<td>-54.79</td>
<td>75.27</td>
<td>67.38</td>
<td>95.71</td>
<td>-198.93</td>
<td>260.98</td>
</tr>
<tr>
<td>JP Morgan</td>
<td>1775</td>
<td>-38.86</td>
<td>40.09</td>
<td>39.08</td>
<td>66.34</td>
<td>-137.43</td>
<td>25.78</td>
</tr>
<tr>
<td>Morgan Stanley</td>
<td>1708</td>
<td>-39.24</td>
<td>61.99</td>
<td>72.01</td>
<td>117.18</td>
<td>-281.70</td>
<td>406.68</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>1730</td>
<td>-33.73</td>
<td>34.43</td>
<td>29.61</td>
<td>45.58</td>
<td>-171.14</td>
<td>38.54</td>
</tr>
</tbody>
</table>

*Table of descriptive statistics. The number of observations range from 1708 to 1775 due to missing values for some series. All other numbers are reported in basis points.*
The bonds used have maturities close to January 2015, June 2016, January 2018, June 2019, January 2021 and June 2022. The series have varying number of observations due to the unavailability of data for certain days. In the interpolation, when a day is missing for one of the bonds, it is also taken off the other bond in order to equate the dates. The CDS data are mid-spread data from senior 5-years CDS. The final credit spread data are also very different in length from the CDS data and some observations are deleted when data are missing in the other series so that both series have the same number of observations.

There is some concern with the data which could have some impact on the basis and the cointegration relationship between the two units of credit risk. First of all, the bond payment and the CDS payment do not necessarily have matching dates. Most of the bonds used have bi-annual coupon payments, although when such bonds are not available for interpolation, monthly notes are used. The bonds thus generally have only two payment dates corresponding to the CDS payments: January and June. Second, the bonds used in the interpolation do not necessarily trade close to par. Some of the instruments have lower required yield and others higher. Third, although the interest rates were low during the time period under study because of the government incentives to recovery from the last financial crisis, the yield curve was not flat. Even though these factors can influence the results of the study, the data are the best available in the circumstances.

When comparing with data used in previous research (e.g. Blanco et al. (2005) and Zhu (2006)), it can be noticed that they are more volatile. The credit markets have changed a lot since the early 2000’s and it has affected the relation between the two units of credit risk. The early 2000’s period was calm and thus the two series were not diverging as much as the series used in this article. In the short-run, CDS prices and credit spreads can differ largely due to market imperfections and thus the data shown in the table are not surprising, mostly since the series cover a longer time span allowing for the presence of more extreme observations. Figure 2 displays the graph of the two series.

As it shows, the credit spread curve is usually over the CDS price curve except for some instance in 2011-2012 when CDS values are much higher than credit spread values. Otherwise the series appear to be moving relatively well together and there seems to be a discernible linkage in their fluctuations.
Figure 2  Credit spread and CDS price series

Bank of America

Citigroup

Goldman Sachs

JP Morgan

Morgan Stanley

Wells Fargo

SWSP  CDS

SWSP  CDS

SWSP  CDS

SWSP  CDS

SWSP  CDS
To analyse the series further, the Augmented Dickey-Fuller test is performed to see if they contain a unit root, as it is expected. The results show that all the series cannot reject the unit root at the 1% significance level. The only problem comes from the CDS series of JP Morgan which rejects the unit root at the 5% significance level. Cointegration tests will still be performed on the series and further test will look deeper into the behaviour of the series within the model. The results are summarized in Table 2.

### Table 2  Unit root tests

<table>
<thead>
<tr>
<th></th>
<th>k</th>
<th>ADF SWSP</th>
<th>ADF CDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of America</td>
<td>2</td>
<td>−1.45</td>
<td>−1.87</td>
</tr>
<tr>
<td>Citigroup</td>
<td>3</td>
<td>−2.08</td>
<td>−2.09</td>
</tr>
<tr>
<td>Goldman Sachs</td>
<td>2</td>
<td>−1.97</td>
<td>−2.09</td>
</tr>
<tr>
<td>JP Morgan</td>
<td>2</td>
<td>−1.31</td>
<td>−3.07</td>
</tr>
<tr>
<td>Morgan Stanley</td>
<td>4</td>
<td>−1.61</td>
<td>−1.81</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>3</td>
<td>−1.65</td>
<td>−2.10</td>
</tr>
</tbody>
</table>

*Augmented Dickey-Fuller with lagged differences equal to k. *** and ** represent the 1%, 5% and 10% significance level respectively.*

### 2.5. Results

Once the series are constructed, the first step in the empirical method is to select the numbers of lags for the vector autoregressive model. The test is made in Oxmetrics with a maximum number of 5 lags to cover for the entire week. When the information criteria give different results for the optimal number of lags, the Hannan-Quinn information criterion is used for the selection. The optimal numbers of lags, reported in Table 3, are 2 for Bank of America, 3 for Citigroup, 2 for Goldman Sachs, 2 for JP Morgan, 4 for Morgan Stanley and 3 for Wells Fargo.

An unrestricted VAR model is then estimated by ordinary least squares (OLS) and observations further than 8 standard deviations from the mean are isolated to create dummy variables. The model is re-estimated with the dummies and a test for integration of order 1 is performed to determine cointegration rank. The unrestricted model is
estimated in Oxmetrics and contains the credit spread series, the CDS price series and a restricted constant. The model has the form

$$X_t = \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + D_{lt} + \mu + Z_t,$$

where $\mu$ is a vector of restricted constants, $D_{lt}$ is a matrix of unrestricted dummy variables for observations further than 8 standard deviations from the mean and $Z_t \sim IID(0, \Sigma)$. The model can be rewritten as

$$\Delta X_t = \Phi X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + D_{lt} + \mu + Z_t,$$

where $\Phi = \sum_{i=1}^{p} \Phi_i - 1$ and $\Gamma_i = -\sum_{j=i+1}^{p} \Phi_j$. The vector $X$ is composed of the two variables $P_{CDS}$ and $P_{CS}$ which represent the price of the credit default swap and credit spread for company $j$. The value of $m$, the full rank of matrix $\Phi$, is 2. The hypothesis to be tested are thus $H_0: \text{rank}(\Phi) = 0$ against $H_a: \text{rank}(\Phi) = 2$. If the test does not reject the null hypothesis, the system is not cointegrated. If it rejects the null hypothesis, a new test is made with the hypothesis $H_0: \text{rank}(\Phi) = 1$ against $H_a: \text{rank}(\Phi) = 2$. If the null hypothesis is not rejected, we can conclude that there is cointegration. However, if the test rejects the null hypothesis, the matrix $\Phi$ is of full rank and $\{X_t\}$ is stationary and I(0); there is no cointegration.

Cointegration is found for 5 of the 6 series: Bank of America, Citigroup, JP Morgan, Morgan Stanley and Wells Fargo. For Goldman Sachs, rank 0 is not rejected and no cointegration relationship is found. From the graph of the series (Figure 2), it can be seen that in the end of 2010 and during 2011, CDS prices appreciated considerably compared to credit spreads and became higher for a long period in 2011. This period is taken away from the observations and new tests are performed on the subseries starting on 3 January 2012. The subseries rejects rank 0, however it also rejects rank 1 at the 10% significance level, but not at the 5% significance level. There is thus weak evidence for cointegration and further tests are performed using the subseries. Cointegration is also tested without dummy variables. The test leads to similar results except for Citigroup where only weak evidence for cointegration (rejection of rank 0 at the 10% level) is found, as opposed to clear rejection when using dummy variables. This gives support to the necessity to exclude outliers from cointegration analysis.
Cointegration is also tested in R using a code written by Catani (2011) for the information criterion. The Bayesian information criterion is more reliable than the asymptotic trace test for finite samples when the residuals of the model show presence of conditional and non-stationary heteroskedasticity (Cavaliere et al. (2015)). Rank 1 is found only for Bank of America and the test rejects cointegration relationships, i.e. it selects $r = 0$, for all other companies. This might be due to the data collection problems discussed in the previous section such as non-matching payment dates, instruments which do not trade at par and a yield curve that is not flat. Although this did not seem to be a problem with the asymptotic test, the information criterion seems to be more sensitive to those considerations and may be too restrictive.

Further tests are performed in Oxmetrics on the series that do not reject cointegration with the Johansen test. The tests are performed with a cointegrated VAR system and restrictions are added to test for the validity of the theoretical cointegration vectors $[1, -1, c]'$ and $[1, -1, 0]'$. The test does not reject the parity relationship up to a constant for two companies at the 5% significance level. For Bank of America, the test rejects the $[1, -1, c]'$ vector at the 10% significance level, but when the restricted model is compared to the model without restriction, the former minimises the information criteria leading to some support of the equal pricing of credit risk up to a constant in the two markets in the long-run for this bank. For the other restriction, the zero long-run basis is not rejected for only one companies; only Citigroup could not reject the vector $[1, -1, 0]'$ at all significance level which means that only this bank prices credit risk equally in the long-run in both market. The results are summarised in Table 3. The last column of the Table shows the largest stationary root of the companion matrix when $r = 1$. The results go from 0.9716 to 0.9885 and support strong persistence in the series.

Many factors can lead to departure from parity in the CDS price and credit spread relationship. Market imperfections can persist for long period of time and decrease the arbitrage opportunities which would lead to persistent mispricing in the credit markets. One of these imperfections is the presence of significant repo costs when shorting bonds which pushes the prices of CDS higher than the bond spreads because of the inability of market participants to capitalise on arbitrage opportunity (Duffie (1999)). Another instance that can push the CDS prices upwards is the presence of cheapest-to-deliver options when there is physical settlement upon default and this option cannot be valued analytically because there is no benchmark on the post default value of deliverable bonds (Blanco et al. (2005)). Another factor which could have a negative effect on cointegration
is the presence of high transaction costs relative to arbitrage opportunities which allows them to persist in time.

### Table 3  Cointegration rank

<table>
<thead>
<tr>
<th></th>
<th>k</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Root</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Johansen trace test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>r=0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bank of America</td>
<td>2</td>
<td>26.12***</td>
<td>5.11</td>
<td>1</td>
<td>3.41*</td>
<td>11.19***</td>
</tr>
<tr>
<td>Citigroup</td>
<td>3</td>
<td>27.67***</td>
<td>1.20</td>
<td>0</td>
<td>0.46</td>
<td>3.77</td>
</tr>
<tr>
<td>Goldman Sachs</td>
<td>2</td>
<td>11.44</td>
<td>3.04</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Subseries</td>
<td>2</td>
<td>26.69***</td>
<td>8.58*</td>
<td>0</td>
<td>9.48***</td>
<td>12.79***</td>
</tr>
<tr>
<td>JP Morgan</td>
<td>2</td>
<td>26.01***</td>
<td>1.20</td>
<td>0</td>
<td>19.18***</td>
<td>22.12***</td>
</tr>
<tr>
<td>Morgan Stanley</td>
<td>4</td>
<td>44.06***</td>
<td>8.13*</td>
<td>0</td>
<td>4.06**</td>
<td>11.59***</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>3</td>
<td>30.41***</td>
<td>2.69</td>
<td>0</td>
<td>12.28***</td>
<td>20.95***</td>
</tr>
</tbody>
</table>

Results of cointegration tests on the model. The constant k represents the number of lags in the model. The next three columns present the LR statistic of $H_0$: rank(Φ) = i against $H_a$: rank(Φ) = 2, where i = {0, 1}, and the optimal rank selected by the Bayesian criterion. $[1, -1, c]'$ and $[1, -1, 0]'$ are the LR statistics of the restriction tested on the cointegration vector. The last column shows the largest root of the companion matrix when r = 1. ***, ** and * represent the 1%, 5% and 10% significance level respectively. Dummy variables were used for outliers further than 8 standard deviations from the mean.

In the beginning of the period, the economy of the United-States was still recovering from the financial crisis. From Figure 2, CDS prices and credit spreads increased, which means that there was an increase in demand for CDS and a decrease in demand for bonds. From Table 1, the volatility for this period is significantly higher than when the whole sample is considered. This could be explained by a “run-to-safety” from the investors demanding more security on their investments and increasing the demand for protection. The demand for protection seemed to be so high that the price of the CDS became higher than the credit spread for most of the banks under observation. These factors could also explain the large departures from equilibrium for the same period in Figure 3 which shows the estimates of the cointegration relations from the VAR models. It can also be seen from the Figure that Goldman Sachs has an extreme departure from cointegration in 2012 before going back to relatively more normal levels around 2014.
Figure 3  Cointegration relationships

Bank of America

Citigroup

Goldman Sachs Subseries

JP Morgan

Morgan Stanley

Wells Fargo
2.5.1. **Price discovery**

For price discovery, the vector error correction model is fitted to the series which could not reject cointegration. To pursue with a similar framework as the previous literature (e.g. Blanco et al. (2005) and Zhu (2006)) the VECM is specified as follow

\[
\begin{bmatrix}
\Delta CDS_t \\
\Delta CS_t
\end{bmatrix} = \begin{bmatrix}
\alpha_1 \\
\alpha_2
\end{bmatrix} (CDS_{t-1} - \beta CS_{t-1} - \rho) + \sum_{i=1}^{p-1} \begin{bmatrix}
\Gamma_{11,i} & \Gamma_{12,i} \\
\Gamma_{21,i} & \Gamma_{22,i}
\end{bmatrix} \begin{bmatrix}
\Delta CDS_{t-i} \\
\Delta CS_{t-i}
\end{bmatrix} + D_{i,t} + Z_t
\]

where \((CDS_{t-1} - \beta CS_{t-1} - \rho)\) is the expansion of the cointegration relationship \(\beta'X_{t-1}\) with the restricted constant \(\rho\). The parameters \(\rho\) and \(\beta\) can be restricted when there is evidence for the cointegration vector to be \([1, -1, 0]'\) or \([1, -1, c]'\). Otherwise, they are allowed to vary freely according to the model. The estimated adjustment coefficients \(\hat{a}_1\) and \(\hat{a}_2\) measure the degree of adjustment of a particular market to eliminate the pricing errors of the long-term relation. By the Granger representation theorem, at least one of the markets has to adjust to the other under the presence of cointegration (Engle and Granger (1987)). If the CDS market contributes to price discovery, the estimated adjustment coefficient \(\hat{a}_2\) will be positive and statistically significant meaning the credit spread market adjusts to compensate for pricing error. On the other hand, if \(\hat{a}_1\) is negative and statistically significant, it means that the CDS market adjusts for errors and that the bond market contributes significantly to price discovery. When both estimates are significant, the relative magnitude of the two coefficients will determine which market contributes most to price discovery.

In order to find out which of the two market leads in term of price discovery, the estimates of error corrections have to be manipulated, leading to the Hasbrouck (Hasbrouck (1995)) and Gonzalo-Granger (Gonzalo and Granger (1995)) measures. The Hasbrouck model assumes that price volatility reflects new information and as such, the market that contributes the most to the volatility of the innovations to the common factor will also be the one that contributes the most to price discovery. On the other hand, the Gonzalo-Granger model does not look at the correlation between the markets, but decomposes the factor itself and attributes the price discovery to the market that adjusts the least to movements in the other market. The expressions for these measures with respect to the contribution of the CDS market are

\[
HAS_{low} = \frac{\alpha_2^2(\sigma_1^2 - \sigma_{12}^2)}{\alpha_2^2 \sigma_1^2 - 2 \alpha_1 \alpha_2 \sigma_{12} + \alpha_1^2 \sigma_2^2},
\]
\[ \text{HAS}_{\text{high}} = \frac{(\alpha_2 \sigma_1^2 - \alpha_1 \sigma_{12})^2}{\alpha_2^2 \sigma_1^2 - 2 \alpha_1 \alpha_2 \sigma_{12} + \alpha_1^2 \sigma_2^2}, \]

\[ GG = \frac{\alpha_2}{\alpha_2 - \alpha_1}, \]

where \( \sigma_{12}, \sigma_1 \) and \( \sigma_2 \) represent the variance-covariance matrix of \( \varepsilon_{1,t} \) and \( \varepsilon_{2,t} \). Analogous expressions are used to measure the contribution of the credit spread market. The Hasbrouck measure gives an interval of the price discovery contribution and Baillie et al. (2002) find that the average of the two bounds of the Hasbrouck interval is a good estimate of the price discovery which can then be compared to the Gonzalo-Granger measure. When the measures are close to 1 it means that the CDS prices lead the credit spread and that the cash market adjust afterwards for price discrepancies. When the measures are close to 0, the credit spread has a leading role in the pricing of credit risk and the derivative market adjust with a delay. When the measures are close to 0.50, both markets contribute equally to price discovery.

For the series which do not reject cointegration and \([1, -1, c]'\) with the Johansen trace test, \( \beta \) was restricted to 1 in the model. For Citigroup, \( \rho \) was also restricted to 0. Otherwise, no restrictions were applied to \( \beta \) and \( \rho \). The models are estimated in Oxmetrics and the results for the estimates of \( \hat{\alpha}_1 \) and \( \hat{\alpha}_2 \) are reported in Table 4 along with the Hasbrouck and Gonzalo-Granger measures.

Five out of the six companies have both coefficients correctly signed; \( \hat{\alpha}_1 \) is negative and \( \hat{\alpha}_2 \) is positive. Furthermore, \( \hat{\alpha}_1 \) is generally not significant which means that the bond market does not play a role in price discovery. \( \hat{\alpha}_2 \) is significant at the 1% level for four companies which gives support to the role of the CDS market in price discovery. The leading role of the credit derivative market is also supported by the Hasbrouck and Gonzalo-Granger measures for four of the five companies. The first one ranges from 0.86 for Citigroup to 0.97 for Bank of America, which means that the CDS market contributes most of price discovery. The Gonzalo-Granger measure gives lower results from 0.75 for Citigroup to 0.84 for Bank of America, but still shows that the clear majority of price discovery happens in the CDS market. For JP Morgan, the measures show that the majority of price discovery happens in the bond market. Surprisingly, the Gonzalo-Granger measure gives a higher contribution to the CDS market than the Hasbrouck measure. The measures are 0.08 and 0.25 for Hasbrouck and Gonzalo-Granger respectively. For Goldman Sachs’ subseries, \( \hat{\alpha}_1 \) and \( \hat{\alpha}_2 \) are both negative and not
significant. Although the measures seem to point for a bigger contribution to price discovery for the bond market, no conclusions can be made for this series. The results are mostly in line with previous studies. As Blanco et al. (2005) mention, price discovery happens in the credit derivative market because of the higher level of liquidity and higher frequency of transactions by informed traders in this market. The results show, however, that this is not always the case.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\sigma}_1$</th>
<th>$\hat{\alpha}_2$</th>
<th>$\hat{\sigma}_2$</th>
<th>$\hat{\sigma}_{12}$</th>
<th>HAS Low</th>
<th>HAS High</th>
<th>HAS Mid</th>
<th>GG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of America</td>
<td>(0.003)</td>
<td>6.835</td>
<td>0.015***</td>
<td>5.933</td>
<td>0.029</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.84</td>
</tr>
<tr>
<td>Citigroup</td>
<td>(0.005)</td>
<td>5.517</td>
<td>0.014***</td>
<td>8.247</td>
<td>(0.049)</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.75</td>
</tr>
<tr>
<td>Goldman Sachs</td>
<td>(0.001)</td>
<td>3.092</td>
<td>(0.0001)</td>
<td>4.263</td>
<td>(0.013)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>(0.11)</td>
</tr>
<tr>
<td>JP Morgan</td>
<td>(0.009)***</td>
<td>3.252</td>
<td>0.003</td>
<td>4.110</td>
<td>0.070</td>
<td>0.08</td>
<td>0.09</td>
<td>0.08</td>
<td>0.25</td>
</tr>
<tr>
<td>Morgan Stanley</td>
<td>(0.002)</td>
<td>8.153</td>
<td>0.013***</td>
<td>8.622</td>
<td>0.010</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.84</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>(0.006)</td>
<td>2.840</td>
<td>0.020***</td>
<td>5.060</td>
<td>0.022</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Results of the VECM. $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are the coefficients on the cointegration vector, $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are the standard deviation of the model residuals and $\hat{\sigma}_{12}$ is the covariance between the residuals. ***, ** and * represent the 1%, 5% and 10% significance level respectively.

For the whole series of Goldman Sachs, no cointegration is found and thus the VECM cannot be used and price discovery cannot be evaluated. The simpler measure of Granger causality is used to analyse the interaction between the variables. The VAR-in-difference model shows that none of the variables Granger cause the other. The coefficients for the past values of CDS price are not individually or jointly significant for credit spread and the past values of credit spread are not individually or jointly significant for CDS price. For this firm, it seems that the two series have moved independently through the period.

2.5.2. Tests

The residuals are tested to demonstrate the adequacy of the models used in the empirical method. The residuals of the cointegrated VAR are saved and both series are modelled
with a GARCH (1, 1) model to test for persistence in volatility. The variance of the model has the form

\[
\sigma_t^2 = \omega + \alpha Z_{t-1}^2 + \beta \sigma_{t-1}^2,
\]

where \(\omega\) is a constant and \(Z_t = \sigma_t e_t, e_t \sim NID(0, 1)\). The diagonal matrices of parameters \(A\) and \(B\) are created with the estimated parameters \(\hat{\alpha}\) and \(\hat{\beta}\), respectively. The stationarity condition of the model is that all the eigenvalues of \(A + B\) are less than 1 in modulus (Engle and Kroner (1995)). Using the notation \(\lambda(A + B)\) to represent the largest eigenvalue of matrix \(A + B\), the stationary condition is \(\lambda(A + B) < 1\) (He and Teräsvirta (2004)). This condition is satisfied only for the subseries of Goldman Sachs. As Carnero et al. (2007) show, GARCH overestimates parameters in the presence of outliers which would cause the estimators to be biased, i.e. even if the data-generating process is stationary, the estimates, in the presence of outliers, would not satisfy the stationarity condition. However, the results support the high persistence in volatility as in Ahlgren and Catani (2017b) which leads to think that the trace test may be oversized. The residuals are also tested for ARCH behaviour, autocorrelation and non-stationary heteroskedasticity. The results show that there is evidence of ARCH effect in all the CDS residual series and in three of the bond series. Cavaliere et al. (2010a) show that ARCH effects does not affect the asymptotic distribution of the cointegration rank statistics. Since the sample under observation is large, conditional heteroskedasticity will not affect the results of the previous section considerably. The Wild Bootstrap (WB) method with heteroskedasticity-consistent covariance matrix estimator (HCCME) developed by Ahlgren and Catani (2017a) is used to calculate the residual serial correlation. Ahlgren and Catani (2017a) find that the asymptotic tests for error autocorrelation (e.g. Breusch-Godfrey LM test) which are based on the IID error assumption will falsely reject the null hypothesis of no serial correlation in the presence of conditional heteroskedasticity in the residuals. The authors also show that the WB test performs well in these conditions. The test is calculated in R with the package VARTests of Belfrage (2017). The results show that there is presence of residual autocorrelation only for the series of JP Morgan when five lags of residuals are considered. The variance ratio test (Lo and MacKinlay (1989)) statistic is not significant for any of the series; all the non-stationary heteroskedasticity is captured by the model. The results are summarised in Table 5 and 6.
### Table 5  Estimates of persistence in volatility

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\alpha} + \hat{\beta} )</th>
<th>( \lambda(\Delta + B) &lt; 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWSP</td>
<td>CDS</td>
<td></td>
</tr>
<tr>
<td>Bank of America</td>
<td>0.959</td>
<td>1.057</td>
</tr>
<tr>
<td>Citigroup</td>
<td>1.008</td>
<td>1.045</td>
</tr>
<tr>
<td>Goldman Sachs</td>
<td>0.645</td>
<td>0.994</td>
</tr>
<tr>
<td>JP Morgan</td>
<td>0.977</td>
<td>1.016</td>
</tr>
<tr>
<td>Morgan Stanley</td>
<td>0.998</td>
<td>1.048</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>0.987</td>
<td>1.026</td>
</tr>
</tbody>
</table>

Sum of the GARCH parameters for the bond series and the CDS series. \( \lambda(\Delta + B) < 1 \) is the stationarity condition. ***, ** and * represent the 1%, 5% and 10% significance level respectively.

### Table 6  Tests on the residuals

<table>
<thead>
<tr>
<th></th>
<th>( Q_{LM,HC3}^{WB} )</th>
<th>Variance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h=1</td>
<td>h=5</td>
</tr>
<tr>
<td>SWSP</td>
<td>CDS</td>
<td></td>
</tr>
<tr>
<td>Bank of America</td>
<td>0.672</td>
<td>78.358 (***)</td>
</tr>
<tr>
<td>Citigroup</td>
<td>31.055 (***)</td>
<td>57.271 (***)</td>
</tr>
<tr>
<td>Goldman Sachs</td>
<td>0.012 (***)</td>
<td>44.779 (***)</td>
</tr>
<tr>
<td>JP Morgan</td>
<td>1.287 (***)</td>
<td>53.686 (***)</td>
</tr>
<tr>
<td>Morgan Stanley</td>
<td>13.262 (***)</td>
<td>73.521 (***)</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>11.004 (***)</td>
<td>53.561 (***)</td>
</tr>
</tbody>
</table>

Tests for model misspecifications. ARCH test (1-5), WB test with HCCME and variance ratio test. ***, ** and * represent the 1%, 5% and 10% significance level respectively.

The results from the tests performed on the residuals show that the model adequately fits the data in most cases; the ARCH effects are attenuated by the large sample under observation, there is no autocorrelation and no non-stationary heteroskedasticity. This strengthens the confidence in the results for cointegration with Johansen Trace test and invalidates the justifications of Cavaliere et al. (2015) for using the BIC in this case. It also shows that the BIC is too conservative and will falsely select rank 0 while there is evidence for rank 1. Figure 4 and Figure 5 plot the residual series.
Figure 4  Plot of the bond residuals

Bank of America

Citigroup

Goldman Sachs

JP Morgan

Morgan Stanley

Wells Fargo
Figure 5 Plot of the CDS residuals

Bank of America

Citigroup

Goldman Sachs

JP Morgan

Morgan Stanley

Wells Fargo
For JP Morgan, further tests with lagged residuals $h = 2, 3, 4$ show that there is evidence for serial correlation starting at lag 4. New VAR models with $p = 3, 4, 5$ are fitted to the data to try to capture all the autocorrelation in the model. The WB test performed on the residuals of the higher order VAR models show that increasing the number of lags up to 5 does not take the autocorrelation away from the residuals, which means that the model is misspecified and does not fit the data sample. The results of JP Morgan should therefore be interpreted with caution.

### 2.6. Conclusion

This article is a small contribution to the literature on cointegration analysis of credit spreads and CDS prices to understand better the dynamics between the two markets. It uses a long time series framework as opposed to many prior studies on the subject.

The purpose of the article is to empirically analyse the equilibrium relationship between CDS prices and credit spreads, and test for the presence of cointegration between the two series. The data show that the basis is much larger than in previous studies. This is due to the different time frame which allows for more deviation from parity and the relative tranquillity of the early 2000's period from which the data are collected in most major papers on the subject (e.g. Blanco et al. (2005) and Zhu (2006)). As in prior studies (e.g. Blanco et al. (2005), de Wit (2006), Norden and Weber (2009), Zhu (2006)), the results show cointegration for most of the companies analysed when the Johansen trace test is used. There is however rejection of cointegration for Goldman Sachs. A subseries is used for this bank for the further tests. When the Bayesian information criterion is used, cointegration is found in only one case, showing that the BIC is a lot more restrictive than the trace test.

On the series which accepted cointegration and the subseries of Goldman Sachs, further tests are made to determine the long-term parity relationship of the two series of credit risk. Two out of six cointegrated series accept the parity up to a constant, but only one firm accepts the zero long-run basis restriction. Because of the market imperfections which make arbitrage sometimes impossible for investors, it is normal that differences in the series can exist for a longer time.

The dynamics of the two markets was also analysed with price discovery to determine which market adapts the most to a change in the cointegration relationship in the previous period. The results show that for four of the series almost all of price discovery
happens in the CDS market, meaning that the bond market adapts the most to changes in the basis in the previous period. With the Hasbrouck measure, the CDS market contributes to 92% of price discovery on average and, with the Gonzalo-Granger measure, to 80%. These results are in line with previous studies that also find a higher contribution of the CDS market to price discovery (e.g. Blanco et al. (2005) and Zhu (2006)). For the other two banks, the results show that the bond market contributes more to price discovery, although the results were not significant for the subseries of Goldman Sachs. For the full series of the latter, which rejected cointegration, Granger causality was used instead of price discovery to analyse the dynamics of the credit markets. The results show that none of the market Granger causes the other, i.e. the past values of one market do not help explaining the current value of the other market, but only the past values of the same market are significant.

Tests on the residuals show that the model adequately fits the data and captures most of the features such as autocorrelation and non-stationary heteroskedasticity, giving support to the cointegration results between the two series. Furthermore, the ARCH effect in the residuals will not have a considerable effect on the results because of the large data sample used. The good specifications of the model lead to conclude that the BIC is too restrictive for tests of cointegration in this case and will select rank 0 even if there are ample evidences for rank 1.

For JP Morgan, the ADF test on the CDS series rejects the presence of a unit root, and further tests show that the residuals of the cointegrated VAR model are autocorrelated. This means that the model is misspecified and does not fit the data sample. The results of that bank should be interpreted with caution.

The article still leaves open many aspects of the relationship between credit spread and CDS price. An interesting direction for future studies would be an event study where the effect of a credit event is analysed and compared for the two markets to see the direct impact on them. This would complement the current understanding of the relationship between the CDS and bond markets so that not only their long-run pricing is understood, but also more directly their short-term reaction to a disruption in the credit risk of the underlying firm.
REFERENCES


Belfrage, M., 2017. VARtests: tests for error autocorrelation and ARCH errors in vector autoregressive models version 1.0.1. Available at: https://CRAN.Rproject.org/package=VARtests


