# THREE ESSAYS ON REIT CDS AND EQUITY RETURNS

by

Ke Shang

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Finance

Charlotte

2019

Approved by:

______________________________
Dr. Kiplan Womack

______________________________
Dr. Richard Buttimer

______________________________
Dr. Yufeng Han

______________________________
Dr. Paul Gaggl

# ABSTRACT


KE SHANG. Three essays on REIT CDS and equity returns. (Under the direction of DR. KIPLAN WOMACK)

The first paper ("An Analysis of REIT Credit Default Swap Pricing") first devise a closed-form solution of a non-arbitrage pricing model with observable factors in the default hazard rate to value Credit Default Swap (CDS). Then, I conduct panel regression to examine the explanatory power of REIT-specific and macroeconomic covariates in explaining the cross-sectional variation of CDS spreads. The high level of R-squared from the regression highlights the role of my list of fundamentals in determining credit risk. The second paper ("The Flow of Credit Risk Information among REIT Securities") discovers the credit risk information flow among REIT stocks, bonds, and CDS markets. In general, information flows from stocks to CDS, and then to bonds. However, there is a reversal of information flow around credit rating downgrades, where CDS leads stocks. Furthermore, I find evidence that large banks active in the CDS market can exploit private information obtained through their direct lending relationships. I conclude that the CDS market appears to be the primary market for trading on REIT credit risk information. The third paper ("The Predictability of REIT Index Returns") applies various machine learning and deep learning models to predict the out-of-sample REIT Index returns. Compared with traditional OLS method, machine learning algorithms significantly improve the predictability of REIT Index returns. To exploit the economic significances of machine learning models, I create a practical investment strategy which produces a substantial profit.

# ACKNOWLEDGEMENTS

First of all, I would like to humbly thank my dissertation chair Dr. Kiplan Womack. Thank you for your continuous support and guidance throughout this journey. Thank you for pushing me to excel whenever I feel like settling and thank you for being a true role model of how an inspirational professor and scholar should be.

I would also like to thank my dissertation committee members, Dr. Richard Buttimer, Dr. Paul Gaggl and Dr. Yufeng Han for your valuable comments and invaluable time spent on my dissertation. I am very lucky to have met you all during my study at UNCC.

Lastly, to all the faculty members and my fellow Ph.D. colleagues/friends in the Department of Finance who have been a part of my transition and growth during my Ph.D. student life, I hope the eternal sunshine will continue shining on your spotless minds.

# DEDICATION

I would like to dedicate my work to my parents, Shijun Shang and Ping Li, who offer unconditional love and support without expectations, and who have always maintained the incredible vision; to my friends in Charlotte who make my long graduate student life enjoyable

谨以此毕业论文献给我的父母，感谢你们开阔的视野以及对我无条件的支持与爱；献给我在夏洛特认识的朋友们，感谢你们让我的漫长的研究生生活变得多姿多彩。

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# INTRODUCTION

In a series of related research, my dissertation investigates the REITs Credit Default Swap market and the application of machine learning and deep learning on the REIT market.

My first essay " An Analysis of REIT Credit Default Swap Pricing " investigates the market quotes of credit default swaps (CDS) in the REIT industry. To better understand the market quotes of credit default swaps (CDS) in the REIT industry, we first devise a closed-form solution of a no-arbitrage model with observable factors in the default hazard rate to value CDS spreads. We then use panel regression to examine the explanatory power of REIT-specific and macroeconomic covariates in explaining the cross-sectional variation of CDS spreads. The R-squared of 86% highlights the role of our list of fundamentals in determining credit risk. Last, we test the impact of the implementation of regulations and standardization of CDS contracts (the “Big Bang”) by constructing an out-of-sample investment portfolio. We find that enactment of the “Big Bang” protocols improved CDS market efficiency and diminished market ambiguity.

In my second essay, "The Flow of Credit Risk Information Among REIT Securities ", I examine the flow of credit risk information among REIT CDS, REIT bond, REIT stock markets. How does credit risk information flow among REIT stocks, bonds, and credit default swaps (CDS)? More specifically, which security first capitalizes credit risk information, where does it spread from there, and are there times in which the established flow changes? Further, are some traders able to exploit private credit risk information in an industry known for transparent assets and operations, and if so, what is the effect? Results from our examination of these critical questions can be summarized as follows.

In general, information flows from stocks, to CDS, and to then bonds. CDS consistently leads bonds, and bonds never lead stocks or CDS. However, there is a reversal of information flow around credit rating downgrades, where CDS leads stocks. In this setting, CDS anticipate credit rating downgrades by nearly two months in advance. We find evidence that large banks active in the CDS market can exploit private information obtained through their direct lending relationships, which are arguably stronger and updated more frequently than in most other industries. We conclude that the CDS market appears to be the primary market for trading on REIT credit risk information.

The third essay in my dissertation package "The Predictability of REIT Index Returns" tests the statistical and economic importance of out-of-sample REITs predicted returns using various machine learning and deep learning models. This paper tests the statistical and economic significance of out-of-sample REITs predicted returns using several machine learning and deep learning, including random forest, XGBoost, stacked autoencoder, and LSTM. Generally, compared with the traditional linear regression model, machine learning considerably improves the predictability of REITs market, especially during the financial crisis when the market experiences extreme fluctuation. Furthermore, we identify that, for different market situations, distinctive predictive signals contribute to explaining the future REITs market movement. Lastly, we show that an investment strategy derived from machine learning algorithm overwhelmingly best the buy-and-hold strategy in the financial crisis and make striking profits. To conclude, this paper justifies the critical role of machine learning and deep learning in the asset pricing of REITs market and investment strategy.

# CHAPTER 1: AN ANALYSIS OF REIT CREDIT DEFAULT SWAP PRICING

## 1. Introduction

A credit default swap (CDS) is an insurance contract which allows a protection buyer to purchase insurance against a contingent credit event on an underlying reference entity by paying a quarter premium, generally referred to as the CDS spread, to the protection seller. The CDS market has grown from a niche venue to a great and active market over the past decade. Concurrently, an increasing amount of literature has shed light on CDS markets and profoundly studies CDS from many aspects, including the CDS valuation, the channel between CDS and related markets, as well as the impact of CDS on corporate finance.[1]

However, compared with CDS markets of other industries, the CDS market on real estate investment trust (REIT) is still fledging and innovative regarding market volume and numbers of contracts. Due to the 90% dividend rule, REITs have to keep a high level of debt for operation and utilize almost twice the leverage than industrial firms (on average).[2] The substantial leverage makes REITs vulnerable to a credit event, which highlights the importance of CDS market on REITs for providing REIT investors protection from default risk.

While there have been very few studies that analyze credit risk within the REIT industry. This study attempts to address this critical gap by examining determinants of CDS

---

[1] For detail related studies, see the survey of CDS market proposed by Augustin et al. (2014)

[2] Giacomini et al. (2015) document that from 1990-2012, REITs have an average market leverage of 46%. The average for industrial firms is 27%. For further comparison, the study reports that high debt capacity retail firms and high debt capacity firms in the top quartile of asset tangibility is 27% and 38%, respectively.

valuation in the REIT industry for the first time. Since default probability primarily determines CDS prices, studying these determinants provides new information concerning credit risk factors for REITs.

Inspired by Doshi et al. (2013), we first devise a discrete-time no-arbitrage model with observable factors for pricing REITs CDS. This model contributes to the structure model for valuing CDS spreads in respects of fitness and economic impact of factors on CDS valuation. Additionally, our study adds to the CDS literature at large, where there has been a debate regarding the usefulness of observable covariates for explaining credit spreads. We use panel regression analysis to examine the explanatory power of a long list of characteristics on valuing REITs CDS. More specifically, we investigate the role of Merton-implied factors, REITs-specific fundamentals, and macroeconomic shocks in determining CDS spreads of REITs. Finally, we document that real estate and firm-specific characteristics remain statistically significant even when controlling for structural default model factors suggested by Merton (1974) as well as an array of macroeconomic factors.

Then, we incorporate a dummy variable for a crucial measure of credit risk, the credit rating, into our panel regression analysis. Similar to Hull et al. (2004), we also find that the CDS spreads on REITs with lower credit rating are more sensitive to adverse shocks. Last, we consider the implementation of "Big Bang" protocol on 2009. This protocol aims to eliminate the ambiguity of CDS markets and make CDS contracts more fungible by standardizing and regulating the trading procedure of CDS. We execute an out-of-sample investment strategy based on the deviations between market quotes and the fundamental valuation and then document the performance of such portfolio in pre- and post- "Big Bang" protocol. The small excess returns after the enactment of protocol provide solid evidence that the “Big Bang” protocol significantly improves the efficiency of CDS markets and makes market quotes convert to the fundamental value.

The remainder of the paper is structured as follows. The next section reviews studies on CDS markets regarding the pricing model, empirical determinants, and the development of REIT CDS. Section 2 introduces the no-arbitrage model for pricing REIT CDS, including a quadratic term structure of interest rates and a quadratic intensity function with observable covariates. In Section 3, we present a case study of the largest (market capitalization) REIT, Simon Property Group, to demonstrate the features of our model. Section 4 describes the data sources and sample construction. In Section 5, we apply panel regression to analyze the performance of a long list of characteristics in explaining the cross-sectional variation of market CDS spreads on REITs. In Section 6, we provide an additional test of our model using an out of sample investment strategy. Section 7 offers concluding remarks.

## 2. Literature Reviews

### *The studies of default risk premium*

A substantial number of studies have shed light on the determinants of default risk premium. Colline-Dufresne et al. (2001) use monthly observable variables derived from theoretical structural models to fit the corporate bond credit spreads, which are the difference between the Treasury curve yield and the bond yield at the same maturity. They suggest that these theoretical variables have limited explanatory power and there is an unknown characteristic that determines the default risk premium. Inspired with the fast development of CDS market, lots of empirical works have focused on the CDS spread, instead of bond credit spreads, to proxy the default premium.

Although CDS spreads are theoretically similar to bond spreads, CDS spreads have some crucial advantages over corporate bond spreads (Ericsson et al. 2009; Han et al. 2017). Firstly, CDS spreads remove the interruption caused by a misspecification of a benchmark risk-free yield curve. Secondly, except for the default premium, a bond spread also contains non-default components, such as the liquidity risk and the market risk (Longstaff et al. 2005). Thirdly, CDS spreads are standardized and collected at a daily frequency, which makes the measure of default risk more accurate and updated.

The existing literature about determinants of default risk premium of REITs is sparse. To our best known, the only relative published study on this topic is Swanson et al. (2002). The authors investigate the relation between REIT portfolio risk premium, which is the spread between the return of REIT equity portfolio and treasury bill rate, and fundamental factors. However, similar with bond spreads, such REIT portfolio spreads are also noised by non-default components. In this study, we use CDS spread of REITs to proxy the default risk premium and investigate the determinants of such credit quality.

*The pricing of credit default swaps*

Recent literature has examined the pricing mechanism of credit default swaps (CDS) from versatile structural credit models and fundamentals of a firm.[3]

*The theoretical model*

Inspired by Black and Cox (1976) and Merton (1974), some pricing frameworks link the CDS spread to the interest rate, the financial leverage, and the business volatility.4

[3] Augustin et al. (2014) summarize a comprehensive review of the CDS literature.
[4] See also Longstaff and Schwartz (1995), Leland and Toft (1996), Collin-Dufresne and Goldstein (2001), Huang and Huang (2012).

However, the performance of such structural models and related observable factors on credit risk pricing are not in common. Huang and Zhou (2008) conduct a GMM-based specification test of five structural models to fit the changes of CDS spreads. They show that the structural models fail to predict CDS spreads. However, both Ericsson et al. (2009) and Bai and Wu (2016) show that the Merton-implied variables explain more than 50% of the variation in the CDS spread.

The reduced-form model, derived from Jarrow and Turnbull (1995) and Duffie (1999),[5] is another essential approach to value CDS spreads. While this method has been proven to be practical and successful, the hazard rate, following a random Poisson process, is a latent factor in this model.[6] The latent factor is barely related to the economic and firm-specific determinants of CDS spread.

To address the weakness of latent factors in reduced-form models, Doshi et al. (2013) introduce a discrete-time no-arbitrage model to determine CDS valuation. The hazard rate in the model is driven by a quadratic form of observable determinants. This quadratic no-arbitrage model not only rules out the negative model spread, which improves the flexibility of model but also links with firm-specific characteristics. In this paper, we follow Doshi et al. (2013) and add a REIT-specific covariate into this no-arbitrage model to fit the variation of CDS spread on REITs.

---

[5] Longstaff et al. (2005) develop closed-form expression within the reduced-form model and indicate that the default component accounts for the majority of the corporate spread across all credit ratings.
[6] For example, Chen et al. (2008) develop a closed-form solution of the reduced-form model to the CDS value, which performs well and improves the computational efficiency.

*The empirical study*

Some papers try to investigate the relation between CDS spreads and some observable determents.[7] At the market level, extant literature has proved the explanatory power of market return, term spread, risk-free rate, and many other macroeconomic factors in valuing CDS spread.[8] Among the firm-level fundamentals, financial leverage, return momentum and stock volatility are the most dominant factors. Considering that a CDS is comparable an out-of-the-money put, Cao et al. (2009) find that implied volatility dominates historical volatility in explaining the change of CDS spreads. Pan and Singleton (2008) derive the risk premium from sovereign CDS spreads and show that the VIX index is critical to this measure. Bharath and Shumway (2008) provide the contribution of Merton distance-to-default factor on pricing CDS. Bai and Wu (2016) combine the Merton-implied measure with a long list of firm fundamentals to value the CDS spread via a Bayesian shrinkage method. The average $R^2$ of 77% in their study indicates a substantially explanatory power of such firm-specific covariates. All the results highlight the significant role of macroeconomic conditions and firm fundamental in explaining CDS spread.

*The studies of CDS on REITs*

Compared with the studies of CDS in other sectors, the researches of CDS on REITs are sparse. To our best known, Bai and Zhu (2017) is the only paper that focuses on the CDS market of REITs. The authors document the importance of property types, geographic location, and economic factors in determining CDS spreads on REITs. Compared with their study, we mainly investigate the explanatory power of firm fundamentals, especially

---

[7] See, for example, Ericsson et al. (2009), Zhang et al. (2009), Cao et al. (2010) and Tang and Yan (2017).
[8] See for example Das et al. (2009) and Tang and Yan (2010)

REITs-specific covariates, in explaining the value of REITs' CDS. Furthermore, by collecting CDS data from three major databases, we contain a more substantial CDS data.

To better fit the spreads of REITs CDS, we also examine some REIT-specific variables from previous literature. Harrison et al. (2011) discuss the determinants of REITs' debt capacity and capital structure decisions, which is also related to the credit risk of REITs. They find that REITs size, interest coverage, property type, umbrella partnership REITs (UPREITs), funds from operations (FFO) payout ratio, historical REITs return, and profitability jointly decide the debt capacity of REITs. This finding is also consistent with the evidence in Brown and Riddiough (2003), Maris and Elayan (1990), and Boudry et al. (2010). We will examine whether such REITs-specific factors contribute to the value of CDS of REITs.

## 3. Model Description

In this section, inspired by Doshi et al. (2013), we develop both a risk-free rate model and a pricing model for CDSs with discrete-time version. Assuming that all factors follow Gaussian VAR processes, we derive a closed-form resolution of both structure models.

### *Risk-free Rate Model*

We adopt a quadratic term structure model (QTSM) to characterize the process of spot interest rate.[9] Compared with affine term structure models, quadratic models remove

[9] Ahn et al. (2002) and Leippold and Wu (2002) develop the continuous-time quadratic models of interest rate, and Realdon (2006) derive the discrete-time quadratic model of interest rate.

the unpractically negative interest rates and also provide better in-sample fit (Ahn et al. 2002). We assume that the quadratic function of interest rate is:

(1) $$r_t = c + b'X_t^r + X_t^{r\prime}AX_t^r,$$

where $X_t^r$ is a vector of N state variables, c is a constant, b is an $N \times 1$ vector, and A is an $N \times N$ matrix. Furthermore, the vector of state variables $X_t^r$ is assumed to follow a first order Gaussian VAR process under the physical measure:

(2) $$X_{t+1}^r = \mu + \Phi X_t^r + \Sigma\eta_{t+1},$$

where $\eta_{t+1}$ is an $N \times 1$ vector of $i.i.d. N(0, I)$ erros, $\Phi$ is an $N \times N$ autoregressive matrix, $\mu$ is an $N \times 1$ vector, and $\Sigma$ is an $N \times N$ matrix.

As any discounted price process under the risk-neutral measure is martingale, we will convert the physical measure to the risk-neutral measure. In the spirit of Dai, Le, and Singleton (2010), we assume time-varying market price of risk that is a linear function of the state variables:

(3) $$\Lambda_t = \lambda_0 + \lambda_1 X_t^r,$$

where $\lambda_0$ is an $N \times 1$ vector and $\lambda_1$ is $N \times N$ matrix. Then$X_t^r$, under the risk-neutral measure $Q$, is described as:

(4) $$X_{t+1}^r = \mu^Q(X_t^r) + \Sigma\xi_{t+1},$$

where $\xi_{t+1} \sim N(0, I)$ and $\mu^Q(X_t)$ is given by:

(5) $$\mu^Q(X_t) = (\Phi - \Sigma\lambda_1)X_t^r + \mu - \Sigma\lambda_0.$$

Finally, the observed zero-coupon bond yield $Y(X_t^r, \tau)$, with maturity $\tau$ at time $t$, is written in the form of the state variables:

(6) $$Y(X_t^r, \tau) = -\frac{c_\tau}{\tau} - \frac{b_\tau'}{\tau}X_t^r - X_t^{r\prime}\frac{A_\tau}{\tau}X_t^r + \epsilon_{t,\tau}$$

where $\epsilon_{t,\tau} \sim N(0, R)$ and $R$ is assumed to be a diagonal matrix, and the explicit expression of the coefficients $c_\tau$, $b_\tau$, and $A_\tau$ are derived recursively.[10]

*Credit default swap valuation*

In the reduced-form model of CDS, the default process is modeled directly by modeling the probability of default itself. Jarrow and Turnbull (1995) describe the credit default as the first event of a Poisson process and assume that an intensity process $\lambda(t)$, known as the hazard rate, decides the default probability. The discrete time survival probability to time T is given by

(7) $\Pr[\tau \geq T] = E_t\left[exp\left(-\sum_{j=0}^{T-1} \lambda_{t+j}\right)\right],$

where $\tau$ indicates the time of default.

Following the quadratic framework of Doshi et al. (2013), we assume that the hazard rate depends on a quadratic form of the default-free rate and a set of observable variables:

(8) $\lambda_t = \left(\alpha_0 + \sum_{k=1}^{n} \alpha_k^r X_{k,t}^r + \sum_{k=1}^{m} \alpha_k^d X_{k,t}^d\right)^2,$

where $n$ is the number of state factors forming the risk-free rate model and $m$ is the number of observable covariates. Stacking $X_t^r$ and $X_t^d$ in $X_t$ a $q \times 1$ vector, where $q = n + m$, we can write

(9) $r_t + \lambda_t = \gamma_0 + \gamma_1' X_t + X_t' \Omega X_t.$

$X_t$ is also specified by autoregressive process under the objective measure:

(10) $X_t = \mu + \rho X_{t-1} + \Sigma e_t,$

where $e_t \sim N(0, I)$, $\mu$ is a $(q, 1)$ vector, and $\rho$ and $\Sigma$ are $(q, q)$ diagonal matirces.

[10] For more details on recursive process, see Realdon (2006) and Doshi et al. (2013)

We first model the value of the premium leg which is paid by CDS buyers to purchase credit protection from a credit event. The premium leg is series of a quarterly payment to maturity or to the time of a credit event. We assume that credit events only happen at the end of each quarter. The present value of the premium leg is

(11) $PB_t = E_t\left[S\Delta \sum_j^h 1_{(\tau>1+j)} A(t,t+j)\right]$

where $S$ is the CDS spread, $S\Delta$ is the amount of each quarterly payment, 1 denotes the indicator function, and $A(t,t+j)$ is the riskless discount rate, which is derived from risk-free structure model. Doshi et al. (2013) display that

(12) $E_t\left[1_{(\tau>1+j)} A(t+j)\right] = \exp\left(F_j + G_j' X_t + X_t' H_j X_t\right),$

where the coefficients $F_j$, $G_j'$, and $H_j$ are derived recursively.

Then we consider the value of protection seller. If a credit event happens, the protection seller makes a payment to the buyer, known as the protection leg. The protection leg is the contingent payment of $(1-R)$, where $R$ is the recovery rate. The present value of the protection leg is

(13) $PS_t = E_t\left[(1-R)\sum_{j=1}^h 1_{(t+j-1<\tau\leq t+j)} A(t+j)\right].$

Assuming that the recovery rate is fixed and predetermined, we then have

(14) $PS_t = (1-R)\left(E_t\left[\sum_{j=1}^h 1_{(\tau>t+j-1)} A(t+j)\right] - E_t\left[\sum_{j=1}^h 1_{(\tau>t+j)} A(t+j)\right]\right),$

where the two parts of right side are in form of the Eq. $(12)$. Finally, we calculate the breakeven default swap spread, which is given by

(15) $PB_t = PS_t.$

*Estimation method of model*

To estimate the CDS spread, we first simulate the QTSM of risk-free rate model. Following Doshi et al. (2013), we cast the model of risk-free rate into a state-space system and apply the unscented Kalman filter (UKF). The nonlinear state-space system is defined as:

(16) $X_t^r = \mu_r + \rho_r X_{t-1}^r + \Sigma_r e_t,$

(17) $Y_t = Z(X_t) + \mu_t,$

where $Y_t$ is a matrix containing observable yields of zero-coupon treasury bond and $\mu_t$ and $e_t$ are diagonal matrix. As the measurement Eq. (17) is nonlinear in the state variables, UKF improves the calculation efficiency by directly estimating the posterior density via a bunch of predetermined sample points. [11]

From the UKF, we acquire a set of efficient estimations on the conditional mean $\overline{Y_t}$ and conditional covariance matrix $F_t$ of Treasury bond yields, and then build the log-likelihood function to derive the state factors in our risk-free rate model

(18) $max\ \ln L = \sum_{t=1}^{T}\left[-\frac{1}{2}\log|F_t| - \frac{1}{2}(Y_t - \overline{Y_t})' F_t^{-1}(Y_t - \overline{Y_t})\right]$

where $T$ is the sample size.

After collecting the estimated state variables from the risk-free rate model, we combine them with other observable factors to price CDS spreads. We set $\rho$ and $\Sigma$ in Eq. (16) as diagonal matrices, so all parameters in Eq. (16) are estimated by an AR(1) regression. To change from the physical measure to the risk-neutral measure, we follow the process of Eq. (3). Under the risk-neutral probability, the remaining parameters are

[11] For a detailed structure of unscented Kalman filter, see Julier and Uhlmann (1997) and Merwe and Wan (2001).

filtered by minimizing the root-mean-square error (RMSE) based on the term structure of CDS spreads.

## 4. A Case Study: Simon Property Group (SPG)

To demonstrate the performance and implication of our pricing model for CDS spreads, we show a detailed study of a single REIT: Simon Property Group (SPG), the largest REIT, the CDS data of which is most available and reliable.

We first set the size of vector $X_t^r$ in Eq.$(1)$ and then extract these two latent factors from the risk-free term structure. According to the canonical formulation of quadratic model proposed by Ahn et al. (2002) and Realdon (2006), we impose restrictions on the parameters in the term structure model to facilitate calculation. $\Phi$, $\Sigma$ in Eq. $(2)$ and $\lambda_1$ in Eq. $(3)$ are diagonal matrices. From Eq. $(1)$, b is equal to zero and A is an identity matrix.

Panel A in Table $(1)$ displays parameter estimates of the QTSM of risk-free rate. Our estimation is consistent with the result in Nyholm and Vidova-Koleva (2012). Furthermore, the RMSE and measure error standard deviations (ME SD) reported in Panel B further confirm the decent fitness of the QTSM. Similar to Doshi et al. (2013) and Chen et al. (2008), we also find that the two-factor QTSM works better for mid-term bond yield than for the short and long bond yield.

We now simulate the dynamics of CDS spreads of SPG via five factors: two state covariates from the interest rate model, leverage, one-year historical volatility, and FFO payout. To make CDS spreads of SPG conform normality distribution, we take a logarithm of CDS data. We also force the similar restrictions on the parameters of CDS valuation model as we impose on the default-free rate model. Panel C in Table (1) reports the loadings of parameters to estimate the CDS spreads of SPG.

In general, the coefficients of all five factors forming the hazard rate are significant, and the signs of observable factors accord with the theoretical implication. Our CDS pricing model indicates that macroeconomic, Merton-implied, and REITs-fundamental factors jointly value the CDS spread of REITs.

## 5. Data Description

We collect the data of CDS and REITs from several sources. The financial statement information of REITs comes from SNL and Capital IQ and market information is retrieved from Bloomberg. As the payoff of CDS is similar to that of a deep out-of-money put option, we collect the one-year 25-delta put option implied volatility from OptionMetrics. Consistent with Bai and Wu (2016) and Doshi et al. (2013), we assume the balance sheet data is available forty-five days after the end of the fiscal quarter. According to Zhang et al. (2009), we adopt both long-term volatility and short-term volatility. Specifically, the long-term volatility is a one-year historical volatility of daily REITs returns, and the short-term volatility is an annualized two-week volatility. Moreover, we follow KMV model and use current liability plus half of long-term liability as the proxy of the debt level to calculate the 1-year distance-to-default (DD).

Collecting the Standard & Poor’s credit rating from SNL, we map the score on a numerical grade scale and set that scales less than 5 are the investment grade and others are the speculative grade. To capture the influence of property type on REITs CDS, we subgroup all REITs into seven categories based on the information provided by SNL: specialty, multifamily, office, shopping center, industrial, healthcare, and hotel. We also sort all REITs into four geographic sections based on the state location of REITs headquarter: Midwest, Northeast, South, and West.

To estimate the risk-free term structure model, we use U.S Treasury zero-coupon yield curve data from the Federal Reserve Bank, which consists of yearly yield observations for maturities of 0.5, 1, 3, 5 and 10 years.[12] Additionally, we also obtain a wide variety of macroeconomic variables from Datastream, including the prime rate, Moody's corporate bond yield, treasury bond yield, and the consumer price index (CPI). Finally, the Table (2) summarizes all variables in this paper for potentially explaining the variation of CDS spreads of REITs.

As CDS contracts are traded over the counter (OTC), there is no standardized and unified database for CDS. Mayordomo et al. (2010) compare the five major databases of corporate CDS prices and show that the CMA, compared with other databases, lead the price discovery process. Therefore, we collect data of REITs CDSs mainly from the CMA database. The CMA CDS data is provided by DataStream before October 2010 and then is offered through Bloomberg and SNL. To construct a comprehensive and accurate data set of REITs CDS, we map the weekly 5-year tenor CDS data into DataStream, Bloomberg, and SNL between Jan 2005 to Sep 2016. Finally, we collect 44 REITs with available CDS information, dead or existing, which are at least contained in two of these the three databases.

As the trading market of REITs CDS is inactive and fledging, in terms of trading volume and numbers of production, we find that our CDS data is contingently

[12] These data are also used in Diebold and Li (2006) and are estimated daily using a cubic spline model.

contaminated with outliers and stale observations. To minimize measurement errors and remove skeptical CDS quotes, we filter our CDS data as the following rules:

- For each REIT, we only include the CDS data since the beginning of REIT status.
- We delete the CDS data within each quarter if more than two third of CDS quotes of a specific REIT are constant, which implies that the market for such CDS contracts is extremely inactive and the CDS quotes may not accurately reflect market information.
- If the CDS spread suddenly jumps or plummets more than 100% in next day, we will remove previous four-weeks CDS quotes, because such quotes may be inefficient.
- We rule out CDS spreads above 5,000 basis points, because these contracts often involve bilateral arrangements for upfront payments.
- Finally, we winsorize the CDS spreads at level 1% and 99% for each REIT

Conclusively, we identify 44 REITs of which the historical CDS data is available and collect 18405 week-CDS observations from January 2005 to September 2016. The Appendix reports the basic information for all REITs in our data set.

## 6. Summary statistics

We provide summary statistics of REITs fundamental characteristics in Table (3). Pooling all week observations, we first summarize the mean, standard deviation, and correlation with CDS spread (logarithm form) of each characteristic. The average credit rating of our REITs sample is moderate, with BBB rating which is the breakpoint of investment grade. We find that all the signs of each variable in correlation column are consistent with the implication of the financial theory. Furthermore, we depict the time-series average of CDS spreads of REITs with different property types in Figure (1). The average spread of multifamily REITs is lowest and average spreads of specialty, hotel, and industrial REITs are the three largest. Due to the distinctive business operation and correlation with the economy, REITs in different industries face distinct level of credit risk which is partly measure by the CDS spreads. The CDS spreads of all property types jump suddenly and extremely during the financial crisis and then fall into a low level.

Additionally, we categorize weekly CDS spreads into quintiles and then show statistics of each covariate in the first and last quintiles in Table (3). In general, all firm-specific covariates change significantly from the first quintile to the fifth quintile. The fact that the spread level of the fourth quintile[13] is even lower than the average level indicates positive skewness in cross-section CDS data. To restrict the distribution of CDS to normality, we take a natural logarithm on CDS spreads for the following analysis. Both market leverage and volatility monotonically increase with CDS quintiles, and a similar process for implied volatility, short sale, and DD measure. All these movements conform to the existent literature. Figure (2) demonstrates the varies of CDS spread and other three covariates in the top and bottom quintile. All characteristics of riskiest REITs (5th quintile) are much more volatile than that of safest REITs (1st quintile), which implies that the explanatory power of determents on CDS valuation is variable for REITs with different credit quality.

However, Checking the correlation columns of each quintile, we find signs of some variables are counterintuitive or insignificant. We believe three reasons contribute to this suspicion. The CDS market is inefficient and opaque before the implementation of the "Big

---

[13] To save table space, we do not present statistics information in all quintiles

Bang" protocol which standardizes the CDS contract and regulates CDS transactions. Additionally, for the small-cap company, credit rating agencies always ask for a stricter criterion to keep the same level of rating, which will disrupt the rational relation between credit measure and fundamental performance. The smaller REITs size in first quintile than in that second quintile suggests that the first quintile includes small REITs with high credit rating. Moreover, the financial crisis may cause such ambiguity, because CDS spreads, especially for small-cap REITs, experience extreme volatility from 2007 to 2010.

To examine the performance of firm fundamentals on fitting the variation of default risk, some previous studies resort to time-series regression (Collin-Dufresne et al. 2012; Ericsson et al. 2009), while others focus on the cross-sectional relation (Bai and Wu 2016; Cao et al. 2010). To exam the fluctuations of fundamentals over time and cross different REITs, we follow Bai and Wu (2016) and estimate two sets of standard deviation. The first measure, named as XS, is the time-series mean of cross-sectional standard deviation at each date. The second test, labeled as TS, is the cross-sectional average of time-series standard deviation for each REIT.

The XS evaluates the volatility of each characteristic across REITs panel whereas the TS assesses the variation of each factor over time for a given REITs. Shown in Table (3), the cross-sectional deviation of most of REIT fundamentals, besides REIT returns, is larger than or close to the time-series deviation. While, the differences are not significant, which is different from the finding in Bai and Wu (2016). Since our study focusing on a single industry, the considerable small cross-sectional standard deviation is understandable. Besides, the financial chaos exaggerates the volatility of whole REITs market and fundamentals of all REITs vary uniformly. However, we still apply panel data regression to investigate the explanatory power of REIT characteristics, because the quarterly-updated financial statements constrain the changes of fundamentals over a short sample period.

Compared with Cao et al. (2010), our market data are more volatile, which is caused by financial crisis and subsequent monetary policy. Generally, the signs of correlation imply that the more prosperous market and economy are, the lower CDS value is.

## 7. Empirical Results

To examine how much the cross-sectional CDS value of REITs can be determined by the list of characters, we propose three panel regressions based on Merton-model variables, REITs-specific fundamentals, and a comprehensive set of factors. The general regression model is expressed as:

(19) $$\mathrm{CDS}_{i,t} = c + b_1 Merton\ Implied_{i,t} + b_2 REITs_{i,t} + b_3 Macro_{i,t} + \varepsilon_{i,t},$$

where the explanatory variables are detailed in Table $(2)$.[14] To adjust for potential bias in each panel regression, we follow Petersen (2009) to examine robustness of P-value by using clustered standard errors.

Inspired by Merton (1974) and Black and Cox (1976), we first test whether CDS spreads are determined by interest rates, leverage of REITs, and REITs volatility. Although the short-term volatility and long-term volatility are highly correlated, the short-term will timely measure an unexpected deviation, which is critical during the financial crisis. Our panel regression includes both long-term volatility and short-term volatility.

Table $(4)$ reports the results of Merton-implied factors model with and without fixed-effect. The high $R^2$ in both regression, 67% and 73%, suggests that the three Merton-

[14] Note that the constant term, $c$, only exists in pooled regression model and will be removed in fixed-effect model to avoid collinearity.

implied measures explain a significant portion of spread variation of REITs CDS. Moreover, the significantly positive coefficient for leverage indicates that CDS spreads efficiently reflect the default risk of REIT, which is proved by the existent literature. Conforming with the conclusion in Zhang et al. (2009), the statistical significance of two volatility measures confirms that a combination of long- and short-term volatilities can better delineate the vibration of REITs CDS. Furthermore, the magnitude of short-term volatilities coefficient is significantly larger than that of long-term volatility coefficient in both regressions, which exhibits the importance of short-term volatilities in measuring credit risk. The significantly negative slope of interest rate is consistent with Longstaff and Schwartz (1995). Moreover, as most of the rent contacts are floating-rate adjust, the higher interest rate will increase operating cash-flow of REITs, which enhances the interest coverage. In general, our findings of Merton-implied model also agree with Doshi et al. (2013) and Wu and Bai (2016).

Inspired by the implication of our CDS value model, we further investigate the explanatory power of REITs fundamentals, such as real estate value, profitability, interest converges, and REITs market performance, on depicting the variation of CDS values. We show the result in Table (5). Similar to the R-squared of the Merton-implied model, the high $R^2$ of both pooled and fixed-effect regression model convincingly reveals that REITs characteristics also play a pivotal role in predicting the default-risk of REITs which is measured by the level of CDS spread. Consistent the previous default-risk studies, we find that a small-size REIT with a high leverage ratio and low-profit margin will be vulnerable to default, and a stagnant REITs market will further deteriorate the probability of default.

In Table (5), the notably improved R-squared from pooled regression to fixed-effect regression implies that some individual-specific immobile covariates that efficiently price REITs CDS. Since the operation-risk is considerably distinct for REITs exposure to different property type and location, we create dummy variables for these two immobile REITs specifications to examine the linkage between REITs CDS and them. After considering the effect of property type and location, the REIT-specific pooled model can explain additional 3% of the variation of CDS spreads ($R^2$ increases to 74%). Moreover, the R-squared of pooled regression including individual-specific dummy (74%) is close to that of fixed-effect regression (75%), which denotes that property type and spatial distribution are the two most principal fixed determinants on valuing REITs CDS.

By comparing the magnitudes of significant coefficients of property dummies, we rank the probability of default from highest to lowest, as the hotel, specialty (mainly focusing on communication sites), office, industry, healthcare, and multi-family. With more sensitive to macro-economy and high vacancy rate, hotel REITs are more likely to face financial stress (Kim et al. 2002). As SPG, the largest REIT, is part of the shopping center property, the CDS spread distribution of shopping center is left-skewed. After excluding SPG from the regression, we notice that the credit risk of shopping center is close to that of healthcare.[15]

The significant coefficients of location dummy in Table (5) further demonstrate the findings of Bai and Zhu (2017) that there exists the effects of geographic concentration and local economic conditions on CDS spreads on REITs. Compared with REITs in other areas, REITs in Southern USA, concentrating in Florida state, face the most substantial default probability. During the financial crisis, the S&P Case-Shiller home price indices of Miami, where is the principal real estate market in Florida, drop in half.[16] Moreover, the

[15] To save the space of table, we do not report the result of regression without SPG data in table.
[16] http://us.spindices.com/index-family/real-estate/sp-corelogic-case-shiller

flourishing local economy in the northeastern and western USA protects regional REITs from bankruptcy.

Last but not the least, we include a comprehensive list of variables and run a kitchen-sink model. As Table (6) shows, the highest R-squared (86%) for a kitchen-sink model with fixed-effect signifies that our empirical model does an excellent work of valuing REITs CDS. Similar to Cao et al. (2010) and Doshi et al. (2013), we observe that implied volatility dominates historical volatility in explaining the CDS spreads. After adding implied volatility, the coefficients of both long- and short-term volatility become insignificant or counterintuitive. Furthermore, the KMV DD measure significantly forecast the value of REITs CDS, which agrees with the finding of Bharath and Shumway (2008) and Bai and Wu (2016).

Our summary of the performance among the market covariates is consistent with Cao et al. (2010). The significantly negative coefficients of interest rate level and slop match with the evidence proposed by Duffee (1998). A low level of interest rate and yield curve always indicates a worse market and economy, which will deteriorate the solvency of REITs. The coefficients of all three measures of market sensitive, the market credit risk, the market liquidity risk, and the VIX index, are positive and significant. Broadly speaking, when the economy is sluggish, so is real estate.

*Extended regression*

The negative coefficients of the credit-rating dummy in previous regressions have confirmed that the CDS spreads reflect the information of credit-rating of REITs. In this section, we will investigate the varied influences of Merton-implied and REITs-specific covariates on describing the CDS spreads on REITs with different credit rating. Hull et al.

(2004) conclude that adverse rating events are much more significantly related to credit spread changes than positive rating events. We add interaction terms between credit-rating dummy and covariates in Merton-implied model and REIT-specific model separately,

(20) $\text{CDS}_{i,t} = c + b_1 Merton\ Implied_{i,t} + b_2 Merton\ Implied_{i,t} \times\ Rating_{i,t} + \varepsilon_{i,t},$

(21) $\text{CDS}_{i,t} = c + b_1 REITs_{i,t} + b_2 REITs_{i,t} \times\ Rating_{i,t} + \varepsilon_{i,t},$

where *Rating* is a dummy viable which is defined as $Rating = 1$ for REITs with investment grade rating and $Rating = 0$ for REITs with speculative grade rating.

Table (7) shows the results of panel regression with interaction items. Panel A presents the result of regression with interaction terms between credit rating and Merton-implied variables. Except for the interaction term of long-term volatility, all others are significant, which implies that the channels of pricing CDS for REITs with different credit-rating are distinct. Explicitly, the coefficients of interaction terms ($Leverage_{i,t} \times Rating_{i,t}$ and $Risk - free\ rate \times\ Rating_{i,t}$) are negative and the coefficient of interaction term between credit rating and short-term volatility is positive. The estimates demonstrate that the magnitudes of leverage and risk-free rate on CDS valuation of REITs with speculative-grade are more significant than that for REITs with investment-grade, but CDS spreads of REITs with investment-grade are more sensitive to changes in short-term volatility, rather than to variations in long-term fluctuations.

Panel B illustrates the results of regression with interaction terms between credit rating and REITs-specific fundaments. We also uncover an asymmetric relation between explanatory power on valuing CDS and REITs with different credit rating. The significances of interaction items suggest that fundamentals of REITs with investment-grade contribute much more than that of REITs with speculative-grade to changes of CDS spreads. The significantly negative coefficients of interaction items ($Profitbility_{i,t} \times Rating_{i,t}$ and $SNL\ Return_t \times\ Rating_{i,t}$) suggest that positive information of REITs with investment-grade offers valuable information in pricing CDS spreads. The negative of interaction item between credit rating and volatility of SNL indicates that REITs with speculative-grade are more vulnerable than REITs with superior credit quality to the chaos of market, which further explains the extreme difference among CDS spreads of REITs during the financial crisis.

*The explanation of coefficient*

To better understand the economic meaning of coefficients in previous regression models, we take the coefficient of hotel dummy, which is 0.69 in Table (5), as an example to show the relation between the magnitude of coefficient and probability of default. The usual assumption is that the CDS spreads only measure the possibility of default because the variations of CDS spreads directly reflect the changes in default probability.

According to the pricing model of REITs CDS which we explain in the Model Section, the probability of default is calculated by

(22) $PD=1-e^{-\lambda}$

where *PD* is the probability of default and $\lambda$ is the hazard rate. We follow Hull and Basu (2016) to calculate the hazard rate as CDS spreads to $(1-\text{Recovery Rate})$ and assume that the Recovery Rate is 40%. Suppose that the CDS spread of Multifamily REIT is 100 basis points, its default possibility is 1.65%. Given all fundaments of this Multifamily REIT is same to that of a Hotel REIT, the coefficient of hotel dummy (0.69) unveils the probability of default of this Hotel REIT is 3.25%. In general, the default risk of a hotel REIT is almost as two times as that of a multifamily REIT given all other determinants are equal.

## 8. Out-of-sample investment strategy

Inspired by Doshi et al (2013) and Wu and Bai (2016), we depict an out-of-sample investment exercise based on the difference between market observations of CDS quotes and the fundamental-regression valuations of REITs CDS.

At the beginning of each month, we first filter the REITs data with available previous 26-weeks CDS spreads. Then, we recursively estimate our kitchen-sink regression model with fixed effect and update the regression by one month at a time. Intuitively, the market CDS quotes will finally convert to the fundamental valuation derived from REITs characteristics. We establish a zero-cost investment portfolio by long the "cheap" CDS, the market value of which is lower than fundamental-implied valuation, and short the "expensive" CDS, the market observation of which exceeds the regression-expected value. The weight of each REITs in the investment portfolio is proportional to the size of the deviation between market value and model-implied spreads, which means the more extreme divergences are, the larger weights in the portfolio are. Finally, the size of short is equal to that of long and is normalized to one. The model of weights of the "cheap" REITs CDS in the portfolio is following:

$$(23) \qquad Not_{t,i} \propto \frac{\left(S_{t,i}^{Regression} - S_{t,i}^{Market}\right)}{S_{t,i}^{Market}},$$

$$(24) \qquad 1 = \sum_i Not_{t,i},$$

where $S_{t,i}^{Regression}$ is the model-implied value, $S_{t,i}^{Market}$ is the market price of CDS, and $Not_{t,i}$ is the proportional weight of each REITs CDS.

Panel A in Table (8) presents the statistics of the excess returns of this investment strategy for horizons from one to four weeks in full data period. Similar to the previous research, we also document positive returns of this investment portfolio up to four weeks. Besides, the excess profit is increasing with more extended investment period, which is also presented by Doshi et al. (2013).

However, as CDS is an OTC contract, we do not consider our investment portfolio as a practical strategy. First, the market values of CDS in our data is not executable quotes, and the ask-bid spreads of CDS contracts may be sizeable. Second, as an OTC contract, the transaction cost of CDS is mainly dependent on the counterparty risk of each party. In consequence, some institutions can benefit from our investment strategy, whereas others may not.

*The influence of the "Big Bang" protocol on REITs CDS*

The "Big Bang" protocol, implemented in May 2009, aims to make the CDS market more consistent and efficient by regulating the procedure for trading CDS contracts. By standardizing CDS contracts and establishing the central clearing counterparty (CCP), the "Big Bang" protocol is expected to improve the efficiency of CDS market and decrease market opaque.

In theory, the inefficiency is the crucial factor in the profitability of our investment strategy. Extensive profits imply that the CDS market is inefficient. To analysis the influence of the "Big Bang" protocol on REITs CDS market, we split our data sample into pre-Big Bang (2005-2009) and Big Bang period (2010-2016), and then examine the performances of out-of-sample investment strategy respectively.

We document the performance of investment portfolio during pre-Big Bang and Big Bang periods in Panel B and Panel C of Table (8). The returns up to four weeks in Big Bang period are significantly less than that in pre-Big Bang, which agrees with the indication that the “Big Bang” protocol facilitates the CDS trading and makes CDS market transparent. Furthermore, the small excess profits in Big Bang periods suggest that the “Big Bang” protocol limits the CDS market prices to the fundamental-implied valuation and diminishes the market friction in OTC market. Besides, the volatility of investment profit almost cuts in half after the enactment of the “Big Bang” protocol, which highlights that the implementation significantly stabilizes the CDS market of REITs.

## 9. Conclusion

The study in CDS market is still evolving, especially for the CDS of REITs. In this paper, we first follow Doshi et al. (2013) to develop a quadratic structural model with observable factors to value the CDS of REITs. Our model endows the reduced-form approach for CDS valuation with the economic characteristics of the REIT underlying the CDS contract and keeps the efficiency and practicability of the reduced-form structure.

Then, we empirically analyze the impact of REIT-fundamental factors and economic covariates on deciding the valuations of REITs CDS. We first examine the explanatory power of three elements derived from Merton (1974) and unveil that such factors can notably explain the changes of CDS spreads on REITs. Subsequently, we perform a cross-sectional analysis of REITs-specific variables on pricing CDS and still conclude that the explanatory power of such variables is considerable. Finally, we consider a long list of characteristics to value CDS spreads on REITs effectively. The large R-squares of regression emphasizes that REITs-specific and economic information can significantly depict the variation of REITs CDS.

Moreover, we also reveal that the effects of characteristics on pricing the CDS of REITs with distinctive credit-rating are uncommon. The CDS spreads on REITs with lower credit rating are more sensitive to adverse fundamentals. Last, we provide evidence about the impact of the "Big Bang" protocol on CDS market. The small excess return of our out-of-sample investment signifies that the "Big Bang" protocol improves the efficiency of CDS market and moderates the market ambiguity.

Our study already discloses that the REITs returns contribute to the valuation of REITs CDS. For further research, the information channel among the REITs CDS market, the REITs bond market, and the REITs equity market will be an interesting topic and one of the most significant concerns on REITs CDS.

# REFERENCES

Ahn, D.H., Dittmar, R.F. and Gallant, A.R., 2002. Quadratic term structure models: Theory and evidence. *The Review of financial Studies* 15: 243-288.

Augustin, P., Subrahmanyam, M.G., Tang, D.Y. and Wang, S.Q., 2014. Credit default swaps: A survey. *Foundations and Trends in Finance* 9: 1-196.

Bai, J. and Wu, L., 2016. Anchoring credit default swap spreads to firm fundamentals. *Journal of Financial and Quantitative Analysis* 51: 1521-1543.

Bai, Q. and Zhu, L., 2017. The Effects of Industry Specific and Local Economic Factors on Credit Default Swap Spreads: Evidence from REITs. *Journal of Financial Services Research*: 1-29.

Bharath, S.T. and Shumway, T., 2008. Forecasting default with the Merton distance to default model. *The Review of Financial Studies* 21: 1339-1369.

Black, F. and Cox, J.C., 1976. Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance* 31: 351-367.

Brown, D.T. and Riddiough, T.J., 2003. Financing choice and liability structure of real estate investment trusts. *Real Estate Economics* 31: 313-346.

Boudry, W.I., Kallberg, J.G. and Liu, C.H., 2010. An analysis of REIT security issuance decisions. *Real Estate Economics* 38: 91-120.

Cao, C., Yu, F. and Zhong, Z., 2010. The information content of option-implied volatility for credit default swap valuation. *Journal of financial markets* 13: 321-343.

Chen, R.R., Cheng, X., Fabozzi, F.J. and Liu, B., 2008. An explicit, multi-factor credit default swap pricing model with correlated factors. *Journal of Financial and Quantitative Analysis* 43: 123-160.

Collin-Dufresne, P. and Goldstein, R.S., 2001. Do credit spreads reflect stationary leverage ratios? *The journal of finance* 56: 1929-1957.

Collin-Dufresn, P., Goldstein, R.S. and Martin, J.S., 2001. The determinants of credit spread changes. *The Journal of Finance* 56: 2177-2207.

Doshi, H., Ericsson, J., Jacobs, K. and Turnbull, S.M., 2013. Pricing credit default swaps with observable covariates. *The Review of Financial Studies* 26: 2049-2094.

Duffie, D., 1999. Credit swap valuation. *Financial Analysts Journal* 55: 73-87.

Ericsson, J., Jacobs, K. and Oviedo, R., 2009. The determinants of credit default swap premia. *Journal of Financial and Quantitative Analysis* 44: 109-132.

Giacomini, E., Ling, D.C. and Naranjo, A., 2015. Optimal Capital Structure and the Effects of Deviations from Target Leverage on REIT Return Performance. *Real Estate Research Institute Journal*.

Han, B., Subrahmanyam, A. and Zhou, Y., 2017. The term structure of credit spreads, firm fundamentals, and expected stock returns. *Journal of Financial Economics* 124: 147-171.

Harrison, D.M., Panasian, C.A. and Seiler, M.J., 2011. Further evidence on the capital structure of REITs. *Real Estate Economics* 39: 133-166.

Huang, J.Z. and Zhou, H., 2008. Specification analysis of structural credit risk models. Working paper, The Penn State University.

Huang, J.Z. and Huang, M., 2012. How much of the corporate-treasury yield spread is due to credit risk? *The Review of Asset Pricing Studies* 2: 153-202.

Hull, J.C. and Basu, S., 2016. *Options, futures, and other derivatives*. Pearson Education India.

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

Jarrow, R.A. and Turnbull, S.M., 1995. Pricing derivatives on financial securities subject to credit risk. *The journal of finance* 50: 53-85.

Julier, S.J. and Uhlmann, J.K., 1997, April. A new extension of the Kalman filter to nonlinear systems. In *Int. symp. aerospace/defense sensing, simul. and controls* 3: 182-193).

Kim, H., Gu, Z. and Mattila, A.S., 2002. Hotel real estate investment trusts' risk features and beta determinants. *Journal of Hospitality & Tourism Research* 26: 138-154.

Leippold, M. and Wu, L., 2002. Asset pricing under the quadratic class. *Journal of Financial and Quantitative Analysis* 37: 271-295.

Leland, H.E. and Toft, K.B., 1996. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance* 51: 987-1019.

Longstaff, F.A. and Schwartz, E.S., 1995. A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance* 50: 789-819.

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

Marts, B.A. and Elayan, F.A., 1990. Capital structure and the cost of capital for untaxed firms: the case of REITs. *Real Estate Economics* 18: 22-39.
Mayordomo, S., Peña, J.I. and Schwartz, E.S., 2014. Are all credit default swap databases equal? *European Financial Management* 20: 677-713.

Merton, R.C., 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance* 29: 449-470.

Nyholm, K. and Vidova-Koleva, R., 2012. Nelson–Siegel, Affine and Quadratic Yield Curve Specifications: Which One is Better at Forecasting? *Journal of Forecasting* 31: 540-564.

Pan, J. and Singleton, K.J., 2008. Default and recovery implicit in the term structure of sovereign CDS spreads. *The Journal of Finance* 63: 2345-2384.

Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies* 22: 435-480.

Realdon, M., 2006. Quadratic term structure models in discrete time. *Finance Research Letters* 3: 277-289.

Swanson, Z., Theis, J. and Casey, K.M., 2002. REIT risk premium sensitivity and interest rates. *The Journal of Real Estate Finance and Economics* 24: 319-330.

Tang, D.Y. and Yan, H., 2010. Market conditions, default risk and credit spreads. *Journal of Banking & Finance* 34: 743-753.

Tang, D.Y. and Yan, H., 2017. Understanding transactions prices in the credit default swaps market. *Journal of Financial Markets* 32: 1-27.

Van Der Merwe, R. and Wan, E.A., 2001. The square-root unscented Kalman filter for state and parameter-estimation. In *Acoustics, Speech, and Signal Processing, 2001* 6: 3461-3464.

Zhang, B.Y., Zhou, H. and Zhu, H., 2009. Explaining credit default swap spreads with the equity volatility and jump risks of individual firms. *The Review of Financial Studies* 22: 5099-5131.

Appendix: REIT CDS sample

| | Institution Name | Ticker | Sample Period of CDS | Acquisition |
|---|---|---|---|---|
| 1 | AMB Property Corporation | AMB | Jan 05 - Jun 07; Oct 09 - Sep 10 | Jun 11 |
| 2 | American Tower Corporation (REIT) | AMT | Jan 12 - Sep 16 | |
| 3 | AvalonBay Communities, Inc. | AVB | Jan 05 - Sep 16 | |
| 4 | Brandywine Realty Trust | BDN | July 07 - Sep 16 | |
| 5 | BRE Properties, Inc. | BRE | Jan 05 - Jun 07; Jun 08 - Mar 14 | Apr 14 |
| 6 | Boston Properties, Inc. | BXP | Jan 05 - Sep 16 | |
| 7 | CarrAmerica Realty Corporation | CARR | Jan 05 - Apr 06 | Jun 06 |
| 8 | Crown Castle International Corp. (REIT) | CCI | Jan 14 - Sep 16 | |
| 9 | Mack-Cali Realty Corporation | CLI | Jan 05 - Sep 16 | |
| 10 | Camden Property Trust | CPT | Mar 08 - Sep 16 | |
| 11 | DDR Corp. | DDR | Jan 05 - Sep 16 | |
| 12 | Duke Realty Corporation | DRE | Jan 05 - Sep 16 | |
| 13 | Equity Office Properties Trust | EOP | Jan 05 -Dec 06 | Feb 07 |
| 14 | Equity Commonwealth | EQC | Jan 10 - Sep 16 | |
| 15 | Equinix, Inc. | EQIX | Apr 16 - Sep 16 | |
| 16 | Equity Residential | EQR | Jan 05 - Sep 16 | |
| 17 | Equity One, Inc. | EQY | Jan 10 - Sep 16 | |
| 18 | FelCor Lodging Trust Incorporated | FCH | Jan 05 - Mar 16 | |
| 19 | First Industrial Realty Trust, Inc. | FR | Jan 06 - Sep 16 | |
| 20 | Federal Realty Investment Trust | FRT | Jan 08 - Sep 16 | |
| 21 | GGP Inc. | GGP | Jan 05 - Sep 08 | |
| 22 | Welltower Inc. | HCN | Oct 05 - Sep 16 | |
| 23 | HCP, Inc. | HCP | Feb 08 - Sep 16 | |
| 24 | Highwoods Properties, Inc. | HIW | Jan 05 - Dec 06; Oct 09 - Sep 16 | |
| 25 | Hospitality Properties Trust | HPT | Jan 08 - Sep 16 | |
| 26 | Healthcare Realty Trust Incorporated | HR | Jan 05 - Jun 07; Feb 10 - Sep 16 | |
| 27 | Host Hotels & Resorts, Inc. | HST | Jan 05 - Sep 16 | |
| 28 | Kimco Realty Corporation | KIM | Jan 05 - Sep 16 | |
| 29 | Liberty Property Trust | LPT | 0v 05 - Dec 06; Apr 09 - Sep 16 | |
| 30 | Nationwide Health Properties, Inc. | NHP | Jan 05 - Jul 07; Jan 09 - Apr 11 | Jul 11 |
| 31 | Prologis, Inc. | PLD | Jan 05 - Sep 16 | |
| 32 | Post Properties, Inc. | PPS | Jul 10 - Sep 16 | Jan 16 |
| 33 | Regency Centers Corporation | REG | Jan 10 - Jun 11; May 13 - Sep 16 | |
| 34 | Rayonier Inc. | RYN | Jan 08 - Aug 10 | |
| 35 | SL Green Realty Corp. | SLG | Apr 10 - Sep 16 | |
| 36 | Simon Property Group, Inc. | SPG | Jan 05 - Sep 16 | |
| 37 | iStar Inc. | STAR | Feb 05 - Sep 16 | |
| 38 | UDR, Inc. | UDR | Oct 05 - Sep 16 | |
| 39 | Uniti Group Inc | UNIT | Aug 15 - Sep 16 | |
| 40 | Vornado Realty Trust | VNO | Jan 05 - Sep 16 | |
| 41 | Ventas, Inc. | VTR | Feb 12 - Sep 16 | |

Appendix: REIT CDS sample (*continued*)

| | Institution Name | Ticker | Sample Period of CDS | Acquisition |
|---|---|---|---|---|
| 42 | Washington Real Estate Investment Trust | WRE | Jan 08 - Sep 16 | |
| 43 | Weingarten Realty Investors | WRI | Jun 05 - Sep 16 | |
| 44 | Weyerhaeuser Company | WY | Jan 10 - Sep 16 | |

*Notes:* Our sample includes all REITs that have CDS pricing available from January 2005 to September 2016. Two REITs did not meet our data filter requirements. MGM Growth Properties and National Retail Properties were omitted from the sample because of insufficient observations and insufficient price quotes within our sample period, respectively. Seven REITs in the sample have some missing pricing data, as noted above. Six REITs were acquired during our sample time frame, as noted above.

Table 1.1. Parameter distribution in CDS pricing model

| Panel A: Risk-free term structure factor loadings and dynamics | | | | | |
|---|---|---|---|---|---|
| | $\mu$*100 | $\phi$ | $\Sigma$*100 | $\lambda_0$ | $\lambda_1$ |
| $X_1$ | 0.19 | 0.93 | 0.89 | 0.2 | 5.81 |
| $X_2$ | 0.03 | 0.92 | 0.42 | 0.27 | 8.68 |
| **Panel B: Model RMSE (bps) and measurement error standard deviation (bps)** | | | | | |
| | 6 months | 1 year | 3 year | 5 year | 10 year |
| RMSE | 0.93 | 0.81 | 0.78 | 0.86 | 1.06 |
| ME SD | 0.94 | 0.8 | 0.68 | 0.59 | 0.49 |
| **Panel C: Loadings of parameters in CDS structure model** | | | | | |
| Constant | $\alpha$_$X_1$ | $\alpha$_$X_2$ | $\alpha$_Leverage | $\alpha$_Volatility | $\alpha$_FFOpayout |
| 1.36 | 47.63 | 5.34 | 0.87 | 0.46 | -0.39 |

*Notes:* Panel A reports the distribution of parameters in the risk-free rate model. The two state factors are estimated by the unscented Kalman filter. We use six-month and 1-, 3-, 5-, 10-year zero-coupon Treasury yields to simulate the risk-free rate. Panel B presents root-mean-squared errors (RMSEs) and measurement error standard deviations (ME SD) for the risk-free model. Panel C reports parameter estimates for the CDS pricing model.

Table 1.2. Variable definitions

| Variable | Source | Definition | Predicted Sign | Economic intuition |
|---|---|---|---|---|
| *Market Cap* | SNL; CRSP | Closing price *outstanding shares | - | Firm with larger market size withstands higher probability of default (PD). |
| *Real Estate Value* | SNL | Market Cap - non-operational real estate asset | - | Small-size REITs are vulnerable to default risk |
| *Market Leverage* | SNL | (Reported debt + reported prefered equity) / real estate value | + | Merton (1974) shows that high leverage increases PD |
| *FFO Payout* | SNL | Declared dividends / fund from operation (calculated by SNL) | +/- | Higher payout means more profit, which decreases PD. However, the regulation of 90% dividend payout on REITs means less residual value to protect from default, which exaggerates default risk. |
| *Profitability* | SNL | FFO (calculated by SNL) / total asset value | - | PD declines when the earning performance improves. |
| *Interest Coverage* | SNL | Net operating income / (interest expense + prefered dividend) | - | Higher interest coverage rate indicates a solid solvency |
| *Historical Volatility* | Calculation | Annualized standard deviation of daily return, over prior 14 and 365 calendar days | + | Merton (1974) framework indicates that larger equity voaltility is more likely to cause credit default. |
| *Implied Volatility* | Optionmatrix | The one-year 25-delta put option implied volatility | + | Cao et al. (2009) find that implied volatility dominates historical volatility in explaining the change of CDS spreads. |
| *Distance to Default* | Calculation | (firm value - half of long-term debt) / (firm value * volatility of firm) | - | KMV model documents that distance to default is negative related to PD. |
| *Momentum* | Calculation | Stock return over past 26 and 52 trading weeks | - | Higher equity value increases the firm capitalization, thereby reducing PD. |
| *Short Sale* | SNL; Compustat | percent of short shares to total outstanding shares | + | A large percentage of short interest reflects a deteriorating future of firm. |
| *VIX* | Bloomberg | the market's expectation of stock market volatility over the next 30-day period. | + | As market's fear gauge, large VIX means that much fluctuation lingers over market |
| *Market Return* | Caculation | The return of S&P 500 index over past one year | - | Higher market returns indicate an improved economic environment. |
| *Market Volatility* | Caculation | Annualized standard deviation of daily return of S&P 500 index over past one year | + | Higher market volatility deteriorates the overall economic conditions, which increase PD of individual firm. |
| *REITs Market Return* | Caculation | The return of SNL REITs index over past one year | - | Better performance of REITs market decrease PD of REITs market |
| *REITs Market Volatiltiy* | Caculation | Annualized standard deviation of daily return of SNL REITs index, over past one year | + | Higher volatility of REITs market causes larger credit spread of individual REIT. |
| *CPI* | Bureau of Labor Statistics | The Consumer Price Indexes | - | Higher CPI indicates a prosperous real estate market. Furthermore, commercial real estate leases usually adjust by CPI. |
| *Market Credit Risk* | Datastream | Moody's Baa yield - Moody's AAA yield | + | A higher market credit risk increases PD of individual REIT. |
| *Market Liquidity Risk* | Datastream | 3 month LIBOR rate - 3 month Treasury bill rate | + | A higher market liquidity risk increases PD of individual REIT. |
| *Risk-Free Rate* | Fred bank of ST.Louis | Daily 1-Year Treasury constant maturity rate | +/- | A higher spot rate increases the risk-neutral drift of the firm value process and reduces PD (Longstaff and Schwartz 1995). Nevertheless, it may reflect a tightened monetary policy stance which discourages investment in real estate, therefore increasing PD of REITs. |
| *Slope of Yield Curve* | Datastream | 10-Year Treasury constant maturity - 5-Year Treasury constant maturity | +/- | A steeper slope of the term structure is an indicator of improving economic activity in the future, but it can also forecast a tightening of monetary policy. |
| *Credit Rating* | SNL | Dummy variable. 1 indicates that the credit rating of REIT is investment grade. 0 is the proxy of speculative-grade REITs. | - | Compared with speculative-grade REITs, investment-grade REITs have low PD. |
| *Property Type* | SNL | There are seven property types of REITs, as specialty, multifamily, office, shopping center, industrial, health care, and hotel. | | Compared with other types multifamily is considered as low-risk industry. The multifamily REITs usually enjoy stable and constant lease rents. Futhermore, Fannie Mae and Freddie Mac subsidise multifamily REITs. |
| *Location* | SNL | We subdivide REITs into four geographic segments, such as northeast, midwest, south, and west. | | The real estate market in northeast is more prosperious than that in other geographic regions, which leads to REITs in northeast with low PD. |

*Notes:* This table provides definitions and data resource for the variables analyzed in this study. The predicted sign is the forecasted sign of each variables in the panel regression to price CDS spread of REITs.

Table 1.3. Summary Statistics

| | Pooled | | | 1st Quintile *CDS* | | | 5th Quintile *CDS* | | | Std Dev | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Variables | Mean | SD | Correlation w/ *CDS* | Mean | SD | Correlation w/ *CDS* | Mean | SD | Correlation w/ *CDS* | Cross Section | Time Series |
| *CDS(bps)* | 200.18 | 288.00 | | 45.63 | 13.71 | | 537.95 | 509.70 | | 151.85 | 149.68 |
| *In(Market Cap.)* | 8.63 | 1.33 | -0.35*** | 8.92 | 0.92 | 0.07*** | 7.67 | 1.26 | -0.50*** | 1.20 | 0.36 |
| *In(Real Estate Value)* | 8.99 | 0.79 | -0.29*** | 9.26 | 0.77 | 0.05*** | 8.66 | 0.75 | -0.08*** | 0.75 | 0.22 |
| *Book Leverage* | 0.53 | 0.11 | 0.24*** | 0.53 | 0.10 | -0.15*** | 0.60 | 0.15 | 0.32*** | 0.11 | 0.04 |
| *Market Leverage* | 0.50 | 0.19 | 0.55*** | 0.40 | 0.10 | -0.09*** | 0.69 | 0.24 | 0.35*** | 0.16 | 0.10 |
| *FFO Payout* | 0.70 | 0.85 | -0.06*** | 0.79 | 1.24 | 0.03 | 0.65 | 0.67 | 0.06*** | 0.51 | 0.42 |
| *Profitability* | 0.01 | 0.01 | -0.19*** | 0.02 | 0.00 | -0.06*** | 0.01 | 0.01 | -0.05*** | 0.01 | 0.00 |
| *Interest Coverage* | 3.04 | 1.32 | -0.27*** | 3.11 | 1.25 | 0.27*** | 2.22 | 1.16 | -0.40*** | 1.14 | 0.59 |
| *Volatility (14)* | 0.30 | 0.32 | 0.60*** | 0.19 | 0.08 | -0.05*** | 0.64 | 0.55 | 0.69*** | 0.11 | 0.12 |
| *Volatility (365)* | 0.34 | 0.29 | 0.66*** | 0.19 | 0.03 | -0.01 | 0.67 | 0.44 | 0.66*** | 0.09 | 0.10 |
| *Implied Volatility* | 0.35 | 0.20 | 0.75*** | 0.23 | 0.05 | 0.06*** | 0.60 | 0.28 | 0.84*** | 0.09 | 0.10 |
| *Dis to Default* | 4.22 | 1.56 | -0.59*** | 5.27 | 0.76 | 0.11*** | 0.60 | 1.31 | -0.47*** | 0.53 | 0.59 |
| *Return(52)* | 0.04 | 0.37 | -0.44*** | 0.13 | 0.14 | -0.35*** | -0.25 | 0.66 | -0.55*** | 0.19 | 0.28 |
| *Return(26)* | 0.02 | 0.27 | -0.33*** | 0.06 | 0.11 | -0.27*** | -0.13 | 0.51 | -0.49*** | 0.14 | 0.21 |
| *Short Sale* | 5.07 | 4.02 | 0.52*** | 3.20 | 2.09 | 0.08*** | 8.85 | 5.63 | 0.46*** | 3.12 | 2.58 |
| *Credit Rating* | 4.12 | 0.65 | 0.47*** | 3.75 | 0.44 | 0.14*** | 4.66 | 0.88 | 0.14*** | 0.64 | 0.19 |
| *Number of REITs* | 44 | | | | | | | | | | |
| *Firm weeks* | 18,405 | | | | | | | | | | |

*Notes:* This table reports descriptive statistics of REITs characteristics for 44 REITs over 613 weeks from Jan 7, 2005 to Sep 30, 2016, a total of 18,405 firm-week observation. The table first provides the average, the stand deviation, and the correlation with logarithm of *CDS spreads* of each REITs fundamental on both the pooled sample and at first and last CDS quintiles. Then, the table provides two measures of standard deviation: (i) XS – time-series averages of the cross-sectional standard deviation estimates on each date; (ii) TS – cross-sectional averages of the time series standard deviation estimates for each firm. '*', '**' and '***' denotes significance at the 10% 5% and 1% level, respectively.

Table 1.4. Merton-Implied Model Estimates

| | Model 1.0 | | Model 1.1 | |
|---|---|---|---|---|
| Variables | Coefficients | t-Stat. | Coefficients | t-Stat. |
| *Intercept* | 4.10*** | (-30.7) | | |
| *Leverage* | 1.14*** | (-3.8) | 0.84*** | (3.5) |
| *Volatility (14)* | 0.90*** | (-10.8) | 0.86*** | (8.9) |
| *Volatility (365)* | 0.56*** | (-6.7) | 0.63*** | (7.1) |
| *Risk-Free Rate* | -0.20*** | (-10.5) | -0.19*** | (-10.8) |
| *Fixed Effect* | No | | Yes | |
| Adj $R^2$ | 0.67 | | 0.73 | |

*Notes:* Table presents descriptive statistics and regression results for panel regressions using the explanatory variables suggested by Merton (1974). The dependent variable is the natural logarithm of CDS Spread. White standard errors which are robust to within cluster correlation are used to compute t-statistics. Compared with Model 1.0, Model 1.1 includes the fixed effect. '*', '**' and '***' denotes significance at the 10% 5% and 1% level, respectively.

Table 1.5. REIT-Specific Model Estimates

| | Model 2.0 | | Model 2.1 | | Model 2.2 | |
|---|---|---|---|---|---|---|
| Variables | Coefficients | t-Stat. | Coefficients | t-Stat. | Coefficients | t-Stat. |
| *Intercept* | 6.76*** | (13.9) | | | 5.61*** | (11.4) |
| *ln(Real Estate Value)* | -0.14*** | (-2.4) | -0.06** | (-1.4) | -0.08** | (-1.6) |
| *FFO Payout* | -0.04*** | (-3.1) | -0.03*** | (-3.0) | -0.04*** | (-3.1) |
| *Profitability* | -3.89 | (-0.8) | -1.11 | (-0.7) | -3.54 | (-0.8) |
| *Interest Coverage* | -0.07*** | (-2.0) | -0.10*** | (-2.3) | -0.08*** | (-2.7) |
| *Short Sale* | 0.03*** | (3.6) | 0.04*** | (3.7) | 0.03*** | (4.1) |
| *SNL Return* | -0.89*** | (-12.7) | -0.82*** | (-11.6) | -0.86*** | (-12.4) |
| *SNL Volatility* | 1.05*** | (9.4) | 1.14*** | (8.7) | 1.08*** | (10.6) |
| *Risk-Free Rate* | -0.20*** | (-12.1) | -0.20*** | (-13.9) | -0.20*** | (-12.4) |
| *Credit Rating* | -0.58*** | (-6.3) | | | -0.29*** | (-3.9) |
| *Specialty* | | | | | 0.60*** | (2.6) |
| *Office* | | | | | 0.32*** | (2.6) |
| *Shopping Center* | | | | | 0.13 | (1.0) |
| *Industrial* | | | | | 0.28*** | (3.1) |
| *Health Care* | | | | | 0.22*** | (3.3) |
| *Hotel* | | | | | 0.69*** | (5.3) |
| *Midwest* | | | | | 0.15** | (1.5) |
| *South* | | | | | 0.24*** | (2.6) |
| *West* | | | | | 0.16** | (1.6) |
| *Fixed Effect* | No | | Yes | | No | |
| Adj $R^2$ | 0.71 | | 0.75 | | 0.74 | |

*Notes:* Table presents descriptive statistics and regression results for panel regressions using the REITs-specific explanatory variables. The dependent variable is the natural logarithm of CDS Spread. White standard errors which are robust to within cluster correlation are used to compute t-statistics. Compared with Model 2.0, Model 2.1 includes the fixed effect. Model 2.2 adds dummy variables which proxy for property types and location into Model 2.0. Controls for self-managed, self-advised, and UPREIT status are omitted due to the homogeneity of these variables within our sample. '*', '**' and '***' denotes significance at the 10% 5% and 1% level, respectively.

Table 1.6. Comprehensive Factors Model Estimates

| | Model 3.0 | | Model 3.1 | | Model 3.2 | |
|---|---|---|---|---|---|---|
| Variables | Coefficients | t-Stat. | Coefficients | t-Stat. | Coefficients | t-Stat. |
| *Intercept* | 3.03 | (0.9) | | | 4.68* | (1.5) |
| *ln(Market Cap)* | 0.03** | (2.1) | -0.15** | (-1.7) | 0.06*** | (3.6) |
| *ln(Real Estate Value)* | -0.18*** | (-3.8) | 0.20** | (1.9) | -0.14*** | (-2.8) |
| *Leverage* | 0.79*** | (2.4) | 0.44** | (1.9) | 0.73*** | (2.1) |
| *FFO Payout* | -0.02** | (-1.8) | -0.01* | (-1.3) | -0.03*** | (-2.4) |
| *Profitability* | -3.91 | (-0.8) | -0.04 | (-0.9) | 0.02 | (0.5) |
| *Interest Coverage* | 0.03 | (0.6) | -3.79*** | (-3.2) | -8.41** | (-2.2) |
| *Volatility (14)* | 0.01 | (0.1) | -0.06** | (-1.7) | -0.04 | (-0.7) |
| *Volatility (365)* | -0.69* | (-1.5) | 0.09 | (0.3) | -0.53* | (-1.3) |
| *Implied Volatility* | 0.53*** | (3.6) | 0.36*** | (3.5) | 0.60*** | (4.0) |
| *Dis to Default* | -0.07*** | (-2.7) | -0.03*** | (-2.6) | -0.04** | (-1.7) |
| *Return(26)* | -0.24*** | (-4.5) | -0.12*** | (-2.8) | -0.19*** | (-4.3) |
| *Short Sale* | 0.00 | (-0.4) | 0.01** | (2.1) | 0.00 | (-0.6) |
| *VIX%* | 0.50*** | (3.1) | 0.61*** | (4.7) | 0.62*** | (4.2) |
| *S&P Return* | -0.16 | (-0.9) | -0.21** | (-1.6) | -0.09 | (-0.6) |
| *S&P Volatility* | 0.71 | (1.1) | 1.48*** | (2.6) | 1.09 | (1.8) |
| *SNL Return* | -0.04 | (-0.5) | -0.03 | (-0.4) | -0.05 | (-0.5) |
| *SNL Volatility* | 0.52 | (1.0) | -0.26 | (-0.5) | 0.38 | (0.9) |
| *ln(CPI)* | 0.60 | (0.9) | -0.18 | (-0.4) | 0.23 | (0.4) |
| *Market Credit Risk* | 0.31*** | (7.9) | 0.30*** | (9.2) | 0.33*** | (10.6) |
| *Market Liquidity Risk* | 0.19*** | (5.6) | 0.21*** | (8.3) | 0.20*** | (6.4) |
| *Slope of Yield Curve* | -0.19*** | (-3.7) | -0.14*** | (-2.5) | -0.14*** | (-2.9) |
| *Risk-Free Rate* | -0.07*** | (-3.4) | -0.08*** | (-3.6) | -0.06*** | (-3.0) |
| *Credit Rating* | -0.54*** | (-5.7) | | | -0.35*** | (-5.7) |
| *Specialty* | | | | | 0.58*** | (3.1) |
| *Office* | | | | | 0.32*** | (2.4) |
| *Shopping Center* | | | | | 0.22*** | (2.0) |
| *Industrial* | | | | | 0.29** | (2.2) |
| *Health Care* | | | | | 0.27*** | (4.5) |
| *Hotel* | | | | | 0.62*** | (4.9) |
| *Midwest* | | | | | 0.06 | (0.6) |
| *South* | | | | | 0.22** | (2.1) |
| *West* | | | | | 0.12 | (1.0) |
| *Fixed Effect* | No | | Yes | | No | |
| Adj $R^2$ | 0.79 | | 0.86 | | 0.83 | |

*Notes:* Table presents descriptive statistics and regression results for panel regressions using a comprehensive list of characteristics. The dependent variable is the natural logarithm of CDS Spread. White standard errors which are robust to within cluster correlation are used to compute t-statistics. Compared with Model 3.0, Model 3.1 includes the fixed effect. Model 3.2 adds dummy variables which proxy for property types and location into Model 3.0. '*', '**' and '***' denotes significance at the 10% 5% and 1% level, respectively.

Table 1.7. Interaction Effect Model

| | Model 1 | | | Model 2 | |
|---|---|---|---|---|---|
| Variables | Coefficients | t-Stat. | Variables | Coefficients | t-Stat. |
| *Intercept* | 4.48*** | 22.18*** | *Intercept* | 6.34*** | (8.5) |
| *Leverage* | 1.12*** | 2.82*** | *ln(Real Estate Value)* | -0.12* | (-1.4) |
| *Volatility (14)* | 0.53*** | 4.033*** | *FFO Payout* | 0.08 | (0.3) |
| *Volatility (365)* | 0.58*** | 6.06*** | *Profitability* | 4.92 | (1.1) |
| *Risk-Free Rate* | -0.01 | -0.93 | *Interest Coverage* | -0.10*** | (-2.8) |
| *Credit Rating* | -0.33* | -1.53* | *Short Sale* | -0.01 | (-0.6) |
| *Leverage*Rating* | -0.27*** | -4.13*** | *SNL Return* | -0.16 | (-1.0) |
| *Volatility (14)*Rating* | 0.46*** | 3.16*** | *SNL Volatility* | 2.30*** | (6.1) |
| *Volatility (365)*Rating* | -0.01 | -0.07 | *Risk-Free Rate* | -0.03 | (-1.1) |
| *Risk-Free Rate*Rating* | -0.19*** | -10.51*** | *Credit Rating* | -0.03 | (0.0) |
| | | | *ln(Real-V)*Rating* | -0.01 | (-0.1) |
| | | | *FFO Payout*Rating* | -0.13 | (-0.6) |
| | | | *Profitability*Rating* | -16.96*** | (-1.7) |
| | | | *Interest Cover*Rating* | 0.05 | (0.8) |
| | | | *Short Sale*Rating* | 0.04*** | (1.8) |
| | | | *SNL Return*Rating* | -0.81*** | (-4.5) |
| | | | *SNL Volatility*Rating* | -1.40*** | (-3.5) |
| | | | *Risk-Free Rate*Rating* | -0.18*** | (-5.6) |
| Adj $R^2$ | 0.74 | | Adj $R^2$ | 0.73 | |

*Notes:* Table presents descriptive statistics and regression results for panel regressions using interaction effect with Investment Rating. The dependent variable is the natural logarithm of CDS Spread. White standard errors which are robust to within cluster correlation are used to compute t-statistics. As a dummy variable, Investment Rating is equal to 1 for all REITs with investment-grade rate, and is equal to 0 for all REITs with speculative-grade rate. The Model 4.0 examines the interaction effects between Merton-implied covariates and credit rating. The Model 4.1 shows the performance of interaction effects between REITs-specific characteristics and credit rating. '*', '**' and '***' denotes significance at the 10% 5% and 1% level, respectively.

Table 1.8. An Out-of-Sample Investment Exercise

| | Full Sample | | | Pre-Big Bang | | | Post-Big Bang | | |
|---|---|---|---|---|---|---|---|---|---|
| Horizon (weeks) | Mean | SD | SR | Mean | SD | SR | Mean | SD | SR |
| 1 | 2.83% | 7.19% | 0.39 | 2.84% | 7.19% | 0.39 | 1.45% | 4.42% | 0.33 |
| 2 | 2.80% | 10.75% | 0.26 | 2.81% | 10.75% | 0.26 | 2.01% | 5.50% | 0.37 |
| 3 | 4.98% | 17.23% | 0.29 | 4.99% | 17.23% | 0.29 | 3.08% | 7.81% | 0.39 |
| 4 | 5.74% | 17.30% | 0.33 | 5.76% | 17.30% | 0.33 | 2.96% | 9.52% | 0.31 |

*Notes:* Table summarizes the average excess return (mean), standard deviation (SD), and the sharp ratio (SR) from an out-of-sample investment exercise over different horizons (in number of weeks). Every month, we execute a zero-cost investment portfolio in the CDS market of REITs based on the deviation between the market CDS quotes and model-implied CDS valuations. Panel A shows the performance of investment strategy across whole sample period. The statistics of investment portfolio before the implementation of "Big Bang" protocol (Jan 2005 – Dec 2009) are documented in Panel B. Panel C describes the excess return of out-of-sample investment after the enactment of "Big Bang" protocol (Jan 2010 – Sep 2016).

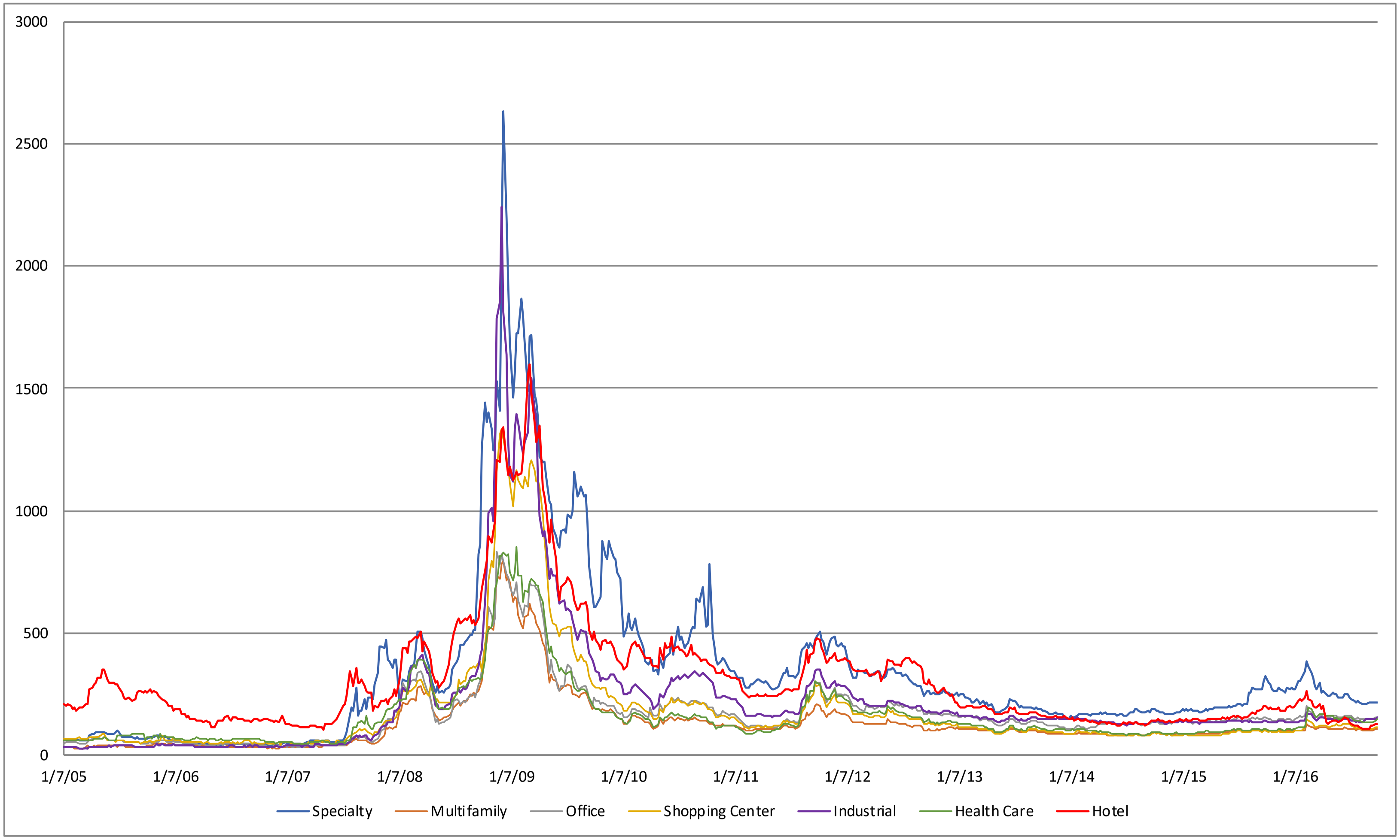



Figure 1.1. The Time-Series Average of CDS Spreads of REITs Property Types. This figure shows our CDS spread data into seven sections based on the property types of REITs, as specialty, multifamily, office, shopping center, industrial, heath care, and hotel. Each week, we calculate the average of CDS spreads of REITs in each property type. This figure describes the time series of average CDS spreads of REITs in seven property types respectively.

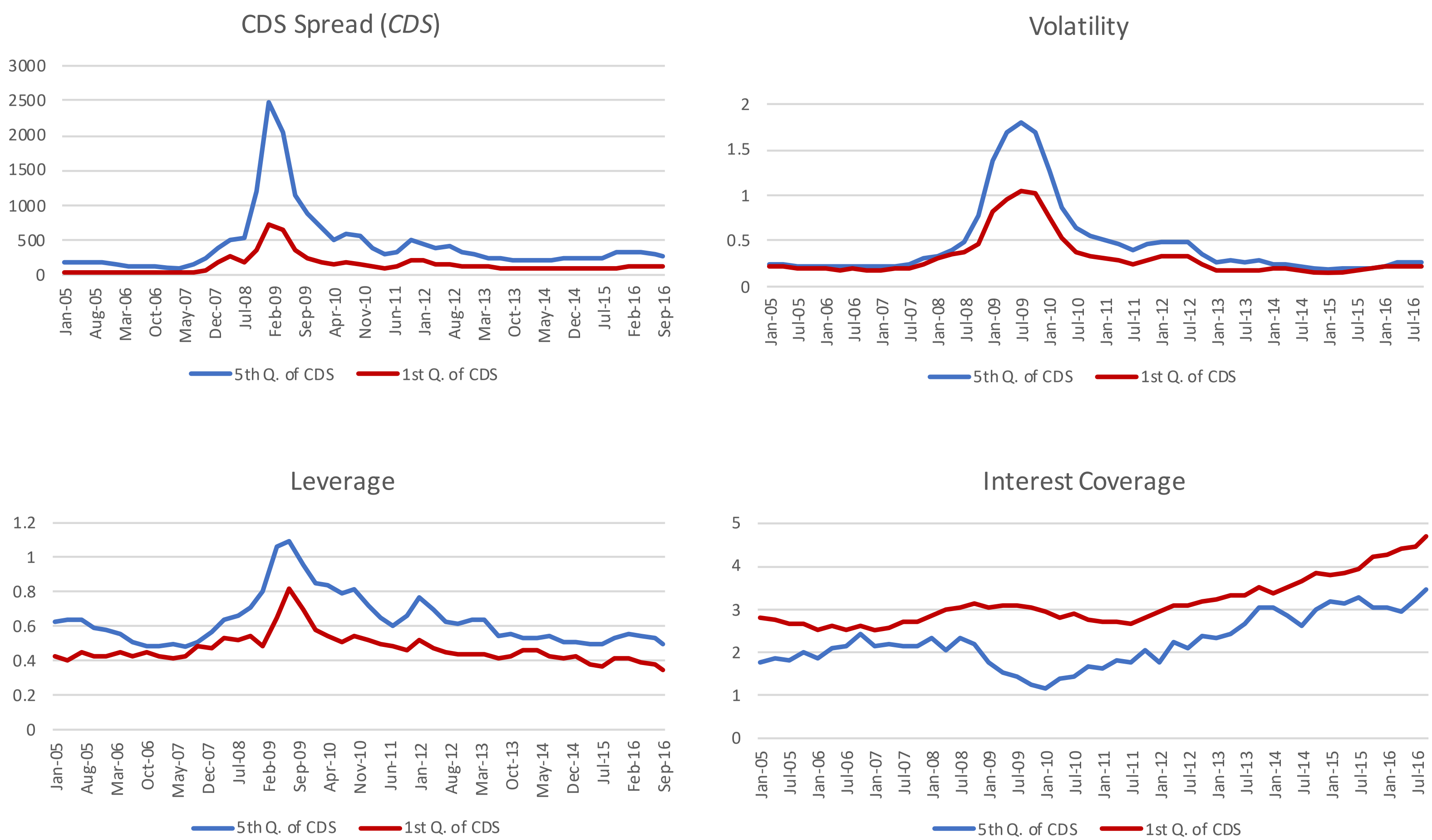



Figure 2. The Time Series of Characteristics by First and Last Quintile of CDS Spread This figure illustrates the time series relationship between our primary explantory variables and CDS spread.

# CHAPTER 2: THE FLOW OF CREDIT RISK INFORMATION AMONG REIT SECURITIES

## 1. Introduction

According to the efficient market hypothesis, in a perfectly active market, credit risk information of the same firm should be timely and efficiently reflected on all securities markets. However, due to market frictions (such as trading cost and regulations) speculators typically choose to trade such information in only one or few markets swiftly. Thus, there is a direction of credit information flow among different securities markets considering the information content, speed, and efficiency.

With the rapid development of credit derivative market, a rapidly growing thread of finance literature has focused on whether the credit default swap (CDS) market is a catalyst for the exploitation of private information or if it deters speculation and price discovery. Acharya and Johnson (2007) is the first paper that sheds light on the manipulation of non-public information in the CDS market. They argue that the largest banks, which are also key players in the CDS market, exploit their privileged information about the default risk in the credit derivative market. Specifically, due to the close relationship between debtors and creditors, banks have exclusive access to material information about firms' credit quality, such as updated income projection or financing plan. For lack of regulated "Chinese walls" within conglomerate banks, it is highly likely that the trading desks of banks profit from such non-public information in the CDS market. Following this masterpiece, on May 2009, the United States Securities and Exchange Commission (SEC) filed its first-ever insider trading case involving CDS against Jon-Paul Rorech, who is a former CDS and bond salesman at Deutsche Bank. (WSJ, May 5, 2009)

In this paper, we further extend previous related literature on the information flow among three securities marketed related to real estate investment (REIT), named as CDS market, bond market, and stock market. Compared with the developed CDS market of other industries, the CDS market on REITs is still fledging and innovative regarding market volume and numbers of contracts. As a niche market, REIT CDS market does not grasp much attention in the finance search. To our best known, it is the first paper that focuses on the information flow related to REIT CDS market.

Although the market size of REITs CDS is inconspicuous, the unique regulation and firm-organization make REITs-related security markets into a distinctive laboratory to test for the argument about information flow among various security markets. First of all, in contrast to other industry, REITs build a closer relationship with the capital market. Given REITs have to pass earning to shareholders, REITs cannot retain extra capital in the balance sheet. To success in real estate investment, a capital-intensive business, REITs must keep consistent access to the financers, such as banks and institutional investors. According to the report by Fitch Ratings (WSJ, April 12, 2017), the REITs significantly increase loan from commercial banks in recent years. For REITs, the recent bank borrowing exposure accounted for 18.8% of total debt on average. Since the large proportion of bank loan, REITs is a proper candidate to verify the findings of Acharya and Johnson (2007).

On the other side, except for bank financing, REITs also heavily rely on capital sources of public, private, debt and equity. Thus, the credit profile of REITs is also exposed to multiple capital markets. Also, a previous study presents that over 60% of REITs' debt is secured debt, compared to less than 5% in non-REITs, and most of the collateral is real estate. According to the regulation, a REIT must invest at least 75 percent of its total assets in real estate. As a tangible asset, the estimation value of a real estate is more accurate and transparent, which diminishes the degree of non-public information about REITs' credit equality. Furthermore, on average, over 90% of REITs' shareholders are institutional investors who are considered as informed investors. The concentration of informed investors in REITs may prevent speculators from exploiting credit-related information in the CDS market.

Furthermore, as an over-the-counter (OTC) market, the CDS market is a lack of regulation and vulnerable to insider transactions. However, as a part of the Dodd-Frank reform, the implementation of "Big Bang 2009" regulation helps standardize the CDS market and build a central clearinghouse for CDS transactions. Loon and Zhong (2016) find that, after the commencement of regulation on the CDS market, the liquidity in CDS market considerably improves and transaction cost decrease. These findings will encourage more information trading in the CDS market. The period of our CDS data is from 2005 to 2016, which reflects the influence of regulation on the CDS market entirely. In general, we examine the lead-lag relation among three REITs-related security markets to analyze if and how these markets react and whether credit-related information is discovered earlier in specific markets than in others.

There is widespread controversy over the direction of credit information discovery among equity, bond, and CDS markets. The evidence on whether equity returns lead CDS returns is mixed. Norden and Weber (2009), Marsh and Wagner (2012), Hilscher, Pollet, and Wilson (2015) and Narayan, Sharama, and Thuraisamy (2014) show that equity market dominates the CDS market in convey of default-related information under most situation, which indicates the absence of price discovery in the CDS market. Hilscher et al. (2015) apply the panel VAR model with daily and weekly CDS data and indicate that lagged equity returns predict CDS returns up to several weeks, not vise vice. They also argue that the considerable trading costs in the CDS market, which are measured by bid-ask spreads, delay the process of price discovery and prevent informed investors from trading in the CDS market. Narayan et al. (2014) test the influence of information on the interaction between CDS and stock market in several diverse sectors. They present that the stock returns significantly lead CDS returns in most sectors and the CDS market contributes to information discovery only in a few stances.

In contrast, Acharya and Johnson (2007) use CDS quotes from 2001 to 2004 to document that the CDS market conveys unique information about upcoming credit news to the equity market. This information discovery only works for adverse firm-specific information events, such as downgrade rating, but not for positive credit news. They further find that the close relationship between banks and underlying firms causes such asymmetric information revelation. This finding implies that speculators primarily profit from information about adverse credit news in the CDS market. The argument that that stock and CDS markets are in a pooling equilibrium and the credit information flows between each other is also buttressed by Bolton and Oehmke (2011), Xiang, Chang, and Fang (2013) and Lee, Naranjo, and Velioglu (2018). Lee et al. (2018) first use daily data to test the lead-lag relation for original CDS returns and stock returns and find that stock prices lead CDS prices. They further the direction of firm-specific the information flow between these two markets. The CDS market is found to contribute to price discovery, while the stock market still predominately reflects information process. Then, they check the exclusive information in the CDS market before the announcement of rating events. They confirm that, before negative credit events, CDS returns strongly predict future equity return. Banks trading CDS with private information on firm-specific credit risk attributes to this price discovery.

Moreover, Longstaff, Mithal, and Neis (2005) reveal that both the CDS and stock market move simultaneously but lead the bond market, while Kryzanowski, Perrakis, and Zhong (2017) suggest that neither stock market nor CDS market generally exhibits an advantage of price discovery. In this paper, we investigate whether the REIT CDS market will contribute to price discovery across related stock and bond markets, and, what factors contribute to the timely and efficient detection of pricing information in REITs CDS market?

Since the publication of Acharya and Johnson (2007), insider information in the credit derivatives market suddenly become a hot topic. A bench of studies discloses that informed banks, which are not only the major players in the CDS market but also keep close loan relationships with CDS reference entities, profit from material nonpublic information in the CDS market.[1] Norden and Wagner (2008) even uncover that the price information in the CDS market positively decides the charge of syndicate loan to firms. Compared with the equity market, the CDS market is mostly unregulated and has limited participants. Reiser (2011) lists three reasons that cause insider trading in CDS market: (1) most of the participants in CDS market possess non-public information; (2) the moral hazards in CDS market fertilize the insider trading; and (3) CDS market is opaque, and trading data is inaccessible. Recently, Fecht, Hackethal, and Karabulut (2018) and Ivashina and Sun (2011) argue that, due to the close lending and other direct business relationships

---

[1] Credit derivatives markets are much more limited in scope. Participants in this market are almost exclusively institutional investors, with banks forming the largest group: 60% of CDS protection in 2006 was bought by banks, 28% by hedge funds and 6% by insurance companies (source: BBA, 2006). Besides providing liquidity, a key motive for banks taking CDS positions is to hedge (about one third of their credit derivatives positions are held in the loan book).

with borrowers, banks can collect and process proprietary information at a lower cost. Banks will profit from such monopoly information from security markets.

However, Hasan and Wu (2015) demonstrate that, although banks trade more CDS on their borrowers, their net CDS positions are mostly unrelated to lending relation. They find no evidence of a bank using CDS to exploit private information. Shan, Tang, and Yan (2016) show that firms tend to resort to new lenders after the inception of CDS trading on their debt, which implied that CDS trading exacerbates firm-bank lending relation and affects borrower's debt structure. In the following part, we will testify the existence of such bank-related price discovery in REITs CDS market.

Since both credit ratings and CDS spreads profile the credit quality of debtors, a growing literature identifies the marginal information about the future rating news in the CDS market. CDS spreads continuously fluctuate more or less, but credit ratings issued by credit agencies rarely change. If both indicators express the same credit information, we expect that the CDS spread changes lead the announcement of a credit event. Moreover, according to the Financial Crisis Inquiry Commission Report (2011), credit agencies manipulated the credit ratings and gamed the credit models. Since there was a severe conflict of interest between credit agencies and bond issuer, credit agencies undertake competitive pressure to lower standards and insist on stability objective. Hull, Predescu, and White (2004), Norden and Weber (2004), Norden (2017), and Lee et al. (2018) display that the CDS market notably reacts around the announcements of rating events. CDS spreads strongly react to upcoming downgrade events, while there is an only ignorable reaction of CDS spread to rating upgrades. Such finding is documented even approximately two to three months before downgrade events.

However, rating agencies have access to unpublished information, and the CDS market of REITs is inertial. There is still a possibility that CDS spread changes lag rating changes. Following previous studies, we examine to the extent that REITs CDS spreads fluctuate before and after a credit announcement.

Generally, previous price-discovery researches agree that the CDS and stock market dominate the bond market in terms of information revelation. (Longstaff et al. (2003), Blanco, Brennan, and Marsh (2005), Norden and Webber (2009), Lee et al. (2018)) Oehmke and Zawadowsk (2015a) conclude that trading frictions in the bond market and speculative motives are among some main reasons why investors trade in the CDS market versus in the bond markets. Considering that the REITs CDS market is relatively inactive, we will test whether or not the information in REITs CDS market still spills over to REITs bond market.

Our study first contributes to the reaction of REITs CDS market to the approaching announcement of a credit event. Due to the small size of REITs CDS, we only report 66 credit events from 2005 to 2016, with 36 positive rating events and 30 negative credit news. Although the event sample is limited, our finding is consistent with previous studies. The asymmetric reaction of CDS spread to rating events also exists in REITs CDS market. CDS spreads of REITs increase 8% as early as two months before the announcement of adverse rating event. For the severe downgrade event, from investment grade to speculative grade, the REITs CDS spreads jump over 20%. Meanwhile, there is no noticeable fluctuation in REITs CDS market prior to positive rating announcements.

Then, we apply the panel vector autoregression model (VAR) to explore the CDS-bond cross-market price discovery. Similar to the conclusion of previous literature, we also document that the sensitivity of REITs' bonds to credit information is always much lower than that of the associated CDS. Inspired by Lee et al. (2018), we further test the firm-specific information flow between REITs CDS and the bond market. After controlling the market trends in both markets, the CDS market still drastically and economically leads bond market in conveying the firm-specific default risk information.

We next examine the lead-lag relation between REITs CDS and equity market. Similar to Lee et al. (2018) and Hilscher et al. (2015), we also show that equity returns overwhelmingly lead CDS returns, unconditionally. However, when we estimate the same model but only during two months before the announcements of downgrade rating, we document that lagged CDS returns strongly predict current stock returns. While the stock market also distributes information to the CDS market simultaneously before the negative credit news, it is the CDS market that dominates the price discovery process.

To further discover the sources of information on future credit events in the REIT CDS market, we examine the influence of bank-relationship and loan size on the co-movement of the CDS and stock market, respectively. Following Lee et al. (2018), we use two measurements to proxy for the relationship between REITs and banks. We find that strong banking relationships cause the predictability of REITs CDS before credit events. Besides, we show that the information revelation for rating events in CDS market also increases with the amount of REITs' bank loan. Our findings support the argument by the study of the exploitation of non-public information in the credit derivative market. The informed banks, the market makers in the CDS market, tend to act in the CDS market with their propriety information on REITs firms through lending relationship.

Finally, we apply the panel VAR model on the integrated three-market analysis and further summarize the channel of credit information flow among all three markets. The results sharply identify that both CDS spread changes and stock returns persistently contribute to price discovery for bond spread changes. After we control the effect of the stock market on the CDS market, the bond market reflects some credit-related information before the CDS market, and then this information is capitalized in the CDS market. However, such information direction disperses before and after the credit events.

Our paper contributes to the literature in four aspects. First, we provide new evidence from REITs CDS market on the disproportionate reaction of the CDS market to future credit events. Second, we present that, without the influence of equity market, REITs CDS market dominates the bond market in realizing credit risk information. Similar to Lee et al. (2018), we also document that the concurrent equity trading information affects the lead-lag relation between bond and CDS market. Third, our study adds new evidence on the information channel between CDS and equity market. Generally, there is unidirectional information flow from REITs stock market to REITs CDS market.

Nevertheless, we find that the price discovery from CDS to equity only exists before the upcoming announcement of downgrade news. The existence of exclusive information about future credit news in REITs CDS market confirms the argument by Lee et al. (2018) that the CDS market is no longer just a sideshow to the equity market. Lastly, we unveil the sources of information trading in REITs CDS market. Adhering to Acharya and Johnson (2007), we demonstrate that strong lending relationship between REITs and banks drives the price discovery in the CDS market before the rating events. We further display that size of bank loan also causes such price discovery.

The remainder of the article is organized as follows: Section 2 describes the data and measurement. Section 3 presents our main results, and Section 4 concludes.

## 2. Data

We collect main variables from several data sources. The St. Louis Fed provides the five-year swap rate and five-year Treasury yield. Both Standard & Poor's credit rating events data and closed REITs prices are retrieved from SNL. The credit rating assesses long-term prospects of the REIT, with the term ranging from half to two years. The credit watches express a short-term concern about the changes in credit rating shortly. Finally, we collect 66 credit events, with 24 downgrades, 36 upgrades, and 6 adverse watch events.

Bond information and daily transaction data are obtained from Mergent FISD and FINRA's Trade Reporting and Compliance Engine (TRACE), respectively. Following Jens Dick-Nielsen (2014) and Bessembinder, Kahle, Maxwell, and Xu (2008) to filter error bond trading transactions, we diminish commission, corrected, and canceled trades from TRACE data. After then, we map the bond identification in TRACE to the corresponding indicator in FISD. Next, we follow Lee et al. (2018) to eliminate the transaction data of bonds with features that may bias trading price. We remove bonds with uncommon coupons, non-US dollar bonds, preferred equity, asset-backed bonds, convertible bond, bonds with warrants. As the underlying reference of standard CDS is a senior unsecured bond, we only keep the trading data of senior unsecured bonds. Since our CDS contracts refer to a constant five-year maturity, we create a synthetic five-year constant maturity bond yield to match our CDS spreads. Following Blanco et al. (2005), We pick up two bonds whose maturity is closest to 5 years and trading price is close to par. By linearly interpolating yields of these two bonds, we will estimate a five-year bond yield. Finally, we apply the cubic spline method to interpolate missing yields of the synthetic five-year constant maturity bond. We create the synthetic five-year constant maturity bond yields for 37 REITs.

The process of collecting CDS data is complex. As CDS contracts are the counter (OTC) derivatives, there is no standardized and unified database for CDS. Mayordomo, Pena, and Schwartz (2010) compare the five major databases of corporate CDS prices and show that the CMA, compared with other databases, lead the price discovery process. Therefore, we collect data of REIT CDSs mainly from the CMA database. The CMA CDS data is provided by DataStream before October 2010 and then is offered through Bloomberg and SNL. To construct a comprehensive and accurate data set of the REIT CDS, we map the weekly 5-year tenor CDS data into DataStream, Bloomberg, and SNL between Jan 2005 to Sep 2016. We collect weekly data instead of daily data because the trading volume of the REIT CDS market is limited and CDS quotes are not daily updated. Indeed, the REIT CDS spreads may not be perfectly synchronized with REITs on a daily basis. As the REIT CDS market is inertia, we set an arduous process to filter our CDS data from abnormal trading. Finally, we collect CDS data for 38 REITs.[2]

We create measurements for bank relationship following the process described by Lee et al. (2018). The Loan Pricing Corporation (LPC) Dealscan database provides detailed loan information as well as the data of lenders on each loan. We match each loan of a REITs to its creditors. After combining all active loans at the announcement of credit events, we count the number of distinct banks at the parent level and also aggregate the total amount of each active loans as the loan size. We devise two variables as the proxy for

---

[2] Collin-Duresne, Goldstein, and Martin (2001) find that the majority of variation in CDS market is explained by the liquidity premia. To reduce the influence of liquidity premia on credit measurement in REITs CDS market, we set comprehensive criteria to filter our data sample.

lending relationships. First, we consider the total number of supervision banks over syndicate loan, such as lead banks, administrative banks. Second, we count the number of banks which are principal players in the CDS markets. Identified by Lee et al. (2018), the list of such banks includes Bank of America, Barclays Banks, BNP Paribas, Citibank, Credit Suisse, Deutsche Bank, Goldman Sachs, JPMorgan Chase, Morgan Stanley, Royal Bank of Scotland.

Table (1) summarizes our data and variables from Jan 2005 to Sep 2016.

## 3. Methodology

### *Measuring CDS reactions*

Following Lee et al. (2017) and Hull et al. (2004), we adjust weekly CDS return by using the change in the same credit level portfolio to remove CDS market trends. According to the S&P credit rating tiers, we sort all REITs CDS into three portfolios by credit level of A, BBB, BB. And then we take the median spread within each credit rating portfolios to account for CDS systematic trends.

There are two distinct methods in previous studies to measure CDS return. One uses absolute changes of CDS spread, and another applies the percentage change of CDS spread. The CDS contract is similar to a credit protection insurance with a decided premium. Thus, the return to the buyer of such implicit insurance contract is the profit from changes in the implied premium. Such returns approximately measure the CDS performance. Thus, we label the CDS return as percentage changes of CDS spread in the following empirical research.

The adjusted spread return of CDS is the difference between CDS spread return and the according CDS market return, and the calculation is following as:

(1) $$Adjusted\ Spread\ Return(ASR)_{i,t} = \frac{(Spread_{i,t} - Spread_{i,t-1})}{Spread_{i,t-1}} - \frac{(Index_{r,t} - Index_{r,t-1})}{Index_{r,t}}$$

where $Spread_{i,t}$ is the weekly 5-year CDS spread of the REIT I at t week, $Index_{r,t}$ is the median spread within the portfolio with r credit rating at date t, and $ASR_{i,t}$ is the adjusted weekly CDS return. Then, we calculate the cumulative adjusted spread return for REIT i within event window $[t_1, t_2]$ as follows:

(2) $$Cumulative\ Adjusted\ Spread\ Return(CASR)_{i,[t_1,t_2]} = \sum_{t=t_1}^{t_2} ASR_{i,t}$$

*Measuring Bond reactions*

In theory, as CDS protects buyers against credit event, a portfolio consisting of a bond and a corresponding CDS will be considered a risk-free investment. Therefore, the 5-year CDS spread should be equal or close to the spread of synthetic 5-year bond issued by the reference REIT over the risk-free rate. However, there is an enduring debate about the proxy for the risk-free rate. Bond traders tend to consider Treasury yield as the risk-free rate, while derivatives traders regard swap zero curves as the risk-free zero curves. Longstall et al. (2003) and Lee et al. (2018) consider the Treasury rate as the benchmark risk-free rate. While, Duffee (1996) and Hull et al. (2004) argue that there are many factors, such as liquidity and taxation, make the Treasury yield depressed relative to the yields of low-risk corporate bonds. Choosing the swap rate as the risk-free rate, Blanco et al. (2003) document that CDS spreads are quite close to bond yield spreads. They also show that price discovery occurs in the CDS market in advance of that in the bond market. Hull et al. (2004) and Norden and Webber (2004) are also consistent with these findings.

We calculate the difference between CDS spread and bond spread over swap rate and the difference between CDS spread and bond spread over Treasury yield. Table (2) summarizes the statistics of such two spreads. Over 75% of CDS spreads are lower than according bond spreads, which also meets our previous assumptions. Furthermore, the bond spreads over swap rate remain closer to CDS spreads than that over Treasury yield. However, for some instance, the CDS spread is larger than the bond spread over the swap rate. According to Blanco, et al. (2005), three imperfections result in such an unusual situation. Repo costs in the bond market decrease the bond spread, the CTD option of CDS contracts escalates the CDS spread, and liquidity premia in the CDS market cause CDS movements contemporarily unrelated to default measurement. For subsequent research, we use the swap rate as the proxy for the risk-free rate to calculate the bond spread. To discovery the idiosyncratic bond spread in firm-specific, we adjust the bond spreads in excess of an equally-weighted index of all bond spread in the whole sample.

*Confounded events*

Following Lee et al. (2017) and Hull et al. (2004), we define a credit event as confounded events if another credit event follows a credit event during a half year window. As confounded events could bias the reactions of CDS to upcoming credit events, we exclude confounded events in our data sample.

## 4. Main Results

*CDS reactions to credit rating events*

Both credit ratings and CDS spreads describe the creditworthiness of bond issuers. In this section, we examine the reaction of CDS spreads to announcements of credit events.

We use the cumulative adjusted spread return (CASR) of CDS to measure the extent of the reaction for approaching credit events. Fig (1) presents the responses of CDS spreads to upcoming rating events within the event window from the prior nine weeks to post five weeks. In the Panel (A), we show the reactions of 50 credit events which include 24 downgrade events and 36 upgrade events. There are distinctly asymmetric reactions for downgrade events and upgrade events. For the forthcoming downgrade events, REITs CDS already reacts as early as nine weeks before such an event. However, for the upcoming positive events, REIT CDS is inertia and fails to predict the happening of such events. Even after the announcement of positive events, the REIT CDS market still shows a weak reaction for the events.

To further discover the prediction of CDS for credit events, we focus on the reaction of two special credit events. One is the upgrade rating from speculative level to investment level, and another is the downgrade rating from investment level to speculative level. Although there is only one notch between such two levels, the repercussions of such two special credit events are severe. The drop to speculative level can show that the REIT firm may quickly run into difficulties in paying debts. Moreover, according to the covenant of institution investors, some institutional investors cannot invest the firms with speculative credit level.

Panel (B) demonstrates that there are substantial responses of the REIT CDS to the downgrade events with credit notch across the investment-to-speculative level threshold. The reaction of CDS is weak and faint for upgrade events where the rating of REIT firms improves from speculative to investment grade.

We report the measurement of CDS reaction to each event in Table (3). We partition the whole event window into three sub-intervals and present the mean of CASR for each sub-interval. Meanwhile, we also separate the credit events into several subcategories according to the notch level of rating events (e.g., Investment grade, Speculative-grade, IG to SG, etc.).

In the panel of downgrade events, REITs with speculative credit rating are more sensitive to the upcoming credit event in the CDS market. In specific, in advance of downgrade events, the mean CASR of speculative-grade REITs CDS is 15.27%, while the mean CASR of REITs CDS with investment grade is just 4.41%. This material difference indicates that the information about deteriorating firms in the CDS market is more efficient and timelier. Furthermore, compared with the CASR for speculative level CDS during the prior-event period, the much lower CASR in the post-event window implies that, for the CDS of struggling REITs, CDS market significantly absorbs the information of upcoming negative credit events. In the particular cases where REITs are degraded from investment level to speculative level, CDS reactions for the approaching downgrade event are considerable with 17.92% CASR, which implies that the CDS market can predict the announcement of credit event under some particular situations.

The panel of upgrade events displays that REIT CDS responses to positive credit events are weak and faint. For the REITs that experience promotion from speculative grade to investment grade, the CASR of CDS before an announcement is positively significant. Such finding may confirm that when a firm starts with an investment grade, the demands of according debt increase and so do that of CDS. Generally, it is hard for the CDS market of REITs to predict the forthcoming upgrade events. Finally, we turn our attention to negative watch events. As the negative watch is a status that the credit-rating agencies are still considering the credit situation of a company, REITs CDS market is insensitive to such blurry information, which is confirmed by the limited CASR before the announcement.

However, after the pronouncement of the negative watch, there are significant CDS responses with 11.69% CASR. Such large reactions support the observation that downgrade events usually follow negative watches.

*The lead-lag relation between CDS and Bond market*

Since both CDS and bond spreads reflect default risk, in this part, we will discover whether the CDS market detects default-risk related information earlier than the bond market, or vice versa.

We first analyze the lead-lag movement between CDS and bond market at the aggregate level. A panel VAR model is proper to capture the co-movement within stationary variables in a simultaneous panel framework.[3] Love and Zicchino (2006) suggest that the normal ordinary lest square procedure with equation-by-equation estimation for VAR models can bias the valuation of panel VAR models. To provide unbiased estimates for panel VAR models, they establish the generalized method of moments (GMM) estimator and use lags of the endogenous variables as instruments to improves the efficiency and accuracy of estimation of panel VAR model. We follow Love and Zicchino (2006) to estimate the following VAR model[4]:

$$(3) \qquad \begin{pmatrix} \Delta S_{i,t}^{Bond} \\ \Delta S_{i,t}^{CDS} \end{pmatrix} = \begin{pmatrix} \beta_{0,i,Bond} \\ \beta_{0,i,CDS} \end{pmatrix} + \sum_{j=1}^{k} \begin{pmatrix} \beta_{j,Bond,Bond} & \beta_{j,Bond,CDS} \\ \beta_{j,CDS,Bond} & \beta_{j,CDS,CDS} \end{pmatrix} \begin{pmatrix} \Delta S_{i,t-j}^{Bond} \\ \Delta S_{i,t-j}^{CDS} \end{pmatrix} + \begin{pmatrix} \varepsilon_t^{Bond} \\ \varepsilon_t^{CDS} \end{pmatrix}$$

where $\Delta S_{i,t}^{CDS}$ is weekly changes in CDS spread, $\Delta S_{i,t}^{Bond}$ is weekly changes of a synthetic 5-year REIT bond spread over according swap rate in *t*, *j* is lag order indicator, and $\varepsilon_t$ is

[3] Our VAR model is similar as previous studies (Lee et al (2017), Hilscher et al (2015), Norden and Weber (2009)).
[4] To estimate the panel VAR, we use a STATA code provided by Abrigo and Love (2016).

disturbance term in *t*. The order of lag (*j*) in our VAR model is determined by the modified AIC, BIC, and QIC discussed by Andrews and Lu (2001). To validate the model, we also exert unit root test and stationarity test for each variable respectively. For a robust test, we also control for clustered errors to adjust for cross-sectional correlation and heteroskedasticity.

We report our panel VAR estimation in Table (4). In the left Panel, we use aggregate CDS and bond spread changes to evaluate the prior panel VAR model. We find that the coefficients of all four lags of CDS spread changes are significantly and economically positive to predict the contemporaneous bond spread changes. While, lagged bond spread changes are irrelevant to current CDS spread changes, which is indicated by the insignificant lagged CDS changes variables in the column of CDS spread changes. To further test the direction of the information channel between these two markets, we report the result of Granger-Causality test (Granger 1969) at the bottom. The Granger test applies the Wald test to check whether all lagged variables of the same reference in each equation of the VAR model are all equal to zero. Academia always uses such analysis to confirm the information flow in multivariate processes. The result of Granger-Causality test in prior model strengthens our finding that the CDS market exploits the credit-related information before the bond market.

To remove the influence of market trend on the information flow between CDS and bond market, we utilize the prior panel VAR model on the idiosyncratic spread changes which are adjusted by the market spread changes. The market spread changes are the equally-weighted average spread changes at same week t in each market. Panel B presents the similar result as Panel A does that all lag of idiosyncratic CDS changes are significantly positive in the column of idiosyncratic bond changes while lagged idiosyncratic bond changes cannot explain the contemporary idiosyncratic CDS changes. We reveal that there is a unidirectional information flow about credit risk from CDS to the bond market after controlling for the variation of the market.

Finally, we employ the VAR model between CDS spread change and bond spread change at the REIT-specific level. We utilize the Granger-causality test to indicate the direction of information flow. From Table (1), 9 of 37 cases display a unidirectional information flow from CDS to the bond market, 18 REITs discover bilateral information flow between these two markets, and only 3 REITs illustrate that the bond market unilaterally conveys information to the CDS market.

All previous results confirm that REITs CDS market dominates REITs bond market in detecting credit information, which is consisted to prior studies (Lee et al. (2018), Blanco et al. (2005) Webber and Norden (2004)). Aside by Blanco et al. (2005), we support that the different mechanism of two markets causes such finding. Due to the short-sales constraints in the bond market, it is hard and costly for an investor to hedge credit risk in the bond market. However, the mechanism of the CDS market makes it most accessible for institutional investors to trade credit risk in the CDS market. Moreover, participants in the CDS market are constrained to the large institutional investors with considerable capital size and high credit rating, which improve the information spillover in the CDS market more efficient. Thus, CDS market principally leads the bond market in the detection of credit information.

*The direction of information flows between CDS and stocks market*

We have exhibited that REIT CDS market dominates the bond market on discovering of credit-related information. In this part, we continue to examine the direction of information flow by considering whether or not REITs CDS spread returns lead REITs stock returns and vice versa.

Applying the similar panel VAR model as in previous researches (Lee et al. (2017), Hilscher et al. (2015), and Norden and Weber (2009)), we regress weekly REITs equity returns on contemporaneous and lagged REITs CDS spread return and meanwhile run the corresponding regressions for CDS spread returns. The detailed model is as following:

$$(4)\qquad \begin{pmatrix} R_{i,t}^{Equity} \\ R_{i,t}^{CDS} \end{pmatrix} = \begin{pmatrix} \beta_{0,i,Equity} \\ \beta_{0,i,CDS} \end{pmatrix} + \sum_{j=1}^{k} \begin{pmatrix} \beta_{j,Equity,Equity} & \beta_{j,Equity,CDS} \\ \beta_{j,CDS,Equity} & \beta_{j,CDS,CDS} \end{pmatrix} \begin{pmatrix} R_{i,t-j}^{Equity} \\ R_{i,t-j}^{CDS} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t}^{Equity} \\ \varepsilon_{t}^{CDS} \end{pmatrix}$$

where $R_{i,t}^{Equity}$ and $R_{i,t}^{CDS}$ are the weekly percentage returns of equity returns and CDS spreads returns for REIT *i* at week *t*. We also cluster the standard errors to control for multiple correlation and heteroscedasticity.

The estimates in Table (5) denotes that neither lagged equity returns or lagged CDS spread returns substantially reflect the information of future REITs returns. However, all three-lagged equity returns notably predict future CDS return even after controlling for the corresponding lagged returns of CDS. Specifically, the significant negative-signs of all lagged equity returns in the prediction of CDS returns are also rational because the reaction of the stock market to credit information should be opposite to that of CDS market to the same credit news. The larger magnitude of lagged stock returns for current CDS returns further confirms that the predictability of stock returns is much stronger than that of the autocorrelation of CDS return.

We further apply our VAR model on each RETIs to summarize the channel of information flow between the CDS and stock market. We measure the direction of information flow by the Granger causality test. The dominant finding is that lagged equity returns contain information for current CDS price changes, while the reverse is rarely the case. Exhibited by Table (1), for 30 REITs of the 40 reference entities, lagged REITs returns contain information of current CDS returns. Only 10 cases exhibit that lagged CDS returns are significant in explaining current stock returns. Such results are consistent with that from the previous pooled VAR. Our analysis of co-movement between the REIT CDS market and REIT equity market reconfirms the conclusion of previous studies that the CDS market is a sideshow to the equity market.5 Furthermore, the large transactions costs in the CDS market make such discrepancy of information discovery persist (Hilscher et al. 2015).

Focusing on the shock of credit rating news, a growing literature has shown that CDS market returns are predictive of stock returns (Acharya and Johnson (2007), Lee et al. (2018), Chava, Ganduri, and Ornthanalai (2016)). This predictability is more prominent for the events containing adverse credit information. Inspired by these previous studies, we further shed light on the importance of REITs CDS in information discovery of credit event. We first re-estimate the previous panel VAR model for lead-lag relation between CDS and stock around the announcement of rating events. Panel A in Table (6) presents the lead and lag relation between weekly CDS returns and stock returns during 18-weeks around each announcement of rating event. However, no finding convinces that the lagged CDS returns can predict current stock returns either before or after the rating event. During the pre-event period, the predictability of stock returns on future CDS returns is economical and substantial. Such predictability disappears after the announcement.

[5] Lee et al (2017) estimate similar panel VAR model and find that the lagged idiosyncratic CDS return will predict current idiosyncratic equity return. However, due to the limited size of our CDS sample, we cannot generate an accurate measurement of idiosyncratic CDS return to verify their finding.

Then, we narrow the attention of our rating event analysis and only focus on the downgrade rating events. The result is showed in Panel B. The significant coefficient of lagged CDS return in the prediction of forwarding stock return in the pre-event period strongly demonstrates that the REIT CDS market contributes to price discovery of negative rating events. The negative sign of lagged CDS returns also indicates that, before the public negative credit news, speculation on credit risk of the REIT CDS market may bring negative information into the stock market. Whereas, after the adverse credit news is available in the market, the effect of CDS on stock wanes, and price discovery no longer exists in the REIT CDS market.

In general, the results in Table (6) verify that REITs CDS market keeps exclusive information about the deteriorating credit of firm and such information is not entirely exposed to the stock market. Before the downgrade announcement, the REIT CDS market spill over unique and credit-related information to the stock market.

*Sources of information about rating events in REITs CDS market*

In this section, we will unveil the sources of unique information in REITs CDS market before credit events. Following Lee et al. (2018), who confirm the existing of unique information in CDSs market, we check whether REITs CDS market reveals proprietary information that can be explained by a relationship between REITs and banks. We first estimate the previous panel VAR model for CDS-equity market relation in [-9, -1] weeks prior to credit events on REITs with above median bank relation.[6]

[6] We use two measurements of bank-relationship, Leadbank depth and CDSbank depth, to estimate the panel VAR model, respectively. The results derived from such two measurements are similar. Thus, we only show the results of CDSbank depth in following analysis.

The Panel A in Table (7) demonstrates that, for REITs with stronger lending-relation with banks which are also major market makers in the CDS market, the lagged CDS spread returns significantly and reversely predict future equity return. The Granger-causality test further confirms that the CDS market conveys unique information to the stock market before the announcement of credit events. To deeply understand the dynamic shock of CDS return on equity return, we resort to impulse-response functions (IRF) and forecast-error variance decomposition (FEVD). As shown in Fig (2), the impulse response of CDS spread returns to equity return is significantly negative and transitory. To test whether the bank relationship contributes to the unique information in the CDS market about future rating events, we control for the loan size and bank relationship as exogenous variables in the previous co-movement model.

After we control the effect of bank relationship and loan size, the coefficient of lagged CDS returns turn into insignificant. Such outcomes further support our finding that the information revelation for credit news in the CDS market is related to the depth of bank dealer. As a comparison, we also estimate the previous panel VAR model on all REITs and REITs with less bank relation, respectively. Lagged CDS returns are irrelevant to contemporary equity returns, while lagged equity returns notably and negatively affect future CDS returns. Under such two controlled circumstances, the stock market predominantly leads the CDS market in price discovery of upcoming rating news. Our empirical research on co-movement between REIT CDS and REIT equity just before the approaching rating news favors the conclusion by Acharya and Johnson (2007) and Lee et al. (2017). Bank-related information advantage in CDS market advances information benefits from future credit events.

D'Auria, Foglia, and Reedtz (1999) point out that an increase in credit to an individual borrower will help banks grab more monopoly information about the creditor. Chava et al. (2017) also find that, for firms that have a relatively large size of bank loans, the reaction of the credit market to rating events is much stronger. Banks are constrained to the capital requirement that changes with the credit level of creditors. Consequently, banks propose tighter oversight on borrowers to which they extend more loans.

In this part, we further extend our information analysis to the influence of loan size. In Panel B of Table (7), we study the lead-lag relation during the [-9, -1] days prior to a rating event in the following three estimation groups: (1) REITs with larger loan size, (2) REITs with larger loan size but control for loan size, and (3) REITs with less loan size. In the first group, for REITs with larger loan size, it is noteworthy that both REITs and CDS returns significantly predict each other prior the credit events. Comparing the magnitude of lagged CDS in the prediction of stock return for REITs with larger bank-relation (-0.366) and REITs with larger loan size (-0.587), we notice that the loan size contributes to stronger information flow from the CDS to equity market. Similar as in Panel A, after controlling for the loan size in the estimation, the predictability of CDS vanishes, and there is no longer information flow from CDS to the stock market. The insignificant coefficient of lagged CDS returns in the third group further supports our argument that, prior to the rating news, the exclusive information from the CDS market to the stock market is stemmed from the loan amount.

*Co-movement among stock, CDS, and bond market*

Having shown the pairwise lead-lag relations between bonds and CDS as well as between CDS and equities, in this section, we jointly examine the co-movement among these three markets. The three-way panel VAR model is similar to the two-way panel VAR model we use in previous sections. The returns of REITs CDS and bond are the weekly percentage change in CDS spread and bond yield spreads, respectively. Table (8) demonstrates the simultaneous co-movement among these three markets. The lagged equity returns have a significantly negative impact on CDS returns and bond returns. Moreover, after concurrently controlling for the stock trading information, CDS trading no longer contributes to price discovery for bonds.

## 5. Conclusion

Focusing on the REITs CDS market from 2005 to 2016, we provide evidence that the CDS market produces unique information spillovers to other REIT security markets before the negative credit news. We also find that CDS spreads significantly react in advance of future downgrade rating events. Moreover, the CDS market persistently dominate the bond market in uncovering credit risk information. We also document that the CDS market contributes significantly to price discovery when firm-specific credit information is deteriorating. Both banking relationship and loan size improve the information revelation in the CDS markets, which implies that large banks active in the CDS market will exploit private information obtained through their direct lending relationships, which are arguably stronger and updated more frequently than in most other industries. Overall, we conclude that the CDS market appears to be the primary market for trading on REIT credit risk information.

# REFERENCES


Abrigo, M.R. and Love, I., 2015. Estimation of panel vector autoregression in Stata: A package of programs. *manuscript, Febr 2015 available on http://paneldataconference2015. ceu. hu/Program/Michael-Abrigo. pdf*.

Acharya, V.V. and Johnson, T.C., 2007. Insider trading in credit derivatives. *Journal of Financial Economics*, *84*(1), pp.110-141.

Bessembinder, H., Kahle, K.M., Maxwell, W.F. and Xu, D., 2008. Measuring abnormal bond performance. *The Review of Financial Studies*, *22*(10), pp.4219-4258.

Blanco, R., Brennan, S. and Marsh, I.W., 2005. An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *The journal of Finance*, *60*(5), pp.2255- 2281.

Bolton, P. and Oehmke, M., 2011. Credit default swaps and the empty creditor problem. The Review of Financial Studies, 24(8), pp.2617-2655.

Chava, S., Ganduri, R. and Ornthanalai, C., 2017. Are credit ratings still relevant?

Collin-Dufresn, P., Goldstein, R.S. and Martin, J.S., 2001. The determinants of credit spread changes. *The Journal of Finance*, *56*(6), pp.2177-2207.

D'Auria, C., Foglia, A. and Reedtz, P.M., 1999. Bank interest rates and credit relationships in Italy. *Journal of Banking & Finance*, *23*(7), pp.1067-1093.

Dick-Nielsen, J., 2014. How to clean enhanced TRACE data.

Duffee, G.R., 1996. *Treasury yields and corporate bond yield spreads: An empirical analysis*. Washington: Federal Reserve Board.

Fecht, F., Hackethal, A. and Karabulut, Y., 2018. Is proprietary trading detrimental to retail investors? *The Journal of Finance*, *73*(3), pp.1323-1361.

Financial Crisis Inquiry Commission and United States. Financial Crisis Inquiry Commission, 2011. *The financial crisis inquiry report, authorized edition: Final report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*. Public Affairs.

Granger, C.W., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pp.424-438.

Hasan, I. and Wu, D., 2017. Credit default swaps and bank loan sales: Evidence from bank syndicated lending.

Hilscher, J., Pollet, J.M. and Wilson, M., 2015. Are credit default swaps a sideshow?

Evidence that information flows from equity to CDS markets. Journal of Financial and Quantitative Analysis, 50(3), pp.543-567.

Hull, J., Predescu, M. and White, A., 2004. The relationship between credit default swap spreads, bond yields, and credit rating announcements. *Journal of Banking & Finance*, *28*(11), pp.2789- 2811.

Ivashina, V. and Sun, Z., 2011. Institutional stock trading on loan market information. *Journal of financial Economics*, *100*(2), pp.284-303.

Kryzanowski, L., Perrakis, S. and Zhong, R., 2017. Price discovery in equity and CDS markets. *Journal of Financial Markets*, *35*, pp.21-46.

Marsh, I. and Wagner, W., 2012. Why is price discovery in credit default swap markets news- specific?

Mayordomo, S., Pena, J.I. and Schwartz, E.S., 2014. Are all credit default swap databases equal? *European Financial Management*, *20*(4), pp.677-713.

Lee, J., Naranjo, A. and Velioglu, G., 2018. When do CDS spreads lead? Rating events, private entities, and firm-specific information flows. Journal of Financial Economics.

Longstaff, F.A., Mithal, S. and Neis, E., 2005. Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance*, *60*(5), pp.2213-2253.

Loon, Y.C. and Zhong, Z.K., 2016. Does Dodd-Frank affect OTC transaction costs and liquidity? Evidence from real-time CDS trade reports. Journal of Financial Economics, 119(3), pp.645-672.

Love, I. and Zicchino, L., 2006. Financial development and dynamic investment behavior: Evidence from panel VAR. *The Quarterly Review of Economics and Finance*, *46*(2), pp.190-210.

Narayan, P.K., Sharma, S.S. and Thuraisamy, K.S., 2014. An analysis of price discovery from panel data models of CDS and equity returns. Journal of Banking & Finance, 41, pp.167-177.

Norden, L., 2017. Information in CDS spreads. *Journal of Banking & Finance*, *75*, pp.118-135.

Norden, L. and Wagner, W., 2008. Credit derivatives and loan pricing. *Journal of banking & finance*, *32*(12), pp.2560-2569.

Norden, L. and Weber, M., 2009. The co-movement of credit default swap, bond and stock markets: An empirical analysis. European financial management, 15(3), pp.529-562.

Oehmke, M. and Zawadowski, A., 2015. Synthetic or real? The equilibrium effects of credit default swaps on bond markets. *The Review of Financial Studies*, *28*(12), pp.3303-3337.

Reiser, A., 2010. Should Insider Trading in Credit Default Swap Markets be Regulated-The Landmark Significance of SEC v. Rorech. *DePaul Bus. & Comm. LJ*, *9*, p.531.

Shan, C., Tang, D.Y. and Yan, H., 2017. Credit default swaps and bank regulatory capital.

Xiang, V., Chng, M. and Fang, V., 2013. Transmigration across price discovery categories: evidence from the US CDS and equity markets. *Journal of Futures Markets*, *33*(6), pp.573-599.

Table 2.1: The Summary of Data

| | Institution Name | Sticker | leaderbank | cdsbank | Loan ($Million) | Credit Rating | cds->equity | equity->cds | cds->bond | bond->cds | observation | CDS period |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AMB Property Corporation | AMB | 0.6 | 1.1 | 226.1 | 4.0 | N | N | | | 177 | Jan 05 - Jun 07; Oct 09 - Sep 10 |
| 2 | American Tower Corporation (REIT) | AMT | 9.8 | 10.0 | 7959.7 | 4.3 | N | N | Y | N | 245 | Jan 12 - Sep 16 |
| 3 | AvalonBay Communities, Inc. | AVB | 5.9 | 8.0 | 1337.6 | 3.9 | N | Y | Y | Y | 610 | Jan 05 - Sep 16 |
| 4 | Brandywine Realty Trust | BDN | 6.3 | 7.3 | 975.2 | 4.0 | Y | Y | Y | N | 373 | July 07 - Sep 16 |
| 5 | BRE Properties, Inc. | BRE | 5.8 | 7.0 | 1348.1 | 4.0 | N | Y | Y | N | 450 | Jan 05 - Jun 07; Jun 08 - Mar 14 |
| 6 | Boston Properties, Inc. | BXP | 5.2 | 5.6 | 934.9 | 3.2 | N | Y | Y | Y | 609 | Jan 05 - Sep 16 |
| 7 | Crown Castle International Corp. (REIT) | CCI | 11.3 | 6.0 | 11642.3 | 4.6 | N | Y | Y | Y | 140 | Jan 14 - Sep 16 |
| 8 | Mack-Cali Realty Corporation | CLI | 4.3 | 6.0 | 1285.0 | 4.2 | Y | Y | N | Y | 609 | Jan 05 - Sep 16 |
| 9 | Camden Property Trust | CPT | 6.2 | 8.0 | 1012.8 | 4.0 | N | Y | Y | Y | 444 | Mar 08 - Sep 16 |
| 10 | DDR Corp. | DDR | 9.2 | 10.0 | 2781.9 | 4.4 | Y | Y | Y | Y | 609 | Jan 05 - Sep 16 |
| 11 | Duke Realty Corporation | DRE | 3.7 | 3.5 | 843.7 | 4.0 | N | Y | Y | Y | 608 | Jan 05 - Sep 16 |
| 12 | Equity Residential | EQR | 5.8 | 3.6 | 1370.1 | 3.7 | N | N | Y | Y | 610 | Jan 05 - Sep 16 |
| 13 | Equity One, Inc. | EQY | 5.4 | 6.0 | 807.4 | 4.0 | Y | Y | N | N | 349 | Jan 10 - Sep 16 |
| 14 | FelCor Lodging Trust Incorporated | FCH | 4.2 | 9.0 | 453.1 | 5.8 | N | Y | N | N | 585 | Jan 05 - Mar 16 |
| 15 | First Industrial Realty Trust, Inc. | FR | 4.8 | 6.0 | 1057.5 | 4.8 | Y | Y | Y | Y | 556 | Jan 06 - Sep 16 |
| 16 | Federal Realty Investment Trust | FRT | 6.9 | 6.0 | 891.6 | 3.6 | Y | Y | Y | Y | 453 | Jan 08 - Sep 16 |
| 17 | Welltower Inc. | HCN | 7.7 | 10.0 | 3225.9 | 4.0 | Y | Y | Y | Y | 570 | Oct 05 - Sep 16 |
| 18 | HCP, Inc. | HCP | 6.4 | 9.0 | 3660.6 | 4.0 | Y | Y | Y | N | 450 | Feb 08 - Sep 16 |
| 19 | Highwoods Properties, Inc. | HIW | 6.6 | 4.0 | 1182.9 | 4.0 | N | Y | N | N | 466 | Jan 05 - Dec 06; Oct 09 - Sep 16 |
| 20 | Hospitality Properties Trust | HPT | 5.7 | 4.5 | 1373.9 | 4.0 | Y | Y | Y | Y | 454 | Jan 08 - Sep 16 |
| 21 | Healthcare Realty Trust Incorporated | HR | 5.2 | 7.0 | 1146.8 | 4.0 | N | N | Y | N | 474 | Jan 05 - Jun 07; Feb 10 - Sep 16 |
| 22 | Host Hotels & Resorts, Inc. | HST | 5.6 | 5.5 | 1178.3 | 5.1 | N | Y | Y | N | 610 | Jan 05 - Sep 16 |
| 23 | Kimco Realty Corporation | KIM | 9.4 | 9.0 | 2881.6 | 3.7 | N | Y | Y | Y | 609 | Jan 05 - Sep 16 |
| 24 | Liberty Property Trust | LPT | 5.5 | 7.4 | 920.0 | 4.0 | N | Y | Y | Y | 441 | 0v 05 - Dec 06; Apr 09 - Sep 16 |
| 25 | Nationwide Health Properties, Inc. | NHP | 4.3 | 4.5 | 593.0 | 4.0 | N | N | N | Y | 244 | Jan 05 - Jul 07; Jan 09 - Apr 11 |
| 26 | Prologis, Inc. | PLD | 10.7 | 13.0 | 6090.7 | 4.0 | N | Y | Y | N | 610 | Jan 05 - Sep 16 |
| 27 | Post Properties, Inc. | PPS | 5.1 | 2.7 | 722.4 | 4.0 | Y | Y | Y | Y | 321 | Jul 10 - Sep 16 |
| 28 | Regency Centers Corporation | REG | 6.6 | 5.0 | 1665.5 | 4.0 | N | Y | N | Y | 254 | Jan 10 - Jun 11; May 13 - Sep 16 |
| 29 | Rayonier Inc. | RYN | 5.0 | 6.0 | 250.0 | 4.0 | N | N | Y | N | 137 | Jan 08 - Aug 10 |
| 30 | SL Green Realty Corp. | SLG | 7.6 | 9.0 | 3365.7 | 4.9 | N | Y | N | N | 333 | Apr 10 - Sep 16 |
| 31 | Simon Property Group, Inc. | SPG | 14.3 | 12.0 | 7080.6 | 3.1 | Y | Y | Y | Y | 609 | Jan 05 - Sep 16 |
| 32 | iStar Inc. | STAR | 4.8 | 12.0 | 4912.0 | 5.3 | N | Y | Y | Y | 602 | Feb 05 - Sep 16 |
| 33 | UDR, Inc. | UDR | 6.3 | 8.0 | 1627.9 | 4.0 | N | Y | Y | Y | 572 | Oct 05 - Sep 16 |
| 34 | Vornado Realty Trust | VNO | 7.0 | 12.2 | 2732.3 | 4.0 | N | Y | Y | Y | 608 | Jan 05 - Sep 16 |
| 35 | Ventas, Inc. | VTR | 9.9 | 10.0 | 4262.5 | 4.0 | N | N | N | N | 242 | Feb 12 - Sep 16 |
| 36 | Washington Real Estate Investment Trust | WRE | 4.8 | 7.0 | 684.7 | 4.0 | N | Y | N | N | 453 | Jan 08 - Sep 16 |
| 37 | Weingarten Realty Investors | WRI | 6.2 | 5.0 | 728.6 | 3.8 | N | Y | Y | N | 589 | Jun 05 - Sep 16 |
| 38 | Weyerhaeuser Company | WY | 7.4 | 7.0 | 2080.8 | 4.0 | N | Y | N | N | 351 | Jan 10 - Sep 16 |
| | Max Value | | 21.0 | 13.0 | 19050.0 | 7.0 | | | | | | |
| | Min Value | | 0.0 | 0.0 | 0.0 | 3.0 | | | | | | |
| | Mean across firms | | 6.5 | 7.5 | 2174.8 | 4.1 | | | | | | |
| | Median across firms | | 6.0 | 7.0 | 1275.0 | 4.0 | | | | | | |
| | STD | | 3.6 | 3.1 | 2440.8 | 0.7 | | | | | | |

Table 2.2: The Summary of Spread Difference

| | swapdif | treasurydif |
|---|---|---|
| mean | -0.41 | -0.68 |
| std | 1.23 | 1.25 |
| min | -27.29 | -27.30 |
| 25% | -0.68 | -0.97 |
| 50% | -0.26 | -0.53 |
| 75% | 0.01 | -0.21 |
| max | 28.81 | 27.84 |

*Note*: The difference between CDS spread and bond spread over swap rate and the difference between CDS spread and bond spread over Treasury yield. This table summarizes the statistics of such spread differences.

Table 2.3: The Reaction of CDS Spread to Credit Event

| | Num | [-9~-1] | [-1~1] | [1~5] |
|---|---|---|---|---|
| *Panel A. Downgrade* | | | | |
| | Num | [-9~-1] | [-1~1] | [1~5] |
| ALL | 24 | 6.67% | 1.30% | 9.17% |
| Investment | 19 | 4.41% | 0.49% | 10.15% |
| Speculative | 5 | 15.27% | 4.37% | 5.48% |
| IG to SG | 4 | 17.92% | 8.00% | 9.78% |
| *Panel B. Upgrade* | | | | |
| | Num | [-9~-1] | [-1~1] | [1~5] |
| ALL | 36 | -1.28% | -1.54% | 0.84% |
| Investment | 22 | -1.95% | -1.50% | 2.25% |
| Speculative | 14 | -0.25% | -1.60% | -1.33% |
| SG to IG | 5 | 5.80% | -2.03% | 1.34% |
| *Pancel C. Negtive Watch* | | | | |
| | Num | [-9~-1] | [-1~1] | [1~5] |
| ALL | 6 | -1.71% | 2.60% | 11.69% |

*Note* : This table reports the measurement of CDS reaction to downgrade, upgrade, and negtive watch credit event.We partition the whole event window into three sub-intervals and present the mean of CASR for each sub-interval. Meanwhile, we also separate the credit events into several subcategories according to the notch level of rating events.

Table 2.4: The Panel VAR Model for CDS and Bond Market

| | | | | | | |
|---|---|---|---|---|---|---|
| *Panel A. Panel VAR Model* | | | | | | |
| | Model 1. Aggregate Difference | | | Model 2. Idiosyncratic Difference | | |
| | | Bond Spread Dif | CDS Dif | | Bond Spread Dif | CDS Dif |
| Bond Spread Dif | t-1 | -0.584 | 0.006 | t-1 | -0.593 | 0.001 |
| | P-value | 0.00 | 0.24 | P-value | 0.00 | 0.73 |
| | t-2 | -0.391 | -0.003 | t-2 | -0.404 | -0.001 |
| | P-value | 0.00 | 0.61 | P-value | 0.00 | 0.79 |
| | t-3 | -0.212 | -0.004 | t-3 | -0.238 | -0.001 |
| | P-value | 0.00 | 0.53 | P-value | 0.00 | 0.95 |
| | t-4 | -0.106 | 0.000 | t-4 | -0.131 | -0.001 |
| | P-value | 0.00 | 0.98 | P-value | 0.00 | 0.73 |
| CDS Dif | t-1 | 0.542 | 0.094 | t-1 | 0.388 | 0.021 |
| | P-value | 0.00 | 0.00 | P-value | 0.00 | 0.42 |
| | t-2 | 0.423 | 0.117 | t-2 | 0.248 | -0.002 |
| | P-value | 0.00 | 0.00 | P-value | 0.00 | 0.76 |
| | t-3 | 0.285 | 0.024 | t-3 | 0.275 | 0.034 |
| | P-value | 0.00 | 0.31 | P-value | 0.00 | 0.13 |
| | t-4 | 0.256 | -0.012 | t-4 | 0.219 | -0.022 |
| | P-value | 0.00 | 0.59 | P-value | 0.00 | 0.52 |
| *Panel B. Causality Test* | | | | | | |
| G-Causality | | Bond | CDS | | Bond | CDS |
| | Bond | | N | Bond | | N |
| | CDS | Y | | CDS | Y | |

Table 2.5: The Panel VAR for CDS and Equity Market

| | | | |
|---|---|---|---|
| *Panel A. Panel VAR Model* | | | |
| | | Equity Return | CDS Return |
| Equity Return | t-1 | -0.045 | -0.247 |
| | P-value | 0.07 | 0.00 |
| | t-2 | 0.024 | -0.122 |
| | P-value | 0.33 | 0.00 |
| | t-3 | 0.032 | -0.052 |
| | P-value | 0.15 | 0.01 |
| CDS Return | t-1 | 0.007 | -0.013 |
| | P-value | 0.34 | 0.22 |
| | t-2 | -0.004 | 0.059 |
| | P-value | 0.64 | 0.01 |
| | t-3 | 0.017 | 0.036 |
| | P-value | 0.01 | 0.07 |
| *Panel B. Causality Test* | | | |
| G-Causality | | Equity Return | CDS Return |
| | Equity Return | | Y |
| | CDS Return | Y | |

*Note*: this tables shows the result of panel VAR between REIT CDS returns and REIT equity returns. We find that that neither lagged equity returns or lagged CDS spread returns substantially reflect the information of future REITs returns. However, all three-lagged equity returns notably predict future CDS return even after controlling for the corresponding lagged returns of CDS.

Table 2.6: The Panel VAR for CDS and Equity Market Before and After Credit Events

| | | | | | |
|---|---|---|---|---|---|
| *Panel A. All Events (n=66)* | | | | | |
| | | Equity return (*t*) | | CDS return (*t*) | |
| | Interval | [-9,-1] | [1,9] | [-9,-1] | [1,9] |
| Equity return | *t-1* | -0.073 | -0.285 | -0.122 | 0.030 |
| | P-value | 0.38 | 0.04 | 0.01 | 0.70 |
| CDS return | *t-1* | -0.096 | 0.047 | 0.061 | -0.139 |
| | P-value | 0.14 | 0.56 | 0.48 | 0.14 |
| *Panel B. Downgrade Events (n=24)* | | | | | |
| | | Equity return (*t*) | | CDS return (*t*) | |
| | Interval | [-9,-1] | [1,9] | [-9,-1] | [1,9] |
| Equity return | *t-1* | -0.061 | -0.290 | -0.130 | 0.004 |
| | P-value | 0.50 | 0.02 | 0.01 | 0.96 |
| CDS return | *t-1* | -0.177 | 0.169 | 0.061 | -0.037 |
| | P-value | 0.05 | 0.12 | 0.59 | 0.76 |

*Note:* this table estimates the previous panel VAR model for lead-lag relation between CDS and stock around the announcement of rating events.Panel A in Table (6) presents the lead and lag relation between weekly CDS returns and stock returns during 18-weeks around each announcement of rating event. Panel B narrows the attention of our rating event analysis and only focus on the downgrade rating events.

Table 2.7: The Source of Unqiue Information in The CDS Market

*Panel A: Effect of Bank-Relation Before Credit Event [-9,-1]*

| | | Stock return (*t*) | | | CDS return (*t*) | | |
|---|---|---|---|---|---|---|---|
| | | Above median bank-relation | Above median bank-relation (with exog var) | Below median bank-relation | Above median bank-relation | Above median bank-relation (with exog var) | Below median bank-relation |
| stock return | *t-1* | -0.114 | -0.015 | -0.186 | -0.125 | -0.054 | -0.363 |
| | P-value | 0.21 | 0.87 | 0.02 | 0.00 | 0.33 | 0.00 |
| CDS return | *t-1* | -0.366 | 0.149 | 0.051 | 0.103 | 0.531 | 0.043 |
| | P-value | 0.01 | 0.48 | 0.34 | 0.35 | 0.00 | 0.67 |

*Panel B: Effect of Loan Size before Credit Event [-9,-1]*

| | | Stock return (*t*) | | | CDS return (*t*) | | |
|---|---|---|---|---|---|---|---|
| | | Above median loan size | Above median loan size (with exog var) | Below median loan size | Above median loan size | Above median loan size (with exog var) | Below median loan size |
| stock return | *t-1* | -0.026 | -0.044 | -0.243 | -0.076 | -0.068 | -0.267 |
| | P-value | 0.44 | 0.78 | 0.02 | 0.02 | 0.67 | 0.00 |
| CDS return | *t-1* | -0.587 | -0.605 | -0.044 | 0.281 | 0.325 | 0.032 |
| | P-value | 0.00 | 0.47 | 0.58 | 0.03 | 0.73 | 0.74 |

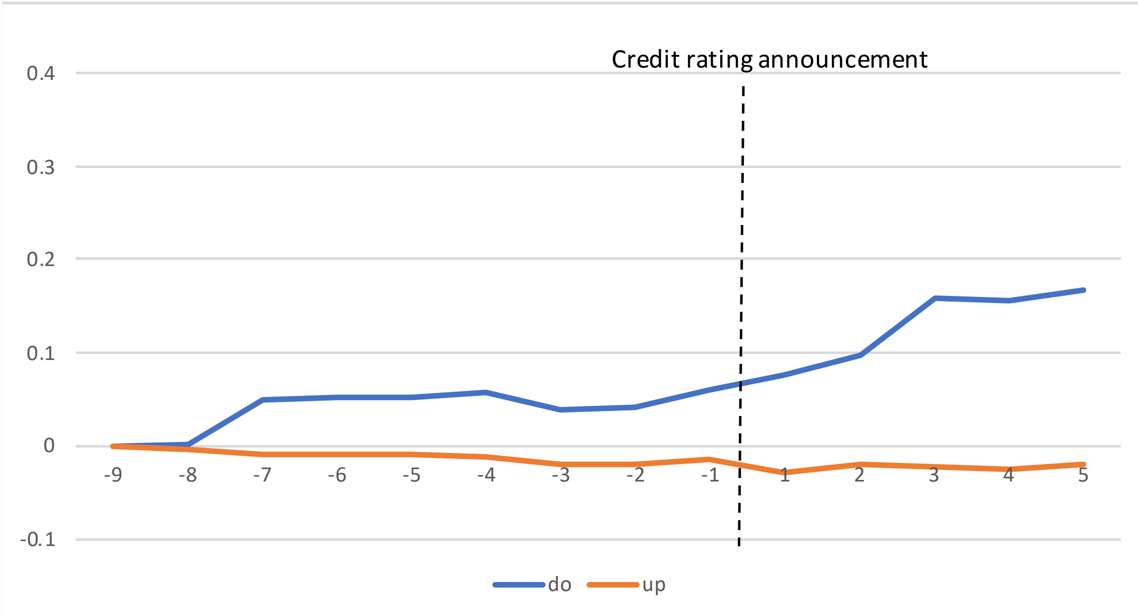
*Note*: this table checks whether REITs CDS market reveals proprietary information that can be explained by a relationship between REITs and banks. We first estimate the previous panel VAR model for CDS-equity market relation in [-9, -1] weeks prior to credit events on REITs with above median bank relation.

Table 2.8: The Panel VAR Model among Bond, CDS, and Equity Marekt

| | | | | |
|---|---|---|---|---|
| *Panel A. Panel VAR Model* | | | | |
| | | Equity chg | CDS chg | Bond chg |
| Equity chg | L1 | -0.058 | -0.195 | -1.538 |
| | P-value | 0.07 | 0.00 | 0.06 |
| | L2 | -0.039 | 0.075 | -2.081 |
| | P-value | 0.42 | 0.45 | 0.06 |
| CDS chg | L1 | -0.036 | 0.089 | -1.142 |
| | P-value | 0.22 | 0.07 | 0.15 |
| | L2 | -0.027 | 0.223 | -0.437 |
| | P-value | 0.05 | 0.01 | 0.18 |
| Bond chg | L1 | 0.000 | 0.000 | 0.000 |
| | P-value | 0.00 | 0.00 | 0.37 |
| | L2 | 0.000 | 0.000 | 0.000 |
| | P-value | 0.00 | 0.00 | 0.65 |
| *Panel B. Causality Test* | | | | |
| G-Causality | | Equity chg | CDS chg | Bond chg |
| | Equity chg | | Y | Y |
| | CDS chg | N | | N |
| | Bond chg | Y | Y | |

*Note*: this table demonstrates the simultaneous co-movement among these three markets. The lagged equity returns have a significantly negative impact on CDS returns and bond returns. Moreover, after concurrently controlling for the stock trading information, CDS trading no longer contributes to price discovery for bonds.

*Panel A. All Credit Events*




*Panel B. Investment-to-Speculative Credit Events*

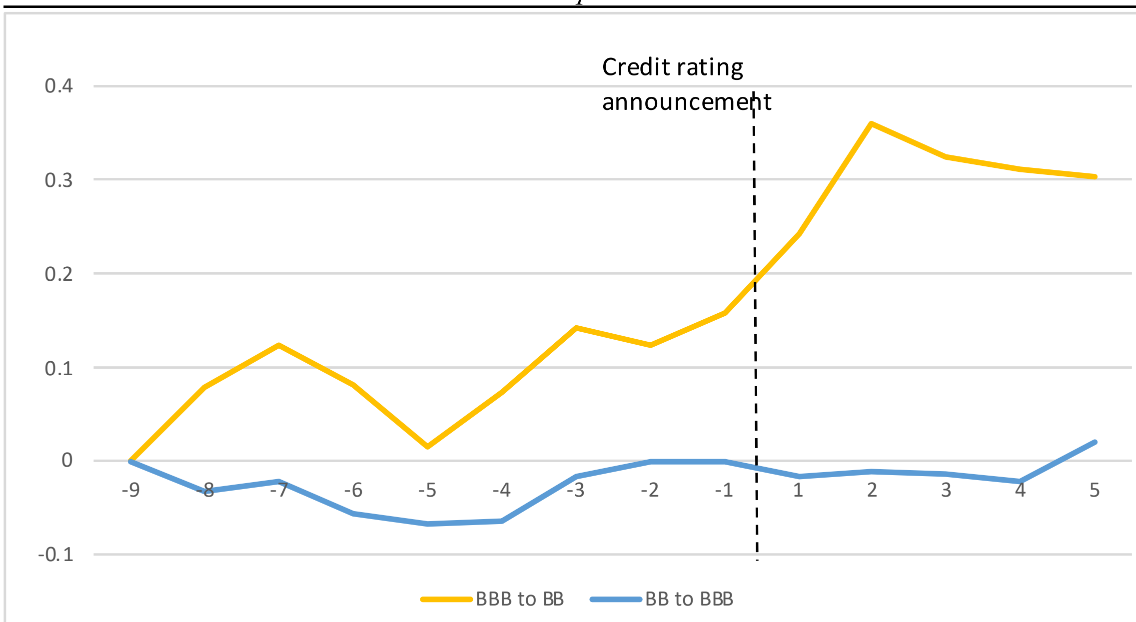



*Figure 2.1. The Responses of CDS Spreads to Upcoming Rating Events.* This figure presents the responses of CDS spreads to upcoming rating events within the event window from the prior nine weeks to post five weeks.

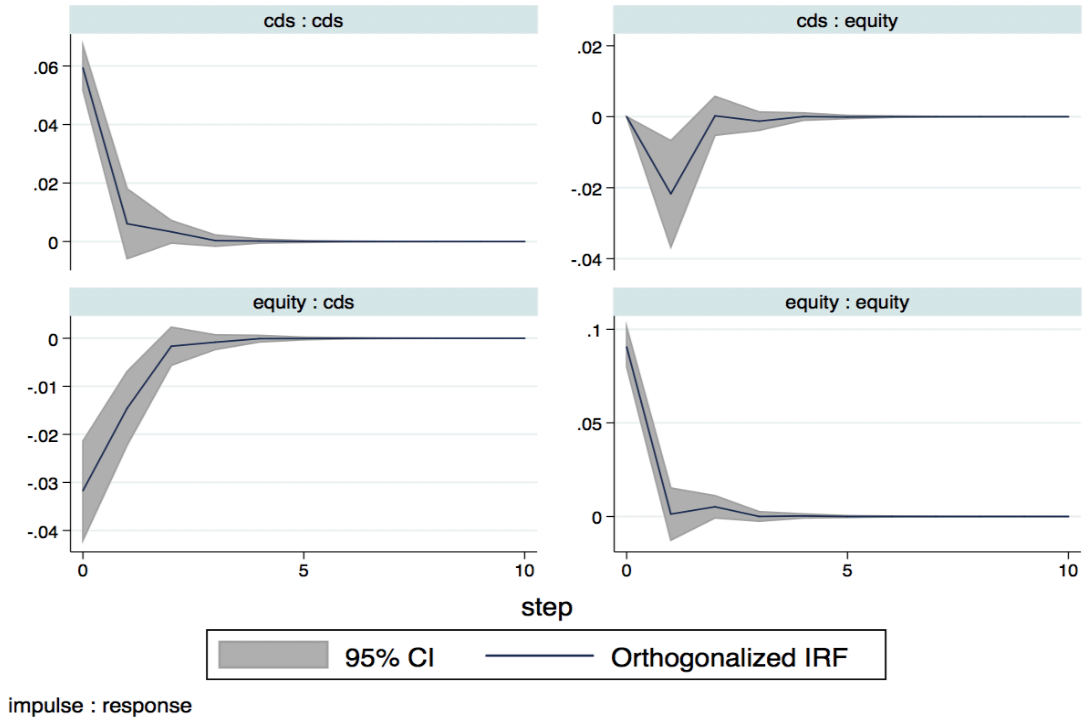



Figure 2.2. Impulse-Response Functions (IRF). This figure shows the impulse response of CDS spread returns to equity return is significantly negative and transitory.

# CHAPTER 3: THE PREDICTABILITY OF REIT INDEX RETURNS: MACHINE LEARNING APPROACHES

## 1. Introduction

Ling, Naranjo, and Ryngaert (2000) construct over 15 features, including macroeconomic variables and financial ratios, to predict the excess returns for equity REITs. Although they show a decent fitness level of the multi-factor model for in-sample prediction, they fail to forecast the out-of-sample return accurately. Especially, their active-trading strategies based on out-of-sample expectations cannot overwhelm buy-and-hold strategies to produce substantial profits. Since Ling et al. (2000), the discussion about REITs prediction already attracts a lot of researchers' attention[1]. Researchers tend to devise more complicated predictive model and factors to explain cross-sectional REITs returns. However, Cochrane (2011) points out that stock return prediction is one of the most challenging studies in current asset-pricing research.

Due to the noise trend and volatile features, how to accurately forecast and explain the REITs returns, especially in out-of-sample data, is still a hot and open topic. In this paper, we devise 10 machine learning and deep learning models to predict the REITs index return, including Kalman filter (KF), the least absolute shrinkage and selection operator (LASSO), classification and regression tree (CART), random forest (RF), adaptive boosting tree (AdaBoost), gradient boosting (GBRT), eXtreme gradient boosting tree

[1] See the details in Carmichael and Coen (2018), Pavlov, Steiner, and Wachter (2015), Hansz, Zhang, and Zhou (2017).

(XGBoost), support vector machine (SVM), stacked auto-encoders (SAE), and long short-term memory (LSTM).

With an astonishing development in algorithm and computer power, machine learning and deep learning models have made significant achievements in the fields of image recognition, data analysis, and natural language process. Recently, there is an exciting trend in the literature that more and more academia focuses on the application of machine learning and deep learning in the financial industry. To our best known, this is the first paper that sheds light on the implementation of machine learning and deep learning models on REIT index return.

*The advantage of machine learning and deep learning*

Compared with the traditional regression model for stock prediction, such as ordinary least squares (OLS) and autoregressive-moving-average (ARMA) models, Gu, Kelly, and Xiu (2018) summarize three advantages of machine learning models. First, machine learning competently handles multi-dimensional data which have various data types in the financial market. A rich mixture of quantitative and qualitative information continuously affects the fluctuation of the equity market. Inspired by the arbitrage pricing theory (APT) which was introduced by Ross (1976), recent financial literature demonstrates a comprehensive set of over a hundred characteristics which contribute to predicting stock return [2]. Such extensive dimensional features set will cause multicollinearity and contradict the assumption of traditional econometric models.

---

[2] Green, Hand, and Zhang (2017) list of 94 firm factors to forecast stock change. Harvey, Liu and Zhu (2016) build 316 characteristics to explain stock returns.

However, the feature extraction and regularization methods in machine learning will overcome the challenge of multicollinearity in multi-factor models.

Second, machine learning models are not restricted by the linearity assumption. The volatile and unstable market situations will consistently change the relation between predictors and stock returns, which finally twists the linearity relationship. To diagnose the linearity among our data sample, we plot residuals of OLS regression versus predicted values in the below Figure (1).

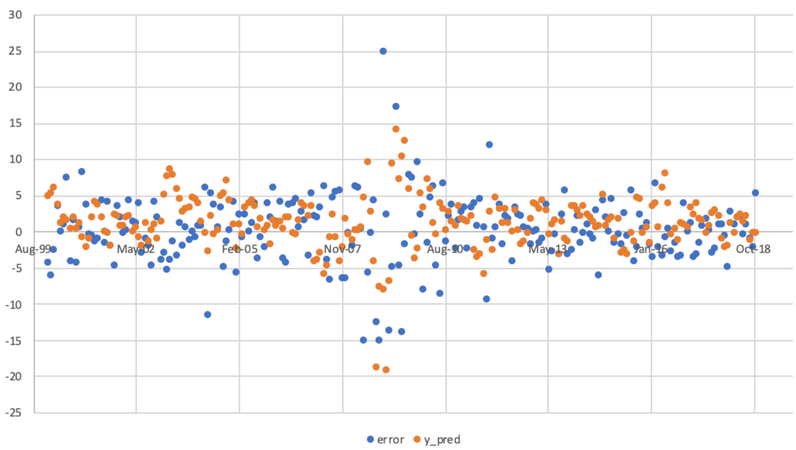



Figure 3.1: The Test of Linearity Relationship

To relieve the multicollinearity problem, we combine OLS with the recursive feature elimination with cross-validation (RFECV)[3]. The RFECV is a feature selection method that consecutively removes the weakest predictors until the specified number of features is reached. The figure shows there is no perfect linearity relation between REITs index returns and picked features because the points in figure scatter asymmetrically around the horizontal line.

[3] We also apply principal component analysis (PAC) to extract feature information and then resort to OLS model. The result of this process is similar as that of RFECV-OLS.

Third, machine learning is famous for its versatile models and robust methods, which helps us to discover the information in REITs index from diverse angles. By setting proper cost functions, we can measure the performance among these models and then compare the predictive power. Moreover, with the outcome of some machine learning models, we gauge the information for return prediction in our 40 predictors and then identify the most critical features for REITs index return in different market statues.

*The unique features of the REITs Index*

Since the efficient market hypothesis came out to the public (Fama 1965), which asserts that stock returns should be unpredictable, there has been a constant battle against the predictability of equity market. However, recent studies exhibit that the developments of complicated time-series models and machine learning models make it possible to discover future market information.[4] Ferson and Harvey (1991) further suggest that, compared with individual stocks, the market index is more predictable because idiosyncratic risks are diversified.

Due to the stable cash flows derived from the generally long-term leases and tangible firm assets in the real estate industry, REITs returns are believed to be more predictable than other stock returns. Serrano and Hoesli (2008) utilize an ARMA-EGARCH model to test the predictability of REITs and conclude that in-sample returns are more predictable for REITs than for stocks. This finding is also confirmed by Liu and Mei (1992) and Liao and Mei (1998). Inspired by all these previous studies, we use REITs

[4] Atsalakis and Valavanis (2009) survey more than 100 published papers that apply soft computing techniques to successfully forecast stock market movement.

index to examine the predictive power of machine learning and deep learning on market returns.

*Main empirical findings and contribution*

Extending the findings by Ling et al. (2000) who employ rolling OLS method to predict future REITs index returns, we test the predictive power of 10 machine learning and deep learning models in this paper. In a nutshell, all these machine learning algorithms overwhelmingly outperform the traditional linear regression on the out-of-sample prediction performance. The root-mean-square error (RMSE) of OLS is 16 times as large as that of machine learning in our whole sample period. In terms of cost function's output, there is no significant difference between our machine learning models. Kalman filter and random forest are the top two best predictive methods. To further study the predictability of different REITs market status, we divide the data sample into three subsets: pre-crisis (Jan 2000 – July 2008), financial crisis (Aug 2008 – Dec 2009), and post-crisis (Jan 2010 – Nov 2018). Due to the extreme market volatility and complicated market situation, it is not a surprise that the predictive powers of all methods deteriorate during the financial crisis. However, compared with linear regression, which barely explains any future REITs market direction, machine learning still describes parts of information about future returns.

Moreover, the decision tree, which is vulnerable to overfitting problem, works better than many other methods in the financial crisis. The sizeable positive autocorrelation of REITs returns in this fluctuant market status contributes to the decent performance of the decision tree. This finding suggests that, to precisely predict future market movement, we should adjust various machine learning methods to unique characteristics of distinctive market situations.

To include comprehensive predictive signals, we select a comprehensive variable set with 14 macroeconomic factors, ten real estate market indicators, eight equity market features, and 9 REIT index characters. By ranking the feature importance which is derived from the tree-architecture algorithm, we detect the variables with the reliable predictive power of the REIT market. Various signals should decide the movement of a relatively stationary REITs market. The REITs-level characteristics contain critical prediction information in such REITs market. However, the value of the REIT index is dominated by the real estate market indicator during the financial crisis. As the crush of real estate market initially causes this crisis, the REIT market closely tracks the performance of the real estate market.

To exploit the economic significance of predicted returns by machine learning, we employ a practical and straightforward investment strategy. The one-month ahead return prediction of machine learning models with a threshold level, which decreases the turnover cost and the noise of prediction error, decides the long/short trading position. Although this strategy hardly beat the buy-and-hold investment in both pre- and post-crisis period, when there is a table increasing-trend in the REITs market, this investment overwhelms the buy-and-hold strategy during the financial crisis and the investment value increases by 70% in this extremely volatile market. This marked achievement justifies the role of machine learning and deep learning in the asset pricing and investment strategy.

In section 2, we will review previous literature and summaries the application of machine learning in return prediction and the real estate industry. Section 3 describes our data set and data splitting schemes for different models. Section 4 briefly explains the theoretical model and algorithm behind machine learning methods in this paper. We show the empirical results in section 5, and the last section offers some concluding comments.

## 2. Literature review

*The application of machine learning and deep learning in stock prediction*

Gu et al. (2018) recently conduct an eye-catching analysis about the predictability of machine learning and deep learning models, including linear regularization models, PCA, random forests, and multi-layer neural networks. Compared with the traditional econometric return-model, machine learning models, in this paper, significantly improve the accuracy of stock return prediction. Specifically, they find that random forests model and neural networks model demonstrate the most forecasting power for the stock return. Additionally, they consider a vast feature set with over 900 baseline signals and finally agree that there is a small set of imperative predictors for return prediction, including return volatility, liquidity, and return momentum. Lastly, they construct a portfolio trading strategy, which is built on the prediction of machine learning models, to testify the economic value of machine learning and deep learning models. Compared with the Sharpe ratio of buy-and-hold investment, the Sharpe ratio of their strategy almost doubles. Inspired by Gu et al. (2018), we expand the number of machine learning models in our paper and generate a more comprehensive conclusion.

Lv, Huang, Li, and Xiang (2019) further study 12 popular machine learning algorithms to generate stock trading strategies in 9 different industries sectors. Their experiment shows that, for different industry sectors, there is no standard best-predictive model. Due to the various theoretical algorithms behind machine learning models, we should update our models setting to different data characters. They also consider a sequence of investment performance evaluation indicators, such as sharp ratio and winning ratio, and then generate a non-parametric test to measure the accuracy of these indicators.

Finally, they denote that the sharp ratio and winning ratio are the best two gauges that evaluate the prediction performance of machine learning models. However, they ignore the REIT market which is a principal component in the capital market.

In the last five years, there is a noticeable trend that machine learning and deep learning methods have sprung up in stock prediction literature. Han, He, Rapach, and Zhou (2018) apply LASSO and elastic net, which are powerful machine learning tools, to retest the predictability of the sizeable signal set in Green et al. (2017). They finally extract 30 critical factors from the 94 firm characteristics for predicting cross-sectional returns. Chinco, Clark-Joseph, and Ye (2018) also apply the LASSO to make rolling return forecasts in the one-minute horizon. Ballings, Poel, Hespeels, and Gryp (2015) evaluate the return predictability of ensemble methods and single classifier models, which are two principal categories of the machine learning algorithm. They reveal that RF is the top algorithm to explain the information about stock price direction.

Furthermore, they emphasize the application of ensembles methods in stock prediction, which combines the results of all predictive models. In this paper, we also follow the implication of ensembles methods to create our investment strategy. Patel, Shah, Thakkar, and Kotecha (2015) combine SVM with RF to build a two-stage scheme to forecast the stock market index changes. They propose that trend-deterministic data can contribute to the predictive power of machine learning models. XGBoost has become a favorite tool in the data science industry, as it is a highly flexible and versatile tool that works on large-scale data for prediction problems. Basak, Kar, Saha, Khaidem, Dey (2019) research on the performance of XGBoost on stock return prediction and demonstrate that XGBoost outperforms other predictive models.

SVM and KF are also widely used to discover forecasting information in the equity market. The masterpieces proposed by Tay and Cao (2001) is the first paper that compares the feasibility of SVM with a simple neural network algorithm in financial time series forecasting. The smaller error measurement in this paper proves that SVM overwhelms a simple neural network for financial market prediction. Afterward, Huang, Nakamori, and Wang (2005) further demonstrate that it is advantageous to use SVM in financial market forecasting. Elliott, Hoek, and Malcolm (2005) propose a novel KF with a mean-reverting smooth model to calibrate the hedging ratio in pairs trading strategy.

Due to the rise in computing power and the development in the backpropagation algorithm, the applications of deep learning methods in the capital market have been attracting massive attention. Although there are versatile neural network algorithms, LSTM, which is devised by Hochreiter and Schmidhuber (1997), is considered as the best deep learning model focusing on financial time series forecasting. The unique design for storing and updating historical information makes LSTM powerful in sequence prediction problems. Cheng, Huang, and Wu (2018) apply LSTM to evaluate the forecasting ability of movement direction of the equity market. They argue that LSTM significantly improves the reliability of prediction in the equity market, which is also confirmed by Chen, Zhou, and Dai (2015). To alleviate the noise information in financial market data, Bao, Yue, and Rao (2017) construct a two-stage deep learning algorithm. They first utilize SAE to extract information from a comprehensive feature set. Then, they feed the high-level features from the first stage into LSTM to forecast the future equity index. Compared with other similar models, such as the recurrent neural network (RNN), this novel algorithm develops both predictive accuracy and profitability performance. In this paper, we follow the idea of Bao et al. (2017) to build deep learning models.

*The application of machine learning and deep learning in real estate industry*

Although the machine learning skills have infused virtually every sector of our economy, the technology has yet to gain a noteworthy foothold in the real estate industry. However, the developments of machine learning and deep learning applications in the real estate industry are astonishing. Today, machine learning, artificial intelligence, and natural language processing are currently and widely used in the real estate industry.

In the last two years, real estate literature has sporadically utilized machine learning skills. Mullainathan and Spiess (2017) present a hands-on approach on how to apply machine learning methods to analyze the real estate value information from numerical house characteristics. Shen (2018) utilize natural language processing algorithms to extract soft information from real estate advertisements. She uncovers that such unobservable information notably captures the information about real estate value. Lindenthal and Johnson (2018) train a deep convolutional neural network (CNN) on images from Google Street View to classify residential buildings styles and then calculate the housing price premium for each architectural vintage.

Besides valuing the real estate, machine learning and deep learning are also widely used to predict mortgage default. Sirignano, Sadbwani, and Giesecke (2018) develop a multi-layer neural network on almost 300 signal variables to assess mortgage risk. Kvamme, Sellereite, Aas, and Sjursen (2018) resort to CNN when they explore the mortgage default information among consumer transaction data. Deep learning methods are proved to be competent for the delinquency models which mainly identify the complex interaction between default risk and various predictors.

## 3. Data

Inspired by Ling et al. (2000), we focus on the predictability of the monthly NAREIT equity index return from January 2000 to November 2018. [5] Figure (2) describes the trend of REITs index during our data period. Although there is a significant and sharp dip during the financial crisis, the index raises gradually and finally increase by almost seven times. Table (1) summarizes descriptive statistics for monthly REITs index returns. In general, the fluctuations of REITs market are substantial. Notably, the volatility of REITs during the financial crisis period is over three times as large as that in the pre- or post-crisis period. Furthermore, the significantly distinct autocorrelations of REITs return in these three periods also imply that it will be challenging to predict out-of-sample returns.

The first important step in building a prediction model is to select predictors of REITs index. To summarize the excellent prediction information of REITs industry, we consider a comprehensive list of features, including the economic factors, the equity market conditions, the signals of technical analysis, and the characteristics of REITs index. Table (2) shows the detailed description of these predictors.[6]

As an essential component of the whole economy, the real estate closely reflects the macroeconomic status. To extract the forecasting information from macroeconomic signals, we follow previous studies and build an extensive collection of economy predictors. This list contains the yield of the 10-year treasury, the curve of treasury yield, and the default spread which is the difference between Aaa and Baa corporate bond yield. We use the one-month lag of these yield rate variables to predict the future of the REIT index

[5] The REITs index return are obtained from NAREIT website.
[6] Except for features in the table, we also considered the predicted value of ARMA, mortgage rate, and other technique indicators. However, due to the existence of multicollinearity, we finally delete these features.

direction. To measure and predict the economic status, we also consider other 11 macroeconomic predictors, including PMI index, unemployment rate, leading economic index, manufacturers' new orders, monetary base, consumer sentiment, CPI index, PPI index, capacity utilization, retail sales, and dollar index. Due to the noise and delay-reporting effects in historical macroeconomic data, which were noticed by Ling et al. (2000), we measure the percentage changes of these features between month $t - 8$ and month $t - 2$ to forecast the REITs index return in month t.

Additionally, we also consider 10 top economic indicators which closely track the real estate market trends. These include total construction spending, Case-Shiller Price Index, FHFA Price Index, NAHB Housing Index, building permits, median sales price, house sold, housing starts, supply of houses, and rental vacancy rate. Except for the rental vacancy rate which is quarterly data, all other are monthly data and the changes of these indicators from month $t - 8$ to month $t - 2$ are predictors of month t REITs index. For the rental vacancy rate, we apply cubic interpolation to convert this data to monthly data and then measure the percentage returns.

Besides the predictive information in the macroeconomy, the fluctuations of the equity market also notably affect the performance of REITs market. We comprise eight equity market indicators to demonstrate information from the capital market, including Fama-French three factors, market momentum factor, margin account balance, VIX, dividend yield of S&P 500, and the PE ratio of S&P 500. All market features are measured with a one-month lag.

Lastly, there are eight characters of REITs index in our paper that are projected to measure the future changes of REITs market. Like equity, the REIT price reflects short- and long-term trading patterns. However, these trading activities are non-stationary and uncertain. To address this challenge, we propose discrete Fourier Transform to capture the short- and long-term prediction of REITs index over time. Moreover, we employ two dummy variables to test the January effect and the influence of the financial crisis. We also list two indicators, MACD and EMA, which are favorite tools in technical analysis. Finally, both dividend yield and momentum are crucial factors to price REITs equity.

4. Sample splitting

Due to the unique data-structure requirement of the various predictive models in our paper, we construct three sample splitting schemes. As the sequential Kalman filter predicts results from the consecutive cycles of prediction and filtering, we cannot split our data sample but input whole sample into the model, and then treat the state estimation as the one-month ahead forecast of REIT return.

For machine learning models in our paper, we split the data into training, validation and testing sets, which is a standard and certified data splitting scheme in machine learning fields.[7] The training set is used to train the model and determine the parameters. After then, we tune hyperparameters and update optimal model settings in the validation set. Finally, we apply the optimal model to predict testing data and measure the predictability of REITs returns. However, we should keep the data structure of time series data when we evaluate our model and tune the hyperparameters of machine learning models. Following Bao et al. (2017), we use previous the 36 months data to train the model, and then search the optimal hyperparameters in the next consecutive three months data sample. After that, we apply the optimal model settings and re-train the model with both training and validation data to

[7] We also compare the predictive power of our data splitting scheme with that of K-folds grid search scheme which shuffle the data structure for hyperparameter tune. We find that the RMSE of our scheme is less than that of K-folds method. Our findings confirm that k-fold cross validation does no work for time series data.

determine the parameters of the optimal model. Finally, we employ this model on the next three-months testing data sample and measure the performance of model prediction. As a rolling scheme, we update our three sub-samples gradually forward in time to include more recent data. This below Figure (3) describes the process of data splitting for machine learning models.

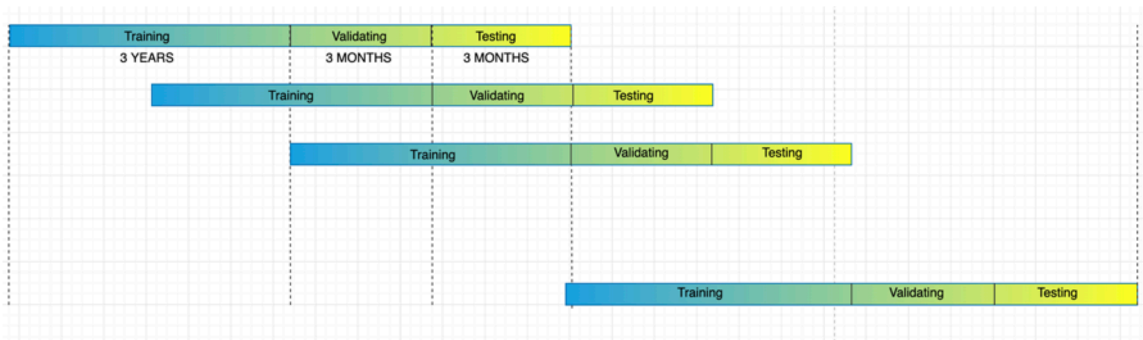



Figure 3.3: The Process of Data Splitting for Machine Learning Models

For the LSTM and Autoencoder model, we are inspired by Gu et al, (2018) and revise the previous data splitting scheme. Due to the complex networks of these two deep learning models, we should train the models in a huge data sample to decide the optimal model sittings. Moreover, the unique network construction of LSTM makes it adapt to the long-term historical data sample. To keep an extensive historical training data set for LSTM and Autoencoder model, we hold the total historical training and validation samples when we update our three split samples in the second split scheme. We illustrate the third split scheme in the following Figure (4).

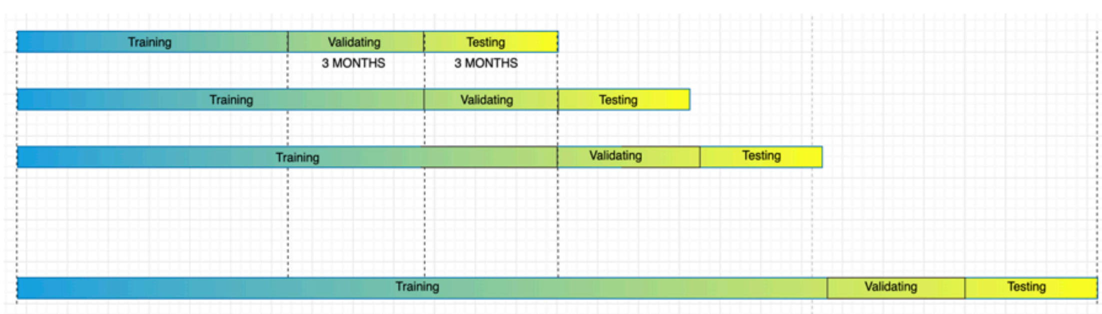



Figure 3.4: The Process of Data Splitting for LSTM Models

## 5. Methodology

In this part, we will briefly discuss the theoretical models behind all the predictive methods and our optimal settings.

### *Kalman Filter*

The Kalman Filter is a prediction-correction process. The filter comprises two function models: the state-transition model and the observation model. As a recursive estimator, the current estimated state is updated by the current measurement and previous estimate error.

In this paper, we construct a sample Kalman Filter and set the monthly REITs index return as both state and observation variable. The estimated state variable is treated as the one-month ahead return prediction.

### *Lasso Regression*

The Lasso regression is a regularization term to the cost function of Linear Regression. The cost function that this regularization tries to minimize is:

$$\sum_{i=1}^{n}(y_i - \sum_{j} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p}|\beta_j|$$

where $\lambda$ is a tuning hypermeter to control the strength of Lasso penalty. An important feature of Lasso is that it tends to completely eliminate the effect of the least important features on the dependent variable.

*Decision Tree and Random Forest*

Decision trees are versatile machine learning algorithms that can be used in both classification and regression models. It is also the fundamental groundwork of Random Forests. The goal of this algorithm is to minimize:

$$\sum_{m=1}^{M} \sum_{i \in R_m} \left(\hat{y} - y^{(i)}\right)^2 \text{ where } \hat{y} = \frac{1}{m} \sum_{i \in R_m} y^{(i)}$$

To decrease the overfitting of decision tree algorithm, we employ the grid-search hyperparameter tuning method to determine the optimal model settings for each testing data sample.

The random forest model is an ensemble of multiple decision trees, which is trained by the bagging algorithm. Instead of searching for the best feature when partitioning a node, which is the algorithm of decision trees, random forests select the optimal predictors only from a random subset of features. Once all random trees are trained, the prediction is the simple average of outputs by all tree regressions. The random subsets of determents make random forests trade a higher bias for a lower variance, which generally yields a better model for out-of-sample prediction.

*AdaBoost, Gradient Boosting and XGBoost*

All these three algorithms belong to the boosting category which applies ensemble methods to combine multiple weak learners into a strong learner. AdaBoost significantly boosts the performance of decision trees. To correct the previous prediction, AdaBoost adjusts the relative weight of wrongly-predicted training instance and then update weights in second regression, and so on. The algorithm stops when the desired number of predictors is reached or when the optimal model is detected. The prediction of AdaBoost is a weighted sum of all predicted values.

Similar to AdaBoost, Gradient Boosting also sequentially correct the fitness of previous predictors. However, instead of changing the relative weight of instances at each iteration, Gradient Boost tries to explain the residual errors from the previous regression model. Gradient Boosting also notably decreases the variance error of decision trees. Chen and Guestrin (2016) first propose the XGBoost algorithm. After the debut, this algorithm soon gained much popularity and has been the algorithm choice of many winning solutions in machine learning competitions. XGBoost is an optimization of Gradient Boosting framework and tends to be highly efficient, flexible and portable. In this paper, after trial and error, we find that the combination of DART booster, a booster method with dropout techniques, and XGBoost produces the most reliable predictive power. [8] In the following part, we only consider the empirical result of XGBoost with DART booster method.

*Support Vector Machines*

Similar to decision trees, SVM is a robust machine learning algorithm and capable of performing both non-/linear regression and classification tasks. As explained by Géron (2017), SVM regression tries to fit as many instances as possible on the street while limiting margin violations. To work on the nonlinear regression, we can employ a kernelized SVM model. Comparing the fitness of SVM with and without kernel algorithm, we show that the SVM with RBF kernel method explains the most substantial information about the moving direction of REIT index. This finding further confirms the nonlinear

[8] We compare the RMSE of three booster methods with XGBoost, including gbtree, gblinear, and dart. The prediction error of dart booster is the least.

relation between REITs return and various features. We only show the performance of SVM with RBF kernel in below part.

*Long Short-Term Memory*

Hochreiter and Schmidhuber (1997) propose the LSTM which is a development of the recurrent neural network. LSTM is considered as the best model for time-series data because it is designed to store necessary information and discard redundant information in the long-term state. The output information of LSTM is extracted from the current state, short-term state, and long-term state. As a type of RNN, the layer structure of LSTM will meet the assumption and requirement of RNN. The critical feature of RNN is that it will utilize the sequential information of a time series data.

In this paper, we apply the regularization and dropout method to decrease the overfitting risk of LSTM. Following Bao et al. (2017), the number of hidden layers and delays are set to 4 and 5 by trial and error.[9]

*Autoencoder*

As an unsupervised deep learning model, autoencoder is a powerful feature detector and acts as a dimensionality reduction. With a "sandwich" structure of multiple neural network, the stacked autoencoder extracts indispensable feature information by the coding layer. We represent an example of architecture of stacked autoencoder in the below Figure (4).

---

[9] The performance of only LSTM and the combination of autoencoder and LSTM is similar for our data sample. We only show the result of the combination of autoencoder and LSTM in this paper.

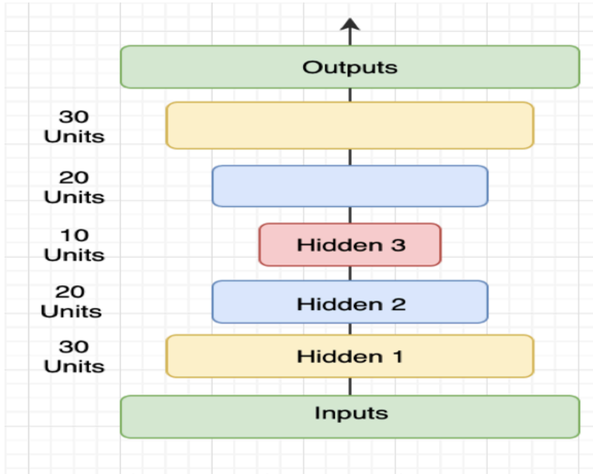



Figure 3.5: The Architecture of Stacked Autoencoder

The number of hidden layers decides the efficiency and quality of feature extraction. After trial and error, we set the depth of our autoencoder into 5 and the number of cells in coding layer is 10. This setting is also recommended by Bao et al. (2017).

6. Performance of out-of-sample prediction

In this paper, we mainly focus on the predictive power of various models on the REITs index return. To set a benchmark measurement, we replicate the linear regression with the same independent factors in Ling et al. (2000). The sample splitting of this benchmark regression will follow our second data splitting scheme, which means we use previous 24 months REITs returns to determine the linear regression and then predict the next three month returns via this model, after that, we will update the training and testing sets as a rolling window.

Moreover, inspired by the idea of the ensemble learning algorithm, we aggregate the prediction value of all machine learning and deep learning models at month t, and then take the average prediction value as the prediction of ensemble method. Géron (2017) explain the predictive power of ensemble algorithm because even if each predictor is a weaker algorithm, the ensemble of these poor predictors can still be a strong predictor if there are enough number of weak models and they are sufficiently diverse. Finally, in this paper, we compare the prediction performance among 11 distinct models.

We consider the error between real REITs index returns and out-of-sample predicted returns as the measurement of predictive power of models. Table (3) summarizes the statistics of prediction performance among all models. All number is the percentage expression. The RMSE represents the square root to the quadratic mean of the differences between predicted values and observed values, which is presented by the following formula:

$$\mathrm{RMSE} = \sqrt[2]{\mathrm{E}((\hat{y} - y)^2)}.$$

The mean absolute percentage error (MAPE) is another loss function for machine learning regression tasks to detect the prediction accuracy. MAPE is defined by the formula:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y-\hat{y}}{y}\right|.$$

Due to the quadratic item in RMSE, RMSE is much more sensitive to outliers than MAPE.

In general, the prediction performances of all machine learning models overwhelm that of linear regression. For the whole data sample, the mean of prediction difference by linear regression is almost four times as large as that of machine learning models. The MAPEs of machine learning models are less than half that of linear regression. To our surprise, there is a massive gap between linear regression and machine learning based on the variance and RMSE measurement. This extraordinary situation implies that linear regression is unable to adapt to an extreme market situation as machine learning is. Among these ten machine learning methods, it is hard to identify the best model with the most reliable predictive power, which is caused by the different measurement formula of various loss functions. Generally, Kalman filter, random forest, XGBoost and ensemble algorithm explain more future return information than other models. The higher levels of RMSE and variance from LASSO, decision tree, and SVR further denote that these three algorithms are less sensitive to an extremely volatile equity market.

To study the performance of models in different market situations. We divide the whole data sample into three subsets: pre-crisis, financial crisis, and post-crisis period. As we show in Table (1), the REITs market experiences severe fluctuation and volatility during the financial crisis and the risk almost triples compared to the other two periods. It is noticeable that the predictive power of linear regression drastically deteriorates during the financial crisis. The extreme values of maximum error and RMSE means linear regression barely explain the information in REITs market, which further suggests that the substantial market fluctuations ruin the linear relation. The results of previous linearity test support this implication. Although the predictive power machine learning is still more convinced than that of linear regression during pre- and post-crisis, the difference is not as substantial as in financial crisis.

Among machine learning algorithms, it is impossible to name one model which dominates in all three REIT market situations based on all cost functions. For relatively stable REITs markets, the performances of out-of-sample prediction are almost similar for all machine learning models except for the decision tree. However, during the financial crisis, the accuracy of prediction by decision tree is much better than that of some other algorithms. The decision tree is the top three best predictive model in the financial crisis in terms of RMSE value. The distinct autocorrelation of returns determines the prediction performance of the decision tree in different market status. Due to the fully-grown architecture, decision tree trades the variance for bias and tends to be overfitting. Table (1) shows that the autocorrelation of REITs returns during the financial crisis is positive, and the absolute value is much larger than that in the pre-crisis period. This positive and substantial autocorrelation suggests that the characteristics of the training data set are similar to that of testing data set, which will decrease the variance and improve the prediction of a decision tree with less bias.

In this section, we first confirm that the out-of-sample prediction by machine learning is much better than that of learning regression. Moreover, although the performances of all machine learning algorithm are comparable, no one model dominates all different market situations. Lastly, due to the significant effect of an extremely volatile market on the predictive power of different algorithms, it is crucial to detect the ominous sign of market crash earlier and decide the proper models.

## 7. Feature importance

When a decision tree grows and splits a node, the tree will search for the most important feature which provides the strongest explanatory power. As a result, the more important features are, the closer to the root of the tree features appear. As a byproduct of the tree-structure algorithm, we can assess features' importance and rank this feature importance in each predictor model. In this paper, there are five tree-structure algorithms: decision tree, random forest, AdaBoost, Gradient Boost, and XGBoost. To measure the feature importance, we aggregate the results of all five models in each training-test set and consider the mean as the measurement of feature importance.

Figure (6) demonstrates the feature importance for the top 10 predictors in four different periods. As the sum of feature importance of a given model is equal to one, the value of each factor importance explains the relative importance for the specific model. The left top chart overall describes the rankings of feature importance for the whole data period. The most influential factor is the short-term variance trend which is extracted from the Fourier Transform. This predictor is a REIT index-level feature and measures short-term trends of REIT index. The second key feature is the dividend yield of REITs index, which is also a REITs index-level predictor. Due to the dividend requirement for REITs, the dividend yield plays a pivotal role in pricing the REITs values. The third-ranked feature is the FHFA house price index, which reflects the overall real estate market status. However, no characteristics significantly dominate the information about future REIT returns. The importance value of the top key feature is less than 0.1, which indicates that we should consider a comprehensive list of predictors to forecast the REITs index returns. Two Fama-French factors, mktrf and smb, also enter into the top 10 critical features, which confirms the small-cap features of REITs market. To our surprise, MACD and ema, two technical analysis indicators, are in the list of top essential factors, which may support the validation of technical analysis on REITs market.

Although various predictors explain the REIT return information in a relatively stable market status, such information is dominated by the real estate market performance during the financial crisis. The left bottom figure reports the feature importance during the financial crisis. The top feature is fhfa, the value of which is 0.21 and the second important factor is the balance of securities margin account whose value is just 0.07. This finding verifies that the REITs market fully reflect the real estate market crash in the financial crisis. There are only two REITs-level features that are the top 10 crucial features for predicting REITs return during the financial crisis. The rank of feature importance concludes that the REITs market is overwhelmed by the whole economy and real estate market status and it is not the safe-haven for the equity market.

## 8. Performance of investment strategy

In this part, we design a novel investment strategy of REITs index to exploit machine learning forecasts directly. At the end of each month, we collect the one-month ahead REITs index return predictions of all algorithms. Due to the considerable turnover trading cost and variance of the predicted value, we set threshold value as 1.5% for adjusting holding position of REITs index. For example, if the predicted return for month t is 1% or -1%, as the absolute value is less than 1.5%, we will keep our trading position of month t-1, no matter whether we long or short the REITs index in month t-1. While, if the forecasting return is 2% or -2%, the absolute value of which is larger than 1.5%, we will change our trading position as long the REITs index if the prediction is positive and short the REITs index if there is negative predicted return.

Table (3) shows the performance of this investment strategy. The number in the table measures how large the investment changes. For example, the first value is 3.45, which means investment value increases by 3.45 times as compared with the beginning value. Generally, only the strategy derived from ensemble algorithm overwhelmingly beats the passive strategy. The profit by ensemble algorithm is twice as much as that of a buy-and-hold strategy. The ensemble learning, which combines the output of several models, will produce better predictive performance compared to a single model because it decreases both bias and variance.

We further analyze the investment performance in our three data subsets. During both pre- and post-crisis periods, buy-and-hold is the best trading strategy in terms of both total return and sharp ratio. This demonstration also supports a controversial dispute that alternative trading strategies cannot persistently beat the market in a long-term investment horizon. However, in the financial crisis, the market experiences unprecedented volatility. The ensemble algorithm deeply extracts predictive information from a volatile market and works better when REIT market changes more complicated.

## 9. Conclusion

To test the predictability of REITs index market, we perform a comparative analysis of 10 machine learning and deep learning models. Compared with the linear regression method which Ling et al. (2000) employ to predict the out-of-sample REITs returns, all the machine learning methods in this paper significantly improve the prediction performance. Notably, during the financial crisis, the complicated market status destroys the linear relation between REITs returns and explanatory factors, which makes linear regression barely reflect any future return information. Although the predictive power of machine learning also drops in the financial crisis, these algorithms still explain some movement of REITs returns. By analyzing the rank of feature importance in distinctive market situations, we find that REIT index-level features fundamentally contribute to detecting return information, and the future market movement should be predicted by a comprehensive feature set in a stable market. However, during the financial crisis, the REITs market is dominated by a real estate market signal and the predictive power of REITs-level features diminishes. Finally, we devise an investment strategy by the prediction of machine learning methods. Generally, it is hard for this strategy to continually beat the passive strategy in a long-term investment horizon, especially when the REIT market is relatively stable. However, the ability of machine learning to capture information from nonlinear and complex relation makes this strategy notably outstand during the financial crisis. In the whole data sample, the profit of ensemble learning strategy is twice as much as that of buy-and-hold investment.

In this paper, we confirm the advantages of machine learning and deep learning methods on pricing and predicting the REIT market. Compared with traditional linear regression, the burgeoning development of machine learning has achieved an excellent success for REIT index return prediction.

# REFERENCES

Ling, David C., Andy Naranjo, and Michael D. Ryngaert. "The predictability of equity REIT returns: time variation and economic significance." *The journal of real estate finance and economics* 20, no. 2 (2000): 117-136.

Carmichael, Benoît, and Alain Coën. "Real estate and consumption growth as common risk factors in asset pricing models." *Real Estate Economics* 46, no. 4 (2018): 936-970.

Pavlov, Andrey, Eva Steiner, and Susan Wachter. "Macroeconomic risk factors and the role of mispriced credit in the returns from international real estate securities." *Real Estate Economics* 43, no. 1 (2015): 241-270.

Cochrane, John H. "Presidential address: Discount rates." *The Journal of finance* 66, no. 4 (2011): 1047-1108.

Hansz, J. Andrew, Ying Zhang, and Tingyu Zhou. "An investigation into the substitutability of equity and mortgage REITs in real estate portfolios." *The Journal of Real Estate Finance and Economics* 54, no. 3 (2017): 338-364.

Ross, Stephen A. "The arbitrage theory of capital asset pricing." In *Handbook of the Fundamentals of Financial Decision Making: Part I*, pp. 11-30. 2013.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. *Empirical asset pricing via machine learning*. No. w25398. National Bureau of Economic Research, 2018.

Green, J., Hand, J.R. and Zhang, X.F., 2017. The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies*, *30*(12), pp.4389-4436.

Harvey, Campbell R., Yan Liu, and Heqing Zhu. "… and the cross-section of expected returns." *The Review of Financial Studies* 29, no. 1 (2016): 5-68.

Han, Yufeng, Ai He, D. Rapach, and Guofu Zhou. "What firm characteristics drive us stock returns." *Unpublished manuscript* (2018).

Chinco, Alex, Adam D. Clark-Joseph, and Mao Ye. "Sparse Signals in the Cross-Section of Returns." *The Journal of Finance* (2018).

Lv, Dongdong, Zhenhua Huang, Meizi Li, and Yang Xiang. "Selection of the optimal trading model for stock investment in different industries." *PloS one* 14, no. 2 (2019): e0212137.

Ballings, Michel, Dirk Van den Poel, Nathalie Hespeels, and Ruben Gryp. "Evaluating multiple classifiers for stock price direction prediction." *Expert Systems with Applications* 42, no. 20 (2015): 7046-7056.

Patel, Jigar, Sahil Shah, Priyank Thakkar, and K. Kotecha. "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques." *Expert Systems with Applications* 42, no. 1 (2015): 259-268.

Basak, Suryoday, Saibal Kar, Snehanshu Saha, Luckyson Khaidem, and Sudeepa Roy Dey. "Predicting the direction of stock market prices using tree-based classifiers." *The North American Journal of Economics and Finance* 47 (2019): 552-567.

Tay, Francis EH, and Lijuan Cao. "Application of support vector machines in financial time series forecasting." *omega*29, no. 4 (2001): 309-317.

Huang, Wei, Yoshiteru Nakamori, and Shou-Yang Wang. "Forecasting stock market movement direction with support vector machine." *Computers & Operations Research* 32, no. 10 (2005): 2513-2522.

Elliott, Robert J., John Van Der Hoek*, and William P. Malcolm. "Pairs trading." *Quantitative Finance* 5, no. 3 (2005): 271-276.

Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9, no. 8 (1997): 1735-1780.

Bao, Wei, Jun Yue, and Yulei Rao. "A deep learning framework for financial time series using stacked autoencoders and long-short term memory." *PloS one* 12, no. 7 (2017): e0180944.

Chen, Kai, Yi Zhou, and Fangyan Dai. "A LSTM-based method for stock returns prediction: A case study of China stock market." In *2015 IEEE International Conference on Big Data (Big Data)*, pp. 2823-2824. IEEE, 2015.

Cheng, Li-Chen, Yu-Hsiang Huang, and Mu-En Wu. "Applied attention-based LSTM neural networks in stock prediction." In *2018 IEEE International Conference on Big Data (Big Data)*, pp. 4716-4718. IEEE, 2018.

Mullainathan, Sendhil, and Jann Spiess. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives* 31, no. 2 (2017): 87-106.

Shen, Yannan (Lily), Information Value of Property Description: A Machine Learning Approach (November 8, 2018). Available at SSRN: https://ssrn.com/abstract=3281223

Lindenthal, Thies and Erik Johnson. Machine Learning, Building Vintage and Property Values (April 15, 2018). Available at https://www.lindenthal.eu/wp-content/uploads/2018/04/manuscript.pdf

Sirignano, Justin, Apaar Sadhwani, and Kay Giesecke. "Deep learning for mortgage risk." *arXiv preprint arXiv:1607.02470*(2018).

Kvamme, Håvard, Nikolai Sellereite, Kjersti Aas, and Steffen Sjursen. "Predicting mortgage default using convolutional neural networks." *Expert Systems with Applications* 102 (2018): 207-217.

Fama, Eugene F. "The behavior of stock-market prices." *The journal of Business* 38, no. 1 (1965): 34-105.

Atsalakis, George S., and Kimon P. Valavanis. "Surveying stock market forecasting techniques–Part II: Soft computing methods." *Expert Systems with Applications* 36, no. 3 (2009): 5932-5941.

Ferson, Wayne E., and Campbell R. Harvey. "Sources of predictability in portfolio returns." *Financial Analysts Journal*47, no. 3 (1991): 49-56.

Serrano, Camilo, and Martin Hoesli. "Are securitized real estate returns more predictable than stock returns?" *The Journal of Real Estate Finance and Economics* 41, no. 2 (2010): 170-192.

Liu, Crocker H., and Jianping Mei. "The predictability of returns on equity REITs and their co-movement with other assets." *The Journal of Real Estate Finance and Economics* 5, no. 4 (1992): 401-418.

Liao, Hsien-hsing, and Jianping Mei. "Risk characteristics of real estate related securities—An extension of Liu and Mei (1992)." *Journal of Real Estate Research* 16, no. 3 (1998): 279-290.

Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794. ACM, 2016.

Géron, Aurélien. Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.", 2017.

Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9, no. 8 (1997): 1735-1780.

Table 3.1: The Summary of REIT Index Returns

| COUNT | MEAN | STD | MIN | MAX | Autocorrelation |
|---|---|---|---|---|---|
| *Panel A: The Whole Data Period (Jan 2000 -- Nov 2018)* | | | | | |
| 227 | 1.08 | 5.94 | -31.67 | 31.02 | 0.036 |
| Panel B: The Pre-Crisis Period (Jan 2000 -- July 2008) | | | | | |
| 103 | 1.30 | 4.44 | -14.58 | 8.74 | -0.050 |
| *Panel C: The Financial Crisis Period (Aug 2008 -- Dec 2009)* | | | | | |
| 17 | -0.07 | 15.79 | -31.67 | 31.02 | 0.179 |
| *Panel D: The Post-Crisis Period (Jan 2010 -- Nov 2018)* | | | | | |
| 107 | 1.04 | 4.26 | -10.97 | 14.28 | -0.169 |

*Note*: this table summarizes descriptive statistics for monthly REITs index returns. In general, the fluctuations of REITs market are substantial.

Table 3.2: The Description of Predictors

| No. | Acronym | Description | Data Source | Frequency |
|---|---|---|---|---|
| | | *Panel A: Macroeonomic Signals* | | |
| 1 | 10y | 10-Year Treasury Constant Maturity Rate | FRED of ST. LOUIS | Monthly |
| 2 | curve | 10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant Maturity | FRED of ST. LOUIS | Monthly |
| 3 | de_spread | Spread between Moody's Seasoned Aaa and Moody's Seasoned Baa coporate bond yield | FRED of ST. LOUIS | Monthly |
| 4 | capacity | Capacity Utilization | FRED of ST. LOUIS | Monthly |
| 5 | consumer | University of Michigan: Consumer Sentiment | FRED of ST. LOUIS | Monthly |
| 6 | cpi | Consumer Price Index | Bureau of Labor | Monthly |
| 7 | dollar | Trade Weighted U.S. Dollar Index: Major Currencies, Goods | FRED of ST. LOUIS | Monthly |
| 8 | lead | Leading Index for the United States | FRED of ST. LOUIS | Monthly |
| 9 | money | St. Louis Adjusted Monetary Base | FRED of ST. LOUIS | Monthly |
| 10 | order | Manufacturers' New Orders: Durable Goods | FRED of ST. LOUIS | Monthly |
| 11 | pmi | IHS Markit Purchasing Managers Index | IHS Markit | Monthly |
| 12 | ppi | Producer Price Index | Bureau of Labor | Monthly |
| 13 | retail | Percent change of Advance Retail Sales: Retail (Excluding Food Services) | FRED of ST. LOUIS | Monthly |
| 14 | unem | Civilian Unemployment Rate | FRED of ST. LOUIS | Monthly |
| | | *Panel B: Indicators of Real Estate Industry* | | |
| 15 | construct | Total Construction Spending | FRED of ST. LOUIS | Monthly |
| 16 | cs | S&P/Case-Shiller U.S. National Home Price Index | FRED of ST. LOUIS | Monthly |
| 17 | fhfa | FHFA House Price Index | FHFA | Monthly |
| 18 | nahb | NAHB Housing Market Index | NAHB | Monthly |
| 19 | permit | New Private Housing Units Authorized by Building Permits | FRED of ST. LOUIS | Monthly |
| 20 | price | Median Sales Price of Houses Sold | FRED of ST. LOUIS | Monthly |
| 21 | sale | New One Family Houses Sold | FRED of ST. LOUIS | Monthly |
| 22 | start | Housing Starts: Total: New Privately Owned Housing | FRED of ST. LOUIS | Monthly |
| 23 | supply | Monthly Supply of Houses | FRED of ST. LOUIS | Monthly |
| 24 | vacancy | Rental Vacancy Rate | FRED of ST. LOUIS | Quarterly |
| | | *Panel C: Equity Market Features* | | |
| 25 | hml | High minus low factor | Fama-French Factors | Monthly |
| 26 | margin | Debit Balances in customers' securities margin accounts | FINRA | Monthly |
| 27 | m_dy | Dividend yield of S&P 500 index | CRSP | Monthly |
| 28 | mktrf | Difference between equity market return and risk free rate | Fama-French Factors | Monthly |
| 29 | m_mom | Momentum factor of market | Fama-French Factors | Monthly |
| 30 | m_pe | PE ratio of S&P 500index | CRSP | Monthly |
| 31 | smb | Small minus big factor | Fama-French Factors | Monthly |
| 32 | vix | The CBOE Volatility Index | CRSP | Monthly |
| | | *Panel D: Characters of REITs Index* | | |
| 33 | crisis | The dummy varaible for financial crisis period | Caculation | Monthly |
| 34 | dvd_y | Dividend yield of REITs Index | NAREIT | Monthly |
| 35 | ema | Exponential moving average of Wilshire US REIT Index | Caculation | Daily |
| 36 | jan | The dummy varaible for January effect | Caculation | Monthly |
| 37 | l_t | Long trend from fourier transform | Caculation | Monthly |
| 38 | macd | Moving average convergence/divergence of Wilshire US REIT Index | Caculation | Daily |
| 39 | reit_mom | 12-2 momentum factor of REITs Index | Caculation | Monthly |
| 40 | return | NAREIT REITs Index Return | NAREIT | Monthly |
| 41 | s_t | Short trend from fourier transform | Caculation | Monthly |

*Note*: this table shows a comprehensive list of features, including the economic factors, the equity market conditions, the signals of technique analysis, and the characteristics of REITs index.

Table 3.3: The Summary of Out-of-Sample Prediction Error

| | Linear | KF | LASSO | CART | RF | AdaBoost | GBRT | XGBoost | SVR | Auto-LSTM | Ensemble |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panel A: The Whole Data Period (July 2003 -- Nov 2018)* | | | | | | | | | | | |
| Min | 0.004 | 0.017 | 0.002 | 0.058 | 0.072 | 0.026 | 0.006 | 0.039 | 0.064 | 0.009 | 0.001 |
| Max | 1392.59 | 32.96 | 32.25 | 27.13 | 32.44 | 30.67 | 32.69 | 30.18 | 33.52 | 32.82 | 30.69 |
| Mean | 16.89 | 4.33 | 4.79 | 5.44 | 4.39 | 4.62 | 4.49 | 4.41 | 4.83 | 4.33 | 4.41 |
| Variance | 10975.6 | 22.07 | 28.93 | 26.09 | 21.42 | 23.81 | 23.63 | 22.23 | 29.85 | 22.47 | 22.12 |
| RMSE | 105.84 | 6.38 | 7.19 | 7.45 | 6.37 | 6.71 | 6.60 | 6.45 | 7.28 | 6.41 | 6.44 |
| MAPE | 5.38 | 2.33 | 3.48 | 4.77 | 1.21 | 2.61 | 2.78 | 2.14 | 3.33 | 2.18 | 2.22 |
| *Panel B: The Pre-Crisis Period (July 2003 -- July 2008)* | | | | | | | | | | | |
| Min | 0.004 | 0.078 | 0.286 | 0.197 | 0.164 | 0.073 | 0.041 | 0.039 | 0.104 | 0.138 | 0.197 |
| Max | 16.05 | 16.28 | 15.48 | 19.16 | 14.21 | 18.06 | 16.49 | 15.42 | 18.14 | 16.64 | 16.65 |
| Mean | 5.24 | 3.92 | 4.01 | 4.97 | 4.21 | 4.06 | 4.45 | 4.11 | 4.21 | 3.83 | 4.03 |
| Variance | 19.05 | 9.87 | 11.95 | 16.46 | 8.64 | 12.48 | 13.04 | 8.82 | 13.58 | 10.17 | 10.37 |
| RMSE | 6.8 | 5.00 | 5.28 | 6.40 | 5.12 | 5.36 | 5.72 | 5.05 | 5.58 | 4.97 | 5.14 |
| MAPE | 5.53 | 3.52 | 4.03 | 7.07 | 1.37 | 3.47 | 4.78 | 3.79 | 3.95 | 3.56 | 3.08 |
| *Panel C: The Financial Crisis Period (Aug 2008 -- Dec 2009)* | | | | | | | | | | | |
| Min | 5.037 | 0.914 | 2.500 | 0.253 | 0.597 | 1.167 | 0.972 | 1.068 | 2.579 | 1.056 | 1.377 |
| Max | 1392.59 | 32.96 | 32.25 | 27.13 | 32.44 | 30.67 | 32.69 | 30.18 | 33.52 | 32.82 | 30.69 |
| Mean | 133.97 | 11.76 | 14.74 | 12.74 | 11.72 | 12.58 | 12.03 | 12.51 | 14.99 | 12.06 | 12.48 |
| Variance | 109978.00 | 106.44 | 102.00 | 86.85 | 104.84 | 95.91 | 103.73 | 97.69 | 109.97 | 104.89 | 91.94 |
| RMSE | 348.50 | 15.44 | 17.70 | 15.62 | 15.36 | 15.77 | 15.57 | 15.76 | 18.12 | 15.62 | 15.56 |
| MAPE | 15.52 | 1.33 | 11.02 | 1.33 | 1.38 | 2.51 | 1.29 | 1.38 | 7.87 | 1.34 | 2.80 |
| *Panel D: The Post-Crisis Period (Jan 2010 -- Nov 2018)* | | | | | | | | | | | |
| Min | 0.002 | 0.017 | 0.002 | 0.058 | 0.072 | 0.026 | 0.006 | 0.082 | 0.064 | 0.009 | 0.001 |
| Max | 26.04 | 13.19 | 17.50 | 19.29 | 13.79 | 18.01 | 14.77 | 13.78 | 19.27 | 13.31 | 15.54 |
| Mean | 4.94 | 3.39 | 3.65 | 4.55 | 3.33 | 3.68 | 3.31 | 3.30 | 3.56 | 3.39 | 3.35 |
| Variance | 20.14 | 6.83 | 10.50 | 13.39 | 6.70 | 8.56 | 7.43 | 7.05 | 9.12 | 6.80 | 6.98 |
| RMSE | 6.66 | 4.27 | 4.87 | 5.83 | 4.21 | 4.70 | 4.28 | 4.23 | 4.66 | 4.27 | 4.26 |
| MAPE | 3.69 | 1.82 | 1.97 | 4.00 | 1.09 | 2.13 | 1.87 | 1.33 | 2.26 | 1.52 | 1.63 |

*Note* : this table summarizes the statistics of prediction performance among all models. All number is the percentage expression.

Table 3.4: the Performance of Investment Strategy

| | Buy-Hold | Linear | KF | LASSO | CART | RF | AdaBoost | GBRT | XGBoost | SVR | Auto-LSTM | Ensemble |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panel A: The Whole Data Period (July 2003 -- Nov 2018)* | | | | | | | | | | | | |
| Return | 3.45 | -0.30 | 3.45 | 3.94 | 0.03 | -0.08 | 2.70 | 3.49 | 1.22 | 2.12 | 3.45 | 7.83 |
| Sharpe Ratio | 0.16 | 0.00 | 0.16 | 0.17 | 0.03 | 0.03 | 0.14 | 0.16 | 0.10 | 0.13 | 0.16 | 0.22 |
| *Panel B: The Pre-Crisis Period (July 2003 -- July 2008)* | | | | | | | | | | | | |
| Return | 1.02 | 0.61 | 1.02 | 0.33 | -0.03 | -0.59 | 0.34 | 0.40 | 1.02 | 0.80 | 1.02 | 0.88 |
| Sharpe Ratio | 0.26 | 0.18 | 0.26 | 0.12 | 0.02 | -0.26 | 0.12 | 0.13 | 0.26 | 0.22 | 0.26 | 0.23 |
| *Panel C: The Financial Crisis Period (Aug 2008 -- Dec 2009)* | | | | | | | | | | | | |
| Return | -0.20 | -0.57 | -0.20 | 0.48 | 0.14 | -0.20 | 0.14 | -0.04 | -0.60 | -0.27 | -0.20 | 0.70 |
| Sharpe Ratio | 0.00 | -0.24 | 0.00 | 0.22 | 0.12 | 0.00 | 0.12 | 0.06 | -0.28 | -0.04 | 0.00 | 0.28 |
| *Panel D: The Post-Crisis Period (Jan 2010 -- Nov 2018)* | | | | | | | | | | | | |
| Return | 1.76 | 0.02 | 1.76 | 1.52 | -0.07 | 1.76 | 1.42 | 2.35 | 1.76 | 1.37 | 1.76 | 1.76 |
| Sharpe Ratio | 0.24 | 0.03 | 0.24 | 0.22 | 0.01 | 0.24 | 0.21 | 0.29 | 0.24 | 0.21 | 0.24 | 0.24 |

*Note* : this table shows the performance of this investment strategy. The number in the table measures how large the investment changes


Figure 3.2: The Description of REIT Index

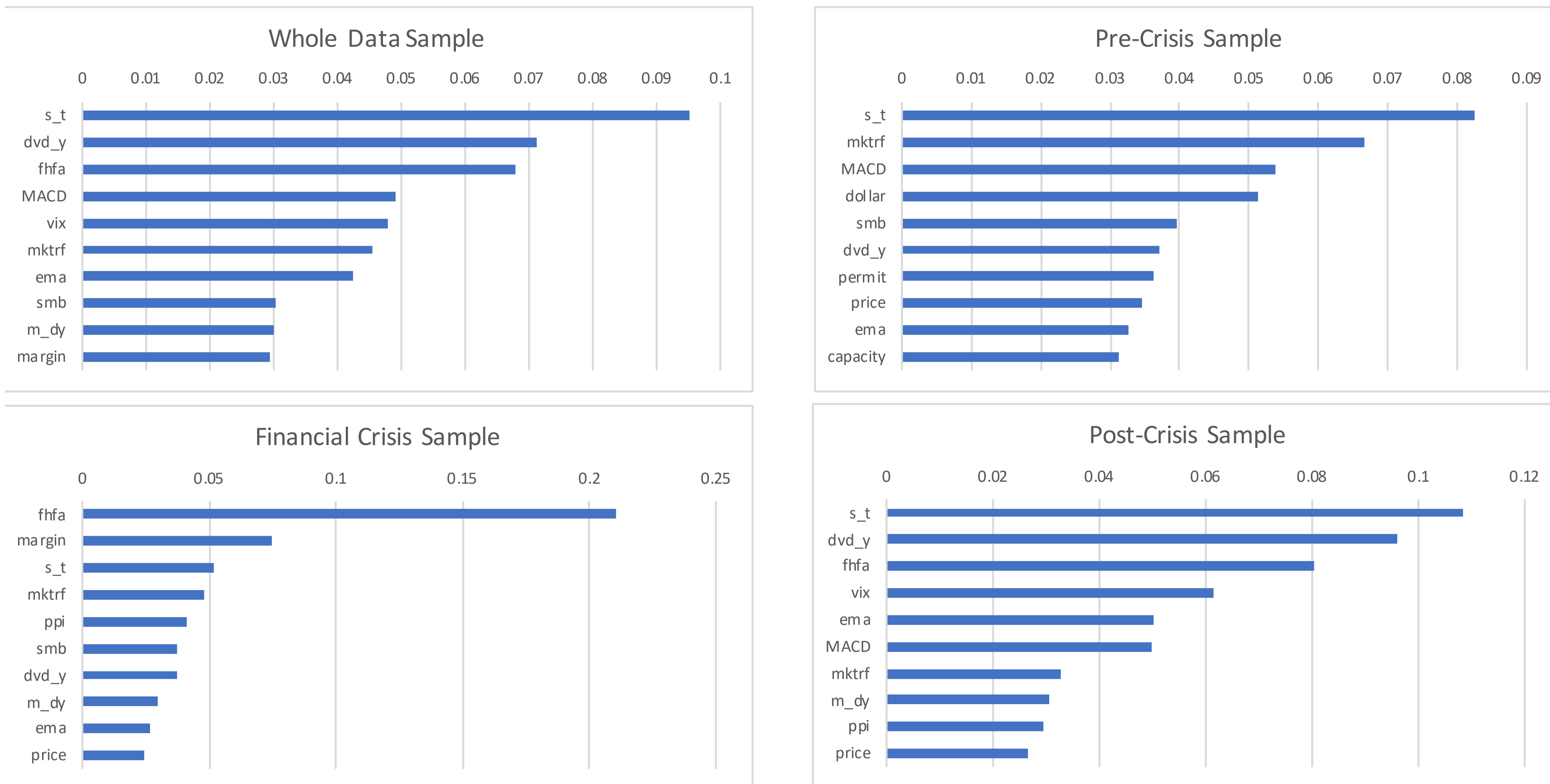



Figure 3.6: The Top 10 Feature Importance. This figure demonstrates the feature importance for the top 10 predictors in four different periods. As the sum of feature importance of a given model is equal to one, the value of each factor importance explains the relative importance for the specific model.

# CONCLUSION

My research topics are generally in the field of empirical asset pricing. My dissertation discovers the valuation of REIT CDS in both theoretical and empirical area, and further shows the credit information flow among three primary trading market, and finally discuss the predictability of the REIT market. In general, my first paper examines determinants of CDS valuation in the REIT industry for the first time. Since default probability primarily determines CDS prices, studying these determinants provides new information concerning credit risk factors for REITs. My second paper contributes to the direction of credit information flow among three different REIT-related markets considering the information content, speed, and efficiency. My third paper applies machine learning and deep learning models to test the predictability of the REIT index for the first time. All these papers shed light on the valuation and development of the REIT industry.