2015

Three Essays on Default Risk in Capital Markets

Laleh Samarbakhsh
Wilfrid Laurier University, lalehsamar@yahoo.com

Doctoral Dissertation

Three Essays on Default Risk in Capital Markets

Author:

Laleh Samarbakhsh

B.Sc. Mathematics, Sharif University of Technology

M.Sc. Financial Mathematics, Wilfrid Laurier University

Submitted in partial fulfillment of the requirements for

Doctor of Philosophy in Management

at

School of Business and Economics

Wilfrid Laurier University

Dissertation Committee Members:

Dr. Ben Amoako-Adu,

Dr. Madhu Kalimipalli (Chair),

Dr. Subhankar Nayak

Internal/External Committee Member:

Dr. Joe Campolieti

Dissertation External Examiner:

Dr. Louis Gagnon

Declaration of Authorship

I, Laleh Samarbakhsh, declare that this dissertation titled "Three Essays on Default Risk in Capital Markets" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a Doctoral degree at Wilfrid Laurier University.
- Where any part of this dissertation has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this dissertation is entirely my own work.
- I have acknowledged all main sources of help.

Signed:

Date:

Acknowledgements

I would like to express my sincere gratitude to my dissertation committee chair, Dr. Madhu Kalimipalli for his insightful feedback, supervision, and encouragement. In addition, I am indebted to my dissertation committee members, Professor Ben Amoako-Adu (who also has generously supervised the first essay of my dissertation) and Dr. Subhankar Nayak for their guidance, insightful comments and support over the dissertation process. Also, I am grateful to the Laurier Finance faculty and my fellow PhD students for their feedback and support. I would like to thank Professor Phelim Boyle for his support and guidance throughout my studies at the School of Business and Economics.

I am grateful to Dr. Joe Campolieti for acting as the internal/external at my dissertation defence. I am thankful to Professor Louis Gagnon, who generously accepted the role of external examiner, for his many helpful comments and insightful recommendations on all three essays, and sincere encouragement.

I acknowledge the comments received on my work at the MFA Annual Conference (2012), the FMA-Europe convention (2012), the EFA Conference (2013), the NFA Conference (2013), the MFA Conference (2015), and from the seminar participants at the Ryerson University, Lethbridge University, Memorial University, University of Massachusetts Boston, and Wilfrid Laurier University.

I dedicate this dissertation to my beloved parents for their unconditional love, and would like to thank my brother and sisters for their never-ending support throughout this process. Last but not least, I appreciate the support and help of my partner (now husband) throughout my doctoral studies. - Thank you!

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION AND OVERVIEW

1.1 Introduction

This dissertation comprises three essays on default risk in capital markets. The three essays investigate major issues surrounding default risk in capital markets and the relationship with the 2008 financial crisis by focusing on important segments of capital markets including investments portfolio (hedge funds), derivatives (equity options market), fixed income (credit default swaps), and equity market.

The first essay "The role of Leverage in Hedge Funds Failure" investigates the role of financial leverage, including the use of margins and derivative products, in the hedge funds failure during the 2008 financial crisis. Motivated by failure of the two Bear Stearns hedge funds at the beginning of the financial crisis in 2007, this paper examines why some hedge funds failed during and after the recent financial crisis, and why some also survived. Leverage is defined in three ways, as (a) debt/equity ratio, (b) the use of margins, and (c) the use of derivative products. The research uses a 15-year panel dataset of 17,202 failed and survived hedge funds from the Lipper TASS Hedge Fund database. The empirical analysis, using probit regression to estimate the likelihood of failure, shows that during the financial crisis period, financial leverage is more significant in increasing the probability of failure, whereas financial leverage becomes insignificant in explaining the probability of hedge fund failure during non-crisis periods after controlling for fund structure, size, incentive fees, prior performance, and off-shore registration. The results are consistent with Ang et al. (2011) who analyze the cyclical leverage for financial institutions and hedge funds and find that financial leverage decreases during a financial crisis period because the funds sell some assets to meet their margin requirements and that forces some funds into liquidation. Further analysis shows that some hedge funds which follow specific styles such as "Emerging Markets", "Equity Market Neutral", "Long/Short Equity Hedge", and "Multi-strategy", which have higher than average betas, are also more likely to fail during the financial crisis.

The second essay, "Does default risk impact Equity Options?", explores the impact of default risk on equity option pricing. The impact is studied in detail by empirically examining to what extent the firm-specific default risk matters in pricing individual equity options. Since credit default swaps (CDS) are similar to put options in that both offer a low cost and effective protection against downside risk, we use CDS spreads as a credit risk proxy to investigate the effects of default risk on put pricing. The recent financial crisis showed that, for many financial firms, equity options experienced high implied volatility (IV) when the underlying CDS spreads went up. By examining an exhaustive sample of US-listed firms with both CDS and put option data available over the period from 2002 to 2010, and studying the primary determinants of option IVs cross-sectionally and over time, the findings show that default risk is a significant factor in the prices of equity options. Moreover, the impact of default risk remains significant after controlling for firm-specific and macroeconomic factors. This study relates to recent literature that explains how default risk can get injected from the fixed income market to the equity options market and why default risk is important in the pricing of equity options and implied volatility.

The third essay, "Forecasting Option-implied Volatility using credit risk", addresses the issue of forecasting option implied volatility which is of interest to option market participants, who routinely formulate volatility and option price forecasts for trading and hedging purposes. As shown in essay two, credit risk matters for option pricing since options are valued on firms with significant trading liquidity, yet subject to default risk, similar to liquidity risk. If credit risk matters for option prices, this essay particularly explores whether better out-of-sample forecasts for option implied volatility (IV) can be developed using lagged credit risk measures. Various time-series forecasts of daily, weekly, and monthly option implied volatilities show that inclusion of default risk as measured by credit default swap (CDS) can significantly improve out-of-sample performance, measured through a decreased mean squared error (MSE) as well as a smaller root mean squared error (RMSE).

1.2 Dissertation Overview

The dissertation is organized in five chapters. Following the introduction, chapter two discusses the risk of failure in hedge funds (as one of the important segments impacted during financial crisis), focusing on the role of excess leverage and its impact on increasing the default risk and consequently failure of hedge funds.

Chapters three and four relate to the issues surrounding credit risk and options market. In Chapter Three by measuring default risk stemmed from credit default swap, and investigating its overlap with options markets, it is discussed to what extent default risk can flow into equity options pricing. In Chapter Four the inter-relationship between implied volatility and the credit default swap, is studied in time-series details and the forecasting power of CDS is identified as well as its role in improving out-of-sample performance of predicting implied volatility.

The reader can consider these three chapters as three independent academic essays. There is no need to read Chapter Two first in order to understand Chapters Three and Four. In each chapter I provide an introduction, an overview of the related literature, a discussion of hypotheses and empirical tests, results and conclusions. In what follows I provide the complete findings in each chapter.

Chapter 2

THE ROLE OF LEVERAGE IN HEDGE FUNDS FAILURE

Abstract

This research investigates the role of financial leverage, including the use of margins and derivative products, in the hedge funds failure during the 2008 financial crisis. Motivated by failure of the two Bear Stearns hedge funds at the beginning of the financial crisis in 2007, this paper examines why some hedge funds failed during and after the recent financial crisis, and why some also survived. Leverage is defined in three ways, as (a) debt/equity ratio, (b) the use of margins, and (c) the use of derivative products. The research uses a 15-year panel dataset of 17,202 failed and survived hedge funds from the Lipper TASS Hedge Fund database. The empirical analysis, using probit regression to estimate the likelihood of failure, shows that during the financial crisis period, financial leverage is more significant in increasing the probability of failure, whereas financial leverage becomes insignificant in explaining the probability of hedge fund failure during non-crisis periods after controlling for fund structure, size, incentive fees, prior performance, and off-shore registration. The results are consistent with Ang, Gorovyy and Inwegen (2011) who analyze the cyclical leverage for financial institutions and hedge funds and find that financial leverage decreases during financial crisis period because the funds sell some assets to meet their margin requirements and that forces some funds into liquidation. Further analysis shows that some hedge funds which are registered under specific styles such as "Emerging Markets", "Equity Market Neutral", "Long/Short Equity Hedge", and "Multi-strategy", which have higher than average betas, are also more likely to fail during the financial crisis.

Keywords: Financial Leverage, Hedge Fund Failure, Financial Crisis, Hazard Model

2.1 Introduction

Hedge funds have increasingly become important investment portfolios in recent years for both rich individuals and institutional investors. Total assets under management for the hedge fund industry grew rapidly from $39 billion in 1990 to $2.3 trillion at the peak of 2007.[1] Jickling and Raab (2006) state that hedge funds account for about 30% of trading volume in US securities markets. Following the global financial crisis in 2008, the hedge fund industry also suffered significant losses. Though hedge funds are huge portfolios which are managed by well-informed and sophisticated managers, they are also vulnerable during the financial crisis. The total assets under management for the industry was cut into half to approximately $1.3 trillion in the first quarter of 2009 when a large number of hedge funds were forced to liquidate or go bankrupt. This research is designed to empirically test the role that financial leverage including margins played in causing some hedge funds to fail while others survived during the recent global financial crisis.

It is known that financial leverage magnifies returns in bull markets and also magnifies losses in bear markets. However, Ang, Gorovyy and Inwegen (2011) in comparing the cyclical behavior of financial leverage of banks and hedge funds, found that in the case of hedge funds, during bear markets, financial leverage ratios drop significantly because the hedge funds liquidate some of their portfolio of assets to meet their margin requirement. In contrast, it is also known that in 1998, Long Term Capital Management failed because of massive financial leverage during adverse market conditions in Russia. With this opposing effect of leverage

[1] Titman and Tiu (2011)

on hedge funds, I empirically evaluate the role of financial leverage in hedge fund failure prior to and after the 2008 financial crisis.

It is argued that the financial crisis period started from the filing of bankruptcy by two Bear Stearns hedge funds in August 2007, and ended in March 2009[2].The objective of this study is to empirically investigate the role of leverage in influencing a hedge fund risk of failure during financial crisis and severe distress times. The linkage between the failure of hedge funds and systemic risk is also of great importance. Systemic risk was greatly observed during the recent financial crisis period. It is also important to note that the correlations between stocks and aggregate market are much larger for downside moves, especially for extreme unexpected downside moves, than for upside moves. This is in-line with findings of Ang and Chen (2002) on "Asymmetric correlation". This fact was also empirically evident during the recent financial crisis. In this paper, we split the sample data with a perceived regime switch and run stability tests to identify the difference in the performance and magnitude of hedge funds' failure drivers among these sub-samples.

I attempt to study the relationship between hedge fund leverage and risk of failure before and during/after the 2008 financial crisis after controlling for structure, and management fees and off-shore location. Leverage magnifies the returns in up markets, as well as magnifying the losses in the down markets. This research will shed light on some key questions about hedge funds' failure, the role of leverage, and its drivers during the financial crisis. The results will help provide

[2]Though these two funds were registered in Cayman Islands, for bankruptcy protection, the New York bankruptcy court ruled against them for credit protection under Chapter 15.

reasons why hedge funds being huge portfolios fail frequently.

With the large number of bankrupt hedge funds during and after the recent financial crisis, the remaining unclear connection is why did many hedge funds fail in the course of the financial crisis and why did many survive? What factors drive the propensity of failure for a hedge fund during financial distress times (and may not necessarily remain a driver during non-crisis periods)? Focusing on 15 years of data, that as well captures the crisis period, this paper's main contribution to the hedge fund literature is to identify factors that increased the risk of failure of hedge funds during the recent financial crisis. Defining financial leverage as debt/equity ratio, margins, and the use of derivative products in the hedge fund investment style, the results indicate that prior to the crisis, financial leverage was not significant in predicting the probability of failure of hedge funds. However, during and after the financial crisis, financial leverage was statistically significant in predicting hedge fund failure. This means the liquidation of investment assets to meet margin requirement forced some hedge funds into liquidation because of the systemic risk of the crisis. This paper adds to the literature by providing empirical analysis of the significance of financial leverage in the failure of hedge funds during the recent financial crisis. By exploring this, it complements the existing literature on hedge funds and makes additional contributions through testing hypotheses to follow.

2.2 Literature Review

There is a large literature on hedge funds which compares their performances with mutual funds, analyzes their investment styles, reviews their risk taking factors and studies their role in systemic risk[3]and bubbles. We will review some of them in this section. However, not much is known about the effect of financial leverage on the failure of hedge funds during the recent market crisis.

Titman and Tiu (2011) record that the size of the hedge fund industry includes more than 11,000 active funds that managed about $2.3 trillion at its peak in 2007, with management fees averaging 1.5% and with 20% of returns in incentive fees. Bali, Brown, and Caglayan (2011) record a total AUM of $2.09 trillion in 2007 which dropped to $1.31 by the end of 2010. Figure 3 shows the total AUM from 1994 to 2010. Such a huge growth in hedge funds has also been accompanied by a growth in the number and severity of failures including forced liquidation and bankruptcy. After the failure of Long Term Capital Management (LTCM) in 1998, investors recognized that hedge funds may provide high expected returns but many of them may be doing so by taking a huge downside risk that is not easily verifiable by merely monitoring traditional risk measures such as standard deviation. Hedge funds usually generate nonlinear and non-normal payoffs due to various factors such as actively trading derivative securities, implementing option-like dynamic trading strategies, having investment styles that experience severe losses during market downturns, and charging an incentive fee that has the same effect to the

[3] The possibility that the failure of a single institution could set off cascading failures throughout the financial system is known as systemic risk. The classic case is a banking crisis, where trouble in one bank can trigger runs on others, including financially sound institutions, if enough cautious depositors decide to withdraw their funds. Source: CRS report for Congress; Dec 4, 2006

investors as shorting a call option on the value of the fund portfolio (Agarwal and Naik (2004), Mitchell and Pulvino (2001), Taleb (2004) and Goetzmann, Ingersoll and Ross (2003).

Many empirical studies provide evidence for nonlinearity and non-normality of hedge funds' returns. Hedge funds on average have negative skewness and excess kurtosis, and the rejection rate of the Jarque-Bera (JB) test for normality is 40.5% to 85.9% depending on the test period and the database used (Cremers, Kritzman and Page (2005), Alexiev (2005), Bali, Gokcan and Liang (2006), Liang and Park (2007)).

There have been recent efforts on analyzing hedge funds' exposures to various financial and macroeconomic risk factors. Bali, Brown, and Caglayan (2011) use univariate, bivariate, and multivariate estimates and investigate performance of these factor loadings (betas) in predicting the cross-sectional variation in hedge fund returns. They find significantly positive (negative) link between default premium beta (inflation beta) and future hedge fund returns.

Christory, Daul, and Giraud (2006) distinguish between three reasons for hedge fund failure. First, financial issues, or losses stemming from unfavorable market moves; second, operational issues, such as errors in trade processing or mispricing complex, opaque financial instruments; and third, fraud, or misbehavior by fund management. The authors state that the most common cause is undoubtedly the first or financial reasons. When hedge funds fail to earn the expected returns, redemptions may exceed inflows so they are subsequently unable to attract new investors, and managers find it unprofitable to continue. As such, the next resolution is to dissolve the fund, in accordance with the partnership agreement, and return remaining assets to the investors.

In line with the relationship between financial leverage and failure during crisis and non-crisis periods, we should also refer to literature on systemic risk and comovement of the financial markets during such distress periods.[4] Ang and Chen (2002) show that correlations between stocks and aggregate market are much larger for downside moves, especially for extreme unexpected downside moves, than for upside moves. This fact was also empirically evident during the recent financial crisis. In their paper they suggest the use of regime switching models which provide greater flexibility in capturing these sudden moves. Figure 4 shows the timeline and perceived regime switch which is used for the analysis of before and after the financial crisis period.

In a more recent paper, Titman and Tiu (2011) show that better-informed hedge funds choose to have less exposure to significant risk factors. They find that hedge funds that exhibit lower R-squared with respect to systematic factors have higher Sharpe ratios, higher information ratios, and higher alphas with higher manipulation-proof performance measures and higher fees.

The question that remains unanswered is why did many hedge funds survive during the financial crisis while many failed? Is there a difference in the role of financial leverage in the failure of hedge funds during distress times? What factors

[4] Systemic Risk is defined as the possibility of a series of correlated defaults among financial institutions, typically banks, that occurs over a short period of time and is often caused by a single major event. (A. W. Lo, 2010)

increase the probability of failure for a hedge fund during financial distress times (and may not necessarily be drivers during non-crisis periods)? This paper's main contribution to the hedge fund literature is to explore the significance of financial leverage in the failure of hedge funds during the recent financial crisis.

2.3 Hedge Funds Background

A hedge fund is an investment portfolio which can use any technique with the primary objective of earning a higher return. Often hedge funds do not benchmark their performance against any index. Unlike mutual funds, hedge funds are not regulated by SEC. The size of the hedge funds industry has almost doubled every two years since 1994 with more than 11,000 active funds that managed more than \$2.5 trillion at its peak in 2008.[5]Hedge funds are often registered as a partnership and then opened to a few rich accredited outside investors. The investment capital from each investor could be large; however, retail investors can invest in hedge funds through fund of funds (FoFs). For example, some mutual funds and Pension Funds invest in hedge funds on behalf of their clients. Because hedge funds are not strictly regulated as compared to mutual funds, the disclosure of information to both the investors and the public is sketchy.[6]

[5] Hedge funds were started in 1949 by Alfred Winslow Jones, a sociologist who was on the editorial staff of Fortune Magazine. Jones used short selling to hedge the risk of his fund and used leverage to increase his returns. These techniques have become a common practice among modern hedge funds. Today, not all hedge funds necessarily hedge. They can be identified by their exemption from the Investment Company Act of 1940 and the unique incentive compensation technique.

[6] A failure example of fraud detection is Bernard Madoff's Ponzi or pyramid scheme whose investors lost more than \$65 billion in US in 2009. Madoff's fund was considered a hedge fund and regulators did not find out about his scheme for over 15 years.

Given the enormous growth of this industry and limited information available on the hedge funds (and the events of the financial crisis), a need for research has emerged which benefits both investors and regulators. Regulators need to identify whether there is a need for regulation to protect investors' interests and what impact hedge funds may have on the stability of the financial systems. Moreover, after the events of the recent financial crisis, there aa re many unanswered questions about the reasons behind hedge funds' failure. In this research we would like to answer some key questions about a hedge funds' failure and its drivers.

2.3.1 Return Generating Process of Hedge Funds

Liang (1999) compares hedge funds with mutual funds and finds that hedge funds have higher returns but also higher risk. So, one should use a risk-adjusted performance measure such as the Sharpe Ratio to compare the two. On average, hedge funds have a 0.47 Sharpe ratio compared to 0.26 for mutual funds, about 0.21 more than mutual funds (Ackerman (1999)).[7]

A number of studies show that hedge funds payoffs are nonlinear due to their use of dynamic option-like trading strategies. Goetzmann, Ingersoll, Ross (2002) show that the Sharpe ratio can be manipulated by use of option-like strategies that can alter the shape of the probability distribution of returns.[8]

[7] Sharpe ratio for hedge funds should use an international risk-free rate, such as LIBOR, because most of the funds invest abroad during opportune times. Also, an international market index should be used for the R_m

[8] Finding a suitable market portfolio for comparing hedge funds performance analysis against has also been discussed in the literature (Liang 1999, Fung and Hsieh 2000, 2008).

2.3.2 Manager's Skill in Hedge Funds

A linear multifactor model can determine two components of hedge funds return, mainly α and β. Beta is the return of the fund related to exposure to different asset classes (a number of factors), and alpha is the return above what can be explained by asset classes (intercept of the model).

Since alpha as explained above is the return of the hedge fund that cannot be explained by exposure to the systematic risk factors, it is normally interpreted as return attributed to manager's skills.

Fung and Hsieh (2001) use data from HFR, TASS and MSCI (Morgan Stanley Capital International) to show empirically that Equity long/short hedge funds have significant positive alphas. Another explanation for alpha can be that models simply are missing a significant risk factor. Some suggestions have been provided in the literature, such as market price of variance risk (Bondarenko 2004), share restrictions of private investment funds (Aragon 2007).

2.3.3 Hedge Fund Management Expense Ratio (MER) and Incentive Fees

The fee structure of a hedge fund in general consists of a management fee based on total assets under management (AUM) and a performance fee on any profits earned beyond a benchmark. This benchmark is typically called high-water mark, which is the highest year end net asset value. The MER is usually a percentage of the asset under management. The hedge fund manager's overall compensation is based on his relative performance compared to the high-water mark. Some hedge funds may charge performance fees only. Aragon (2007) show that 68% of hedge funds in TASS database use the high-water mark.

The performance fee is an important component of the hedge fund manager's compensation which may also be called incentive fee. The aim of the incentive fee is to align the interests of manager and investors by encouraging better performance. On the other hand, it is an incentive for taking higher (and sometimes excessive) risk.

One key difference between mutual funds and hedge funds is in the management's compensation. While mutual funds charge just Management Expense Ratio (MER) of about 1%-5%, hedge funds share in the returns by taking 10%-50% of the annual returns in addition to the MER that they charge to manage the funds.

2.3.4 Failure of Long-Term Capital Management

The 1998 failure of Long-Term Capital Management (LTCM), an American hedge fund which managed more than $100 billion in assets at its peak, is an important historical case.[9] Between 1994 and 1998 the fund showed a return on investment of more than 40% per annum. However, in one month, LTCM lost $1.9 billion. Prior to the bailout, the fund was down in value to $2.9 billion. Perceived as a financial disaster, the fund's collapse had significant international monetary implications, jeopardizing the financial system itself. Prompted by deep concerns about LTCM's numerous derivative contracts, in order to avoid a panic by banks and investors

[9] Among LTCM's principals were former distinguished finance professors, including two Nobel Prize-winning researchers.

worldwide, the Federal Reserve Bank of New York stepped in to organize a bailout with the various major banks at risk. There are several articles that studied the case of LTCM.[10]

These studies highlight the risks associated with high leverage in the hedge fund industry by the failure of LTCM. Unanimously, it is shown that for a hedge fund, if the leverage is too high, the potential losses can seriously exhaust and even wipe out the hedge funds net worth once the creditors call in their loans. Therefore, high leverage increases the risk of forced liquidation. LTCM had $125 billion of assets on a base of $4 billion of investors' money shortly before its collapse. This translates into a leverage ratio of about 31.25 debt-to-equity.[11]

2.3.5 Hedge Funds Strategies

Hedge funds employ many different trading strategies, which are classified in many different ways. There is no standard system but there are key categories. A hedge fund will typically commit itself to a particular strategy, particular investment types and leverage limits via statements in its offering documentation, thereby giving investors some indication of the nature of the particular fund. Each strategy can be said to be built from a number of different elements such as Style, Market, Instrument, Exposure, Sector, Method and Diversification. The main strategy groups are based on the investment style and have their own risk and return characteristics.

[10] Jorion (2000) investigates the red flags in the risk management at LTCM, Kabir and Kabir (2005) study the "too-big-to-fall" hypothesis for LTCM and the bailout decision at the time, while Dungey, Fry, Gonzalez, and Martin (2006) study the contagion impact of bond market in the LTCM Failure. Edwards and Franklin (2006) also study the governance and regulations implications for hedge funds based on lessons learned from LTCM case.

[11] Source: http://riskinstitute.ch

In this research, we perform the study on each group of the strategies. As such, we group the hedge funds into 12 main strategy groups. This will be further explained in the methodology section.

2.4 A Hazard Model of Hedge Fund Failure

In order to perform a survival study we need to model hazard rate. In their paper, Liang and Park (2010) distinguish attrition rate from real failure rate of hedge funds by implementing a survival analysis for each of the two definitions of failure. Attrition is used for the traditional definition (i.e. the fund simply doesn't exist in the live database) and real failure, the new definition.[12]We use the same method for identifying real failure. The hazard function follows the same specification as in Shumway (2001) and He, Chong, Li, and Zhang (2010).

Suppose that Fund i can either fail or survive. The time to leave the live database t and the status j are observed, where $j = 1$ corresponds to the case of failure, and $j = 0$ corresponds to the case of survival. Define hazard rate by $\lambda_i(t)$ which specifies the instantaneous rate of failure of Fund i at time t conditional upon the fund's survival up to t. For each state j, there is a latent duration T_j , which is the time that elapses before the fund chooses route j in the absence of any other risks. Thus, the actual exit time from solvent to insolvent state can be interpreted as the realizations of random variable T which is defined as follows:

$$T = min(T_j, j = 0, 1)$$

[12] Liang and Park (2010) suggest new criteria to define "real failure" of hedge funds as: i) those included in a database but stopped reporting, ii) negative average rate of return for the last six months , and iii) decreased AUM for the last twelve months

At each point in time, the hazard function for failure is defined as:

$$\lambda_i(t) = \frac{lim_{\Delta t \to 0} Pr(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} \tag{2.1}$$

In the risk-specific hazard function with the proportional hazard (PH) Cox (1972) model, a vector of fund characteristics is introduced to explain the hazard rate. Vector z's components are called "covariates".

$$\lambda(t \mid z_i(t)) = \lambda_0 exp(z_i^t \beta) \tag{2.2}$$

where z^T denotes the transpose of the vector z (a vector of time-dependent covariates for fund i), and $\lambda_0(t)$ is the baseline hazard function. The vector β is assumed to be the same for all funds. To estimate β, Cox (1972, 1975) introduced the partial likelihood function, which eliminates the unspecified baseline hazard $\lambda_0(t)$ and accounts for censored survival times. The partial likelihood function, $L_i(\beta_j)$ can then be calculated with the following specification:

$$L_j(\beta) = \prod_{i=1}^{k} \frac{exp(Z_i(t_i)\beta)}{\sum_{t \in R(t_i)} exp(Z_I(t_i)\beta)} \tag{2.3}$$

where k refers to the number of funds in specific hazard group j (i.e. failed since here $j = 1$), and $t_1 < ... < t_k$ denotes the k ordered failures of hazard group. $R(t_i) = I \mid t_I \geq t_i$ is the set of funds that have not failed at time t_i (See He, Chong, Li, and Zhang (2010) and Kalbeisch and Prentice (2002) for further details.) Brown, Goetzmann and Park (2001) use the Cox model for the first time to analyze hedge fund failure. They find that performance, risk and fund age are significant in the fund failure. Their measure of the "fund risk" is the standard deviation of the fund return over twelve months before the termination. The higher the standard deviation, the higher the hazard rate of the fund. In a following paper Gregoriou (2002) uses a similar model and argues that performance, size and leverage can be used to predict the survival of hedge funds.

2.5 Hypotheses

In order to determine the factors that drive the probability of hedge fund failure, we define some characteristics (or attributes) for every hedge fund. Some potential factors deemed to be significant include Financial Leverage as measured by the Debt/Equity Ratio, Portfolio Composition Type (whether the fund invested in primitive financial assets as opposed to derivatives and structured products), Use of margins (whether the hedge fund is leveraged through margins), Ownership (whether the hedge fund is independent or owned/backed by larger banks or financial institutions), Registered off-shore head office (or on-shore), Incentive fees and Management skills, Reputation and credibility (to capture management fraud or misbehavior).

In this section, we present the hypotheses on the four key factors that are expected to increase probability of failure during financial distress times.

2.5.1 Hedge Fund Failure and Leverage

The risk associated with high leverage in the hedge fund industry has been expected to be significant, mainly after the high financial leverage of the failed LTCM. For a hedge fund, if the leverage is too high, the potential losses can seriously exhaust and even wipe out the hedge fund's assets. Therefore, a high leverage is expected to increase the risk of forced liquidation or bankruptcy. Measuring financial leverage, however, is not a direct task. A fund can utilize leverage through different venues. There are three main sources of financial leverage: (a) First, the direct financial leverage ratio, measured as debt/equity is one source of financial leverage. (b) Second, a hedge fund can employ leverage by using margins. (c) Third, a hedge fund indirectly uses leverage by holding investment positions such as derivative products, futures, or short sales. [13]

In this research, using TASS hedge fund database, the average leverage ratio as measured by Debt/Equity is observed for each fund. We also observe binary variables for the derivatives, futures, and margins which to some extent provide the information needed. Additionally, we like to focus on the "excess leverage" employed by hedge funds and as such, at each point in time we can also define a demeaned measure of leverage as below:

$$ExcessLeverage_{i,t} = FinLeverage_{i,t} - \frac{\Sigma_{n=1}^{N}(Leverage_{n,t})}{N} \tag{2.4}$$

Evaluating a comprehensive leverage value for a fund as explained above, can be more precise when access to rich private databases such as proprietary hedge

[13]Note that (b) and (c) can be related in many cases.

fund holdings is available.[14]

H1: Hedge funds with higher financial leverage ratio are more likely to fail during financial distress times.

This hypothesis is very well-backed by the case example of LTCM. We like to see whether leverage ratio has a direct effect on the risk of failure or not. In other words what we expect is that the higher the leverage ratio, the more probable the hedge fund will fail in distress times.[15]

2.5.2 Hedge Fund Failure and Asset Types

This hypothesis looks into the relationship between financial asset types owned by hedge funds and the risk of failure.

H2: Hedge funds with larger investment portion in derivatives positions and futures (as opposed to primary financial asset ownership such as stocks, bonds and private assets) are more likely to fail during financial distress times.

Apart from being one source of leverage, another rationale for this expectation is related to the fact that a primary financial asset has higher collateral value compared to derivative products, most of which are time-decaying contracts with

[14]For example, Ang, Gorovyy and Inwegen (2011) use a private database on a limited HFs and measure the comprehensive leverage precisely for each fund based on detailed confidential data provided by managers.

[15]We later include a binary variable as an indicator to isolate the impact of distress.

expiration dates.[16] For example, consider options as opposed to stocks. In cases of financial distress products like options or futures cannot retain their original value; however, a stock or a bond ownership is still valuable though with significantly reduced asset value. Derivative products also tend to be a decaying security with expiration period. As such, we expect to see higher ownership in derivatives compared to pure assets among distressed hedge funds that fail.

It is noteworthy to mention that since derivative securities tend to have huge implied leverage, we expect to find a relation between hypothesis 1 and 2. As such an interaction term will also be tested for the mutual impact of leverage and use of derivatives and structured products.

2.5.3 Hedge Fund Failure and Margins

This hypothesis explains the role of margins on failure by relating failure of the hedge funds to financial leverage through use of margins.

H3: Hedge funds which employ margins as a source of leverage are more likely to fail during financial distress times.[17]

The logic behind this hypothesis is that normally hedge funds which use margins will be provided with additional leverage which is also more volatile during distress times. This may interrupt the fund managers' risk management and imply

[16] Forwards and futures (and other linear instruments) do not exhibit time decay.

[17] Margins at first glance may be perceived as simply another form of leverage. However, during distress periods margins have the tendency to increase, making them a more risky form of leverage.

forced liquidation of the investments to meet the required (increased) margins in the distress time. The result of this hypothesis will shed light on this point.

Comparing Hypotheses 1 and 3, we should also test for the potential interaction between these two factors. As such, an interaction term must be added to the model. This allows for nonlinearities in our econometric model which uses the product of explanatory variables. We test for the presence of these nonlinearities by examining the significance of the interaction term's estimated coefficient. If it is significant, the interaction term is needed to capture the relationship.

2.5.4 Hedge Fund Failure and Incentive Fees

A large number of hedge funds charge incentive fees on top of MER or management fees. The incentive fee (also called performance fees) encourages the manager to take on more risky investments in order to improve his share of the payoff.[18]
The hypothesis below will investigate this question.

H4: The hedge funds with higher incentive fees, are more likely to fail in financial distress times.

Higher incentive fees will induce management to invest in risky investments. Such funds are more likely to fail during financial crisis.

[18] In this sense, incentive fees act like executive stock options.

2.5.5 Hedge Fund Failure and Off-shore Head Office

A large number of hedge funds that are active in United States register their head offices off-shore in order to save on taxes. This protection however, can have a two-sided effect on the failure of hedge funds. Off-shore registered funds can save on income taxes (for the manager(s)) and hence have additional value that can be used in distress times. Moreover, off-shore registration might result in relatively larger risk-taking and aggressive borrowing (that is perceived to increase the risk of failure). Hedge funds register off-shore mainly due to regulations differences. In 2009, about 60% of all US hedge funds were registered off-shore (according to TASS database). Using the same database, we show that this number was risen to 63% at the end of 2011. Brown, Goetzmann, and Ibbotson (1999) perform a study on off-shore vs. on-shore funds and find that off-shore funds as a group have positive risk-adjusted performance when measured by Sharpe ratios and by Jensen's alpha.

Moreover, Liang and Park (2008) examine the liquidity premium in the hedge fund industry for off-shore vs. onshore hedge funds and record that the illiquidity premium is higher for off-shore than onshore funds because of a higher correlation between share restrictions and asset illiquidity.

The two cases of Bear Stearn hedge funds failure, Bear Stearns High-Grade Structured Credit Strategies Master Fund, Ltd., and Bear Stearns High-Grade Structured Credit Strategies Enhanced Leverage Master Fund, Ltd., registered in the Cayman Islands are examples of hedge funds that were registered off-shore and requested bankruptcy protection in U.S. These two Bear Stearn funds began to fail in early 2007 due, in part, to volatility in the subprime mortgage market and by May 2007, the Funds had begun to suffer significant devaluations of their asset portfolios, leading to margin calls, which the Funds were not able to meet. [19]

To shed light on the above mentioned relationship, the hypothesis below can explain the relation between off-shore vs. on-shore registration of the fund and probability of failure by relating failure to the registration of the hedge fund's head office.

H5: The hedge funds registered off-shore are more likely to survive in financial distress times because of the management tax savings added to the fund's value to provide a cushion.

2.6 Methodology

The goal of this paper is to find the role of financial leverage in increasing the risk of hedge fund failure during financial crisis. The definition and estimation procedure of various forms of leverage will be described later. There are three main stages in order to perform the analysis. First, we need to identify significant factors that drive failure. This is done using the hazard model specification. Second, we need to compare these factors for two samples of failed hedge funds and survived hedge funds. This is done using a dichotomous dependent variable regression model such as probit model. Finally, we need to identify the crisis periods from non-crisis periods and set up matching groups of hedge funds for each period. Robustness checks by using the Heckman model as well as propensity

[19] See article "U.S. court denies federal bankruptcy protection to hedge funds for liquidation proceedings in Cayman Islands", Sep 2007 for more details.

score matching are performed in the final stage.

Using a panel data of failed hedge funds matched with survived hedge funds, the data comprises 3156 survived funds and 7076 defunct funds, and will cover the 1996 to 2011 period. (TASS database reports failed funds from 1996). The estimation period will be broken into two sub-periods. The pre-crisis period (1996 to June 2006) and the crisis period (July 2006 to December 2011) which is displayed in Figure 4 and will be further explained.

Hedge fund studies are usually subject to certain potential data biases such as the survivorship bias which should be fixed for. Survivorship bias is the result of exclusion of the returns of non-surviving hedge funds, which causes the reported hedge fund performance in that database to look better than the true actual hedge fund performance. However, since in this study we have both defunct funds database and live funds database eliminates the possibility of survivorship bias in our analyses.

2.6.1 Probit Model Specifications

The dichotomous dependant variable regression specifications will be utilized to test the various hypotheses in this paper. Since the observed data points are large enough, we can assume a normal distribution and select a probit regression model. A probit model is specified with the dependent variable "Failure" equal to 0 or 1, and the continuous independent variables $X_i(t)$'s are estimated in:

$$Prob(Failure_{i,t} = 1) = \Phi(X_i(t)^T \beta) \tag{2.5}$$

Here, x_i is a vector of the explanatory variables, β is the (vector of) parameter(s) to be estimated, and Φ is the normal cumulative density function. The following equation will be estimated in order to test hypotheses 1, 2, and 3.

$$\begin{aligned} Prob(Failure_{i,t} = 1) &= \alpha_{0,t} + \alpha_{1,t}(FINLEV)_{i,t} + \alpha_{2,t}(DERIV)_{i,t} \\ &\quad +\alpha_{3,t}(MARGN)_{i,t} + \alpha_{4,t}(INFEES)_{i,t} \\ &\quad +\alpha_{5,t}(OFFSHR)_{i,t} + \alpha_{6,t}(FINLEV)_{i,t}(DERIV)_{i,t} \\ &\quad +\alpha_{7,t}(FINLEV)_{i,t}(MARGN)_{i,t}) \\ &\quad +\sum_{j=1}^{n}\gamma_{i,j}(ControlVar's)_{i,t} + \sum_{k=1}^{5}\beta_{i,k}D_{i,k} + \epsilon_{i,t} \end{aligned} \tag{2.6}$$

where $Failure_{i,t}$ is a binary variable equal to 0 if fund i failed by the time t and equal to 1 if fund i is in the live database at time t. Time t refers to time series observations (i.e. at different months) and $D_{i,k}$ is the dummy variable for hedge fund strategies. (Note that a hedge fund can be registered for more than one investment style).

In equation (2.6) above, FINLEV is the measure of financial leverage of the fund and the probability of failure is expected to be positively related to leverage; FINLEV can be obtained in three ways, including direct Debt/Equity ratio, use of margins, and derivatives. DERIV is the portfolio composition (or asset types) of the fund in the form of derivatives or futures; and we expect to find a positive relationship between portfolio composition and probability of failure. MARGN is the margin indicator (whether uses margin as a source of leverage or not), INFEES is the % incentive fees and OFFSHR is a dummy variable that shows whether the fund is registered off-shore or not. A set of control variables are also used as well as dummy variables for the strategy of the hedge fund. The control variables include a vector of cross-sectional factors such as fund size, fund age, volatility, prior performance, high water mark, MER, and risk.

Currently we use regression models with dichotomous dependent variables, probit model due to large number of observations. In the literature, logit, probit and Cox models have also been widely used.[20] The choice of probit model is a result of dealing with a large panel dataset since the underlying distribution tends to converge to normal (law of large numbers), and as such the dependent variable will be best tested with the probit model. For robustness we repeat the tests with logit and the results are consistent.[21]

2.6.2 Propensity Score Matching, Choice of Time Periods and Stability Tests

We would like to study the hedge funds failure during financial crisis, compared to non-crisis periods. In order to identify the role of leverage into why some hedge funds failed and why some of them survived the financial crisis period, we need to build matching samples of failed funds and survivors across the crisis and non-crisis periods. The observed data is a panel data set that includes monthly observations of the hedge funds characteristics and financial attributes. Our study examines a comprehensive sample of hedge funds, (databases will be discussed in the next

[20] Liang (2000) and Malkiel and Saha (2005) use the probit model, while Chan, Getmansky, Haas and Lo (2005) use the logit model. Lunde, Timmermann and Blake (1999) argue that since the probit model requires strong parametric as well as distributional assumptions, the Cox model adopts a more flexible approach. Based on a large data points assumption, we use the probit model.

[21] In robustness checks, the same regressions were performed using the logit model and the results stay the same. For the Cox model, it is expected that the findings are robust irrespective of the statistical method used.

section) over the January 2000 to December 2011 time period.

Figure 2 shows the level of S&P 500 Composite Index during this period. As we can see, the period starting from Oct 2007 until March 2009 shows sharp fall of the price level. But the crisis period that we like to study does not terminate in 2009. The reason is that comparing this figure with figure 1 we can see that failure of the hedge funds continues well after 2009 and prolongs into 2010. Hence, a rational conclusion is that the effect of the financial crisis on the failure rate of the hedge funds can be more prominent than the effect observed during normal periods.

As such, we split our time period into two sub-periods of Jan 2000 to Jul 2007 and Jul 2007 to Dec 2010. (The set up of these sub-periods may vary based on stability tests in other areas of financial crises studies, but our indicator is bankruptcy of Bear Stearn hedge funds.) Figure 4 displays this timeline. The two sub-periods will form the test of stability to see whether the impact of the determinants changed during the crisis period. After identifying the crisis period from non-crisis period we need to focus on the matching procedure. The matching procedure will be based on many characteristics. A matching sample is utilized because we are examining the effects of different structures of hedge funds on risk of failure (i.e. two different groups of failed hedge funds and survivors).

The propensity score matching (PSM) technique enables us to reduce large biases as much as possible by using a wide range of matching variables. The algorithm for PSM includes three main steps: First, run logistic regression on the dependent variable with appropriate conditioning (instrumental) variables and obtain the propensity score. Second, match each participant to one or more non-participants on propensity score by using techniques such as "Nearest neighbor matching", "Caliper matching", "Stratification matching" or Difference- in-differences matching (kernel and local linear weights). Finally, we run the multivariate analysis based on new sample using appropriate analysis for non-independent matched samples.[22]

The matching characteristics include a range of variables such as Strategies Effect (We use dummy variables D_i to adjust for the investment strategy effect.), Average Size (Average AUM during the previous year is used to measure the size of a fund), Size volatility (Standard deviation of the funds AUM is used to measure the size volatility of a fund), Age (Months of fund's age); We also include dummy variables to specify funds with high water mark (HWM), and a lockup provision.[23]

Finally, we need to run stability tests to compare the difference in results across different time periods. As such, we divide the full period into two sub-samples and use a Chow test to test for stability of regression coefficients between two periods. The value of the Chow statistic as well as the associated F-value are calculated. The null hypothesis, that there is no structural change over the full period, is rejected when the Chow statistic is greater than the associated F-value.

[22]For more details please see: Dehejia, R. and S. Wahba (2002).

[23]It is noteworthy to discuss the cons of using propensity score matching algorithm as well: (a) Data shrinkage - the process reduces the dataset since we eliminate the non-matched firms. (b) The resulting dataset is vulnerable to choice of variables/parameters for matching. - For robustness, I perform the regressions both before and after the PSM.

2.7 Data

Various database searches have been performed to identify the most comprehensive data sources for this research. For this research we need both the live database and the defunct (or graveyard) database. TASS Hedge Fund database provides such data.

As an example, Bali, Brown, and Caglayan (2011) employ data on hedge funds to evaluate different measures of systematic risk that can affect hedge funds returns. The hedge fund data for their study is provided by Lipper TASS database, and it contains information on a total of 14,228 defunct and live hedge funds. The TASS database, in addition to reporting monthly returns (net of fees) and monthly assets under management, also provides information on certain fund characteristics including returns, size, age, management fees, and incentive fees. Their paper uses only one hedge fund database and this will be the primary source of data for my study as well. Obtaining all the required data from one source eliminates most of the biases associated with overlaps and back-fills.

2.7.1 TASS Hedge Funds database

A total of 17,202 hedge funds reported monthly returns to TASS for the years between Jan 1996 and Dec 2011 in this database. From this number, 10,455 are defunct funds and 6,747 are live funds. Hedge funds in this database report their monthly AUMs in their primary currency value. As such, I eliminate any funds that report their AUMs in a different currency than USD. For the specific tests in this paper that involves AUMs analysis, the data is limited to USD-reported funds. A total of 47% of the Live Funds (3,156) report their AUM in USD currency, while 69% of the Failed Funds (7,076) report in USD currency.

This means by removing non-USD-currency funds our total sample consists of 7076 Failed and 3156 Survived firms. The question that comes to mind next is whether this removal will create any bias in our sample, especially for tests of off-shore (i.e. HFs registered in other than United States). However, based on Appendix B, it can be seen that even after focusing only on USD funds, there is a consistent variation in the distribution of HFs among different registered domicile countries. Another important observation, as seen in Appendix C, is that about 46.5% of all the hedge funds in the sample are registered in either Cayman Islands or British Virgin Islands which supports testing the hypothesis on off-shore registered funds.

2.7.2 Potential Data Biases and Mitigation

Hedge fund studies are usually subject to certain potential data biases such as the survivorship bias which should be addressed. Survivorship bias is the result of exclusion of the returns of non-surviving hedge funds, which causes the reported hedge fund performance in that database to look better than the true actual hedge fund performance. However, since in this study we have both defunct funds and live funds in our database, the possibility of survivorship bias in our analysis is eliminated.

Moreover, TASS database reports the reason why the fund is dropped. As such, we can distinguish between failure and attrition.[24] The reasons for failure or exit from the live database are recorded in Lipper TASS database by assigning a status code to the funds in defunct (or graveyard) database. These codes include liquidation, no longer reporting, unable to contact, closed to new investment, and merged into other equity.

Another common data bias while working with hedge fund databases is the back-filling bias. Back-filling bias exists since once a hedge fund is included in the database, previous returns are automatically added. This increases the chance that only successful hedge funds may report the true previous returns to the database. To overcome this, we use the advantage that TASS database reports both "when added" and "first reported performance date". The first reported performance date is usually about 1 year later than first return. As such, there are two mitigation plans for correcting back-filling biases. We can either delete the first year of returns or employ the return bias that has been empirically studied into account.[25]

The other source of bias is the self selection bias. This is due to the fact that reporting is voluntary, so a bad fund has less or no incentive to report. At the same time, a fund that is very successful may close quickly and there is no further incentive to attract new investors by reporting (perceived advertising[26]). In their work, Fung and Hsieh (1997b) however, claim that these effects offset each other, and should not create a bias in the empirical tests. Nevertheless, this bias can be studied later in our study.

[24] See Fung and Hsieh (2009) for details on "real" failure.

[25] For example, empirical studies show that there exists approximately 1.5% bias in the returns of the back-filled data. See Fung and Hsieh (2009) and Bali, Brown, and Caglayan (2011)

[26] Titman and Tiu (2011)

2.8 Empirical Results

2.8.1 Descriptive Statistics

Tables 1 and 2 summarize the data used in this study. There are two types of characteristic variables available in TASS database. First are the continuous variables which vary in value. These variables include but are not limited to “Average Leverage”, “Management Fees”, and “Incentive Fee”. Second set of characteristics variables are the binary or dichotomous variables including but not limited to “Personal Capital” (Do principals have money invested?), “Margin” (Does the fund leverage using margin borrowing value?), “Derivatives” (Does the fund leverage using derivatives value?), “Futures” (Does the fund leverage using futures value?), and “Leveraged” (Does the fund use leverage?). Table 1 shows these variables for the two subsets of live and defunct databases.

Table 3 and Figures 1 and 3 display the time series features of our data set. Figure 1 reports the number of hedge funds that failed and the number of hedge funds that survived over time. Figure 3 displays a bar chart that show the cumulative AUM sustained in this industry over time.

Based on Tables 2 and 4, and Appendices C and D, categorical characteristics of the hedge funds can be compared across failed funds and survived funds. These categorical groupings include “Primary Category” of the fund, or investment style (Table 4), “Domicile Country” or registered headquarter (Appendix C) and “Base Currency” or primitive investment currency for the fund’s investments (Appendix

D). It can be seen that comparing to survived funds, a larger ratio of failed hedge funds were registered in Cayman Islands and Virgin Islands, indicating a higher probability of failure for such off-shore registered funds.

Table 5 provides pair-wise correlation coefficients between different fund characteristics.

2.8.2 Asymmetric Correlation Results

One key question in this paper is to uncover the relative difference in the drivers of HFs' failure during crisis vs. non-crisis times. To give some context to the hypotheses surrounding the crisis period, Table 6 shows the correlation between monthly returns of the HFs and monthly returns of the S&P500 index (proxy for the market) for each Investment Strategy category. Overall, the correlation increases both in value and significance during the crisis period. This is more prominent for certain categories.

This phenomenon is known as "asymmetric correlation" in equity portfolios, or bigger downside correlation during crisis periods. Ang and Chen (2002) show that correlations between stocks and aggregate market are much larger for downside moves, especially for extreme unexpected downside moves, than for upside moves (See Figure 2 for the level of S&P 500 Composite Index). This fact was also empirically evident during the recent financial crisis. In their paper Ang and Chen suggest use of regime switching models which provide greater flexibility in capturing these sudden moves. In this paper, we split the sample data with a perceived regime switch and run stability tests to identify the difference in the performance and magnitude of hedge funds' failure drivers among these sub-samples. Figure 4 displays this timeline and the perceived regime switch. Brunnermeier and Pedersen (2009) also suggest that the start of the recent crisis period should be identified as January 2007. They separate the time period into crisis, and recovery periods. In this study, we are mostly interested in identifying the beginning of the period.

2.8.3 Role of Leverage and Regression Results

The preliminary results indicate that the factors that increase probability of failure have different significance and impact during the crisis period vs. non-crisis period, which is consistent with the hypothesis being tested. Table 7 shows the regression results for average leverage, as well as leverage induced by use of margins and derivative products, controlling for the rest of the factors, including size, prior performance, and the fund's characteristics. This regression is run on three different models and for before and after the crisis period.

The first test uses a "crisis" dummy variable on all-time sample to test whether the difference in propensity of failure is statistically significant in crisis and prior non-crisis periods. As it can be seen, this dummy variable is significant at 1% across all three regression models being tested which approves the motive to move on to the two sub-sample tests. Next, is to find the economic difference between the role of leverage during crisis and non-crisis periods. The two sub-sample tests both confirm the significance of leverage during the crisis and lack of significance during non-crisis periods in driving failure of hedge funds. The results are shown in Table 8.

For the sub-sample tests, we split the interval into Pre-Crisis and, during and after-Crisis periods. We find that leverage is only significant during and after the crisis period and insignificant before. This is after controlling for size, performance, incentive fees, and off-shore location. This shows the impact of leverage is significant regardless of how big or how successful the hedge fund is. The results are robust after controlling for use of margins and derivatives.

2.8.4 Other Determinants of Failure and Regression Results

Apart from the results on the role of leverage, Tables 7 and 8 have interesting observations on other determinants of failure to note. First, the off-shore variable is significant and positive only during crisis period. It shows that funds that are registered off-shore are more likely to fail during and after the crisis period, but there is no difference among domicile country of the fund during non-crisis period. This can be due to the fact that as a result of difference in regulations and transparency, off-shore hedge funds aggressively invest in more risky portfolios which increase risk of failure during crisis.

Second, the funds that apply High Water Mark benchmark are consistently less likely to fail. We expect this is due to the fact that usually smarter funds who can tease out the market risk better, are also those who use a High Water Mark performance benchmark (Titman and Tiu, 2011).[27]

[27] Note that we do not investigate/show intelligence factor in this paper.

Third, the management fees are consistently insignificant. Regardless of which period, the management fee will not increase or decrease the propensity of failure of the hedge fund. This however, is different for the incentive fee. Incentive fees are significant and positive which signals that higher incentive fees increase the risk of failure of the hedge fund.

Last, it is important to note the role of size. Size is significant only during the pre-crisis period. The results indicate that larger funds are more likely to survive in pre-crisis period. However, as soon as the crisis period begins and during and after crisis the size cannot determine the probability of failure and becomes insignificant. This once again proves that the "too big to fail" notion does not apply during the crisis period, especially during the systemic risk period.

Prior performance of the fund, as measured by lagged return of the fund, is consistently significant in the results regardless of the time period. This shows that failure of the hedge funds consistent with other financial institutes is a decaying process which can well be captured and tested by the use of a hazard model.

2.8.5 Hedge Funds Strategies and Regression Results

To find the key hedge fund styles that are more prone to fail, we run the regression on 12 different Investment Strategy groups. Overall, leverage is significant in driving probability of failure across indicated Investment strategies. The results for these strategies are shown in Table 9. The z-values of one or more forms of leverage models are consistently significant and positive, indicating that higher leverage of the fund may not necessarily increase the risk of failure during the non-crisis periods. These styles include Emerging Markets, Long/Short Equity Hedge, Equity Market Neutral, and Multi-Strategy hedge funds.

We further investigate the commonality of these four styles with significant probability of failure due to leverage. To do so, we measure the average beta across each of the investment styles. We find that since these investment styles focus mostly on equity market, they have relatively higher betas, and hence larger exposure to the total equity market. The higher beta risk can be one reason that makes these specific hedge fund categories vulnerable to the use of leverage during the financial crisis period.[28]

We report the regression results for the rest of the styles in Tables10a and 10b. It is interesting to note that some of these hedge fund categories, such as "Managed Futures" and "Options Strategy" tend to have portfolios that cover more than equities, for example they invest in currencies, commodities, etc. This diversification allows them to survive during crisis periods. Note that prior performance continues to be an important indicator of all funds' failure regardless of the particular style being followed, indicating persistence in hedge funds performance.

We report the results of probit regression, however, we run both logit and probit regressions for the entire analysis and the results are robust regardless of which model we use. Because the number of observations is large enough, it is

[28]To measure beta return, remember that as explained in section 2.3.2, we split return on each investment style into three components of alpha, beta, and management fees. Following Ibboston (2006), the average systematic beta return for these four styles is calculated at 6.42, which is also higher than the average non-significant investment styles.

recommended to use probit regression since the underlying distribution is close to a normal distribution.

2.9 Conclusions

We find that during crisis periods the average financial leverage turns out to be more significant in increasing probability of failure. This is by far the first attempt to disentangle the impact of hedge funds failure determinants during crisis time from non-crisis time and the conclusion that hedge funds with less leverage can survive the crisis period better. This finding is consistent with the literature on the role of leverage in increasing overall risk of the firm.

Leverage, as defined by the ratio of debt/equity is not the only source of leverage for the funds. The use of derivative products also implies the usage of leverage. Moreover, the use of margins is another indicator of the leverage risk for the funds. Furthermore, the excess leverage at any point in time can be defined and measured as the deviation from the average leverage across all funds. We find our findings robust with respect to the definition of leverage.

The findings of this paper are also in-line with the recent theoretical works on the role of leverage in systemic risk and the policy implications. Schwarcz (2011) assesses the progress of identifying and managing systemic risk (through recent policy advances) and indicates that one way that Dodd-Frank attempts to avoid the need to make emergency loans caused by financial crisis is by requiring banks and other designated "systemically important" financial firms (such as hedge funds) is to enforce a range of capital, leverage, and liquidity requirements and periodic "stress testing."

This paper is also another empirical support for the work of Brunnermeier and Pedersen (2010) on the market liquidity and funding liquidity. The fact that the role of leverage is highly significant and increases the probability of failure during the crisis period can also explain how forced liquidation may be triggered by hedge funds' margin increase during bear markets. Hedge funds are generally compensated for investments in illiquid assets and the leverage together with illiquid holdings plays an important role in hedge funds' risk of failure.

Finally, we also show that after controlling for crisis periods, certain investment styles prove to show a more prominent and significant impact of leverage on the fund's risk of failure. Hedge funds registered under "Emerging Markets", "Equity Market Neutral", "Long/Short Equity Hedge", and "Multi-strategy" investment strategies continue to be subject to larger and more significant risk of leverage in their lifetime both before and after the crisis and distress periods. These funds have higher than average betas exposing them to significant equity market risk. In contrary, hedge funds styles that have portfolios including variety of assets in addition to equities (e.g. currencies, commodities, etc.) tend to survive.

2.10 References

Ackermann, C., McEnally, R., and Ravenscraft, D., 1999, The Performance of Hedge funds: Risk, Return, and Incentives, Journal of Finance 54, 833-874.

Adrian, T., 2007, Measuring Risk in the Hedge Fund Sector, Current Issues in Economics and Finance, Vol. 13, No. 3, Federal Reserve Bank of New York.

Agarwal, V., and Naik N., 2004, Risks and portfolio decisions involving hedge funds, Review of Financial Studies 17, 63-98.

Agarwal, V., N. D. Daniel, and N. Y. Naik, 2009, Role of Managerial Incentives and Discretion in Hedge Fund Performance, Journal of Finance 64 (5), 2221 - 2256.

Alexander, G., and Baptista, A., 2003, Portfolio Performance Evaluation Using Value at Risk, The Journal of Portfolio Management, Summer 2003, 93-102.

Alexiev, J., 2005. The impact of higher moments on hedge fund risk exposure. The Journal of Alternative Investments, 7(4), 50-65.

Amin, G. and H. Kat, 2003a, Hedge Fund Performance 1990-2000: Do the Money Machines Really Add Value? Journal of Financial and Quantitative Analysis 38, 251-274.

Amin, G. and H. Kat, 2003b, Welcome to the Dark Side: Hedge Fund Attrition and Survivorship Bias over the Period 1994-2001, Journal of Alternative Investments 6, 57-73.

Andersen, P. K., Borgan, O., Gill, R.D., Keiding, N, 1992, Statistical Models Based on Counting Processes, Springer-Verlag.

Andersen, P. K. and Gill, R. D. 1982, Cox's Regression Model for Counting Process: A Large Sample Study, Annals of Statistics, 10, 1100-1120.

Ang, A. and J. Chen, 2002, Asymmetric Correlation of Equity Portfolios, Journal of Financial Economics, 63 (3), 443-494.

Ang, A., Gorovyy S., and Inwegen G., 2011, Hedge fund leverage, Journal of Financial Economics, 102(1), 102-126.

Aragon, G. O., 2007, Share Restrictions and Asset Pricing: Evidence from the Hedge Fund Industry, Journal of Financial Economics 83 (1), 33 - 58.

Artzner, P., Delbean, F., Eber, J.-M., and Heath, D., 1999, Coherent Measure of Risk, Mathematical Finance, Vol. 9, No. 3, 203-228.

Bali, T. G., Gokcan, S. and Liang, B., 2007, Value at Risk and the Cross-Section of Hedge Fund Returns, Journal of Banking and Finance, 1135-1166.

Bali, T.G., Brown S.J.,Caglayan M.O., 2011, Do hedge funds' exposures to risk factors predict their future returns? Journal of Financial Economics.

Baquero, G., Horst, J., and Verbeek, M., 2005, Survival, Look-Ahead Biases and the Persist in Hedge Fund Performance, Journal of Financial and Quantitative Analysis, Vol. 40, No. 3, 493-517.

Bollen, N. P. B. and V. K. Pool, 2008, Conditional Return Smoothing in the Hedge Fund Industry, Journal of Financial and Quantitative Analysis 43 (2), 267-298.

Bondarenko, O. (2004, March). Market price of variance risk and performance of hedge funds. AFA 2006 Boston Meetings Paper.

Brav, Alon , Wei Jiang, Frank Partnoy, and Randall Thomas, 2008, Hedge fund activism, corporate governance, and firm performance, The Journal of Finance 63, 1729-1775.

Brown, Stephen J., William N. Goetzmann, Roger Ibbotson, 1999, Offshore hedge funds: survival and performance, Journal of Business 72, 91-117.

Brown, Stephen J., William N. Goetzmann, Roger Ibbotson and Stephen A. Ross, 1992, Survivorship Bias in Performance Studies, The Review of Financial Studies, Vol. 5, No. 4, 553- 580.

Brown, Stephen J., William N. Goetzmann, and James Park, 2001, Careers and survival: Competition and risk in the hedge fund and CTA industry, Journal of Finance 56, 1869-1886.

Brunnermeier, M. K., and Pedersen, L. H., 2009. Market liquidity and funding liquidity. Review of Financial studies, 22(6), 2201-2238.

Chan N., M. Getmansky, S. Haas, and A. Lo, Systemic Risk and Hedge Funds, prepared for the NBER Conference on the Risks of Financial Institutions, August 2005, p. 51.

Christoffersen, P. F. and Goncalves, S., 2005, Estimation Risk in Financial Risk Management, Journal of Risk, v 7, 1-28.

Christory C., Daul S., and Jean-Rene Giraud, Quantification of Hedge Fund Default Risk, Journal of Alternative Investments, fall 2006, pp. 71-86.

Clifford, Christopher P., 2008, Value creation or destruction? Hedge funds as share- holder activists, Journal of Corporate Finance 14, 323-336.

Cox, D. R., 1972. Regression Models and Life Tables, Journal of the Royal Statistical Society, Series B, 20, 187-220.

Cox, D. R., 1975. Partial likelihood. Biometrika, 62(2), 269-276.

Cremers, Kritzman and Page, 2005, Optimal Hedge Fund Allocations: Do Higher Moments Matter?, Journal of Portfolio Management, v 31, i3, p70(12).

Dehejia, R., Wahba, S., 2002, Propensity Score Matching Methods for Non experimental Causal Studies, Review of Economics and Statistics 84 (1) 151-161.

Dungey, M., Fry, R., González-Hermosillo, B., Martin, V., 2006, Contagion in international bond markets during the Russian and the LTCM crises, Journal of Financial Stability, v 2, i1, 1-27.

Edwards, Franklin R., Hedge Funds and the Collapse of Long-Term Capital Management, The Journal of Economic Perspectives , Vol. 13, No. 2 (Spring, 1999), pp. 189-210, Published by: American Economic Association.

Fung, William, and David Hsieh, 1997, Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds, Review of Financial Studies, 10, 275- 302.

Fung and Hsieh, 2000a, Performance Characteristics of Hedge Funds and CTA Funds: Natural versus Spurious Biases, Journal of Financial and Quantitative Analysis, 35, 291-307.

Fung and Hsieh, 2000b, Measuring the Market Impact of Hedge Funds, Journal of Empirical Finance, 7, 1-36.

Fung and Hsieh, 2001, The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers, Review of Financial Studies, 14, 313-341.

Fung and Hsieh, 2002a, Benchmarks of Hedge Fund Performance: Information Con- tent and Measurement Biases, Financial Analyst Journal, 58, 22-34. Fung and Hsieh, 2002b, Asset-based Style Factors for Hedge Funds, Financial Analyst Journal, 58, 16-27.

Fung, W. and D. Hsieh, 2009, Measurement Biases in Hedge Fund Performance Data: An Update, Financial Analysts Journal 65, 36-38.

Getmansky, Mila, Andrew Lo, and Igor Makarov, 2004, An Econometric Model of Serial Correlation and Illiquidity in Hedge Funds Returns, Journal of Financial Economics 74, 529-610.

Getmansky, Mila, Andrew Lo, and Shauna Mei, 2004, Sifting through the Wreckage: Lessons from Recent Hedge-fund Liquidations, Journal of Investment Management, 2(4), fourth quarter, 6-38.

Goetzmann, W., Ingersoll, J. and Ross, S., 2003, High-water Marks and Hedge Fund Management Contracts, Journal of Finance, v 58, n 4, 1685-1717.

Gregoriou, G. N., 2002. Hedge Fund Survival Lifetimes, Journal of Asset Management, Vol. 2, No. 3, 237-252.

Griffin, John M., and Jin Xu, 2009, How smart are the smart guys? A unique view from hedge fund stock holdings, The Review of Financial Studies 22, 2531-2570.

Gupta, Anurag, and Bing Liang, 2005, Do Hedge Funds Have Enough Capital? A Value at Risk Approach, Journal of Financial Economics 77, 219-253.

He, Q., Chong, T. T.-L., Li, L. and Zhang, J., 2010, A Competing Risks Analysis of Corporate Survival. Financial Management, 39: 16971718.

Heckman, J. 1979, Sample selection bias as a specification error. Econometrica 47 (1): 153-162. Jorion, Philippe, 2000b, Risk Management Lessons from Long Term Capital Management, European Financial Management 6, 277-300.

Jorion, P. (2000), Risk management lessons from Long-Term Capital Management. European Financial Management, 6: 277–300. doi: 10.1111/1468-036X.00125

Jickling, Mark and Raab, A., 2006, Hedge Fund Failures, CRS Report for U.S. Congress.

Jickling, Mark, 2009, Hedge Funds: Should they be regulated?, CRS Report for U.S. Congress.

Kabir, M. & Kabir Hassan, M. (2005). The Near-Collapse Of LTCM, US Financial Stock Returns And The Fed. Journal of Banking & Finance, Vol 29(2), pp. 441-460.

Kalbeisch, Jack D. and Ross L. Prentice, 2002, The Statistical Analysis of Failure Time Data, New York: Wiley.

Kosowski, Robert, Narayan Y. Naik, and Melvyn Teo, 2007, Do hedge funds deliver alpha? A bayesian and bootstrap analysis, Journal of Financial Economics 84, 229-264.

Lhabitant F., 2006, The Handbook of Hedge Funds. John Wiley and Sons, London.

Liang, B., 1999, On the Performance of Hedge Funds, Financial Analysts Journal 55, 72-85.

Liang, B., 2000, Hedge fund: the Living and the Dead, Journal of Financial and Quantitative Analysis 35, 309-326.

Liang, B., and Park, H., 2007, Risk Measures for Hedge Funds: A Cross- Sectional Approach, European Financial Management, Vol.13, No.2, 317-354.

Liang, B., and Park, H., 2008. Share restrictions, liquidity premium and offshore hedge funds. Unpublished working paper.

Liang, Bing and Park, Hyuna, 2010. Predicting Hedge Fund Failure: A Comparison of Risk Measures. Journal of Financial and Quantitative Analysis, 45, pp 199-222.

Lo, A., 2001, Risk Management for Hedge Funds: Introduction and Overview, Financial Analysts Journal 57, 16-33.

Lo, A., 2010, Hedge funds: An analytic perspective. Princeton University Press.

Lunde, A., Timmermann, A., and Blake, D., 1999. The hazards of mutual fund underperformance: A Cox regression analysis. Journal of Empirical Finance, 6(2), 121-152.

Malkiel, B. G., and Saha, A. 2005. Hedge funds: risk and return. Financial analysts journal, 61(6), 80-88.

Mitchell, M. and Pulvino, T., 2001, Characteristics of Risk and Return in Risk Arbitrage, Journal of Finance, vol. LVI, No. 6, 2135-2176.

Schwarcz, Steven L., Keynote Address – Identifying and Managing Systemic Risk: An Assessment of Our Progress (March 16, 2011). Available at SSRN: http://ssrn.com/abstract=

Shumway, T., 2001. Forecasting bankruptcy more accurately: A simple hazard model*. The Journal of Business, 74(1), 101-124.

Sortino, F. and R. van der Meer, 1991, Downside Risk, Journal of Portfolio Management, summer, 27-31.

Taleb, N. N., 2004. Bleed or blowup? Why do we prefer asymmetric payoffs?. The Journal of Behavioral Finance, 5(1), 2-7.

Titman, S. and C. Tiu, 2011. Do the Best Hedge Funds Hedge?, Review of Financial Studies, 24(1): 123-168 first published online October 18, 2010.

2.11 Essay I Appendices

2.11.1 Appendix A: Complete Variables Glossary

In this section, the detailed definitions for each of the variables are provided.

- **Fund's Assets Under Management ($M):** The total value of a fund's assets under management (AUM).

- **Return (%):** Monthly net-of-fee returns.

- **Return Volatility (%):** Standard deviation of a fund's monthly returns over a calendar year. (This variable is winsorized at 5% and 95% to remove potential outliers.)

- **Crisis:** A dummy variable equal to one if the year is in the crisis period, which includes years 2001 and 2002 (the period when technology bubbles burst) and years 2008 and 2009 (the most recent financial crisis). This is used in robust check analysis.

- **Financial Leverage:** Equals the ratio of "Debt/Equity" for the fund and indicates the average leverage that the fund uses.

- **High Water Mark:** A dummy variable equal to one if the fund has a high water mark provision.

- **Incentive Fee (%):** A fund's incentive fee as a percent of fund assets.

- **Offshore:** A dummy variable equal to one if the fund is located offshore (anywhere but United States), and zero if the fund is located in the United States.

- **Size:** The natural logarithm of the value of the assets under management of a fund.

2.11.2 Appendix B: Complete Styles Glossary

The table below shows the list of hedge funds styles in our dataset.

Table 2.1: Hedge Funds Investment Styles
(based on TASS database)

Style Index	Style Name
1	Convertible Arbitrage
2	Dedicated Short Bias
3	Emerging Markets
4	Equity Market Neutral
5	Event Driven
6	Fixed Income Arbitrage
7	Fund of Funds
8	Global Macro
9	Long/Short Equity
10	Managed Futures
11	Multi-Strategy
12	Options Strategy
13	Other/Undefined

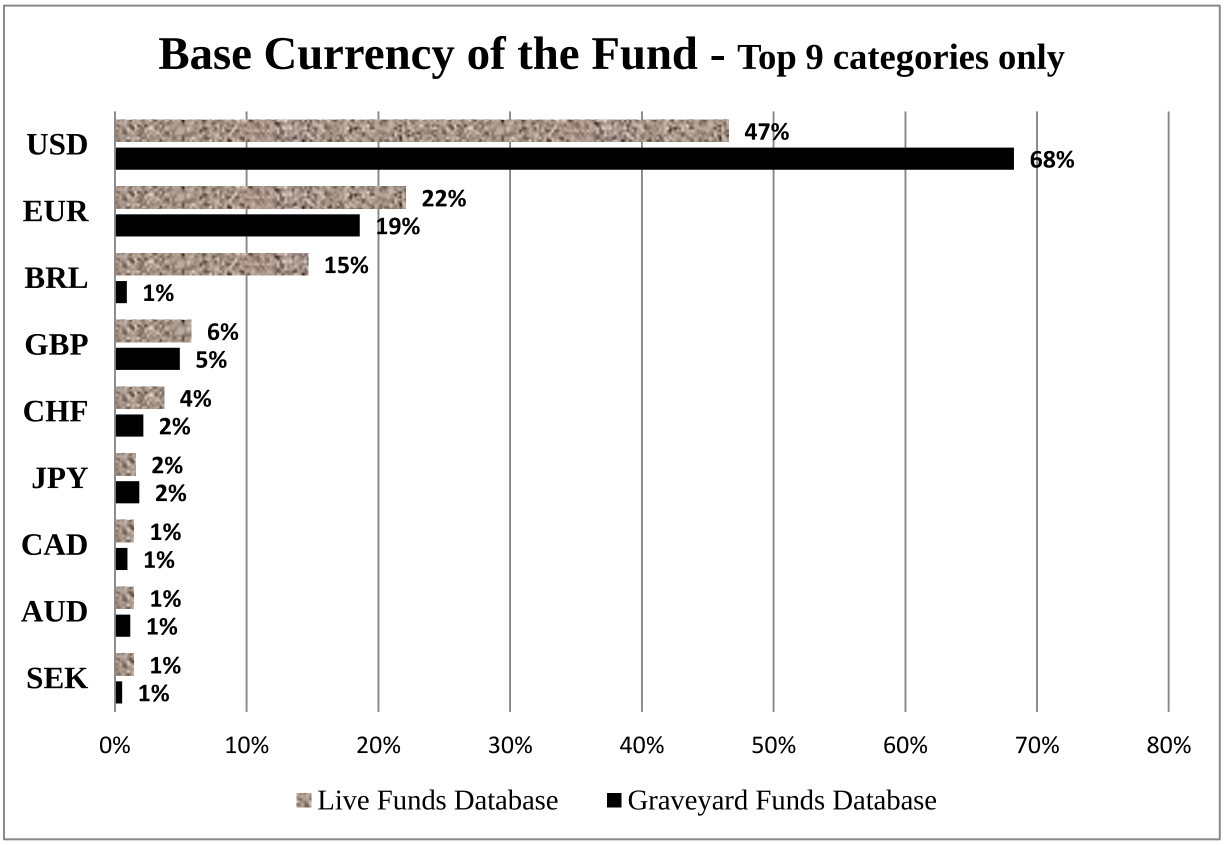
Appendix C. Domicile Country Distribution of All Funds with reported USD Currency

Note that 69% of the funds are registered off-shore.

Domicile Country	Count of Funds	%Total (10076)	Domicile Country	Count of Funds	%Total (10076)
Cayman Islands	3761	37.33%	Saint Martin	11	0.11%
United States	3150	31.26%	Australia	8	0.08%
Virgin Islands (British)	923	9.16%	Netherlands	7	0.07%
Bermuda	572	5.68%	Saint Vincent And The Grenadines	4	0.04%
Luxembourg	423	4.20%	Singapore	4	0.04%
Bahamas	254	2.52%	Bahrain	4	0.04%
Ireland	246	2.44%	Sweden	3	0.03%
Guernsey	224	2.22%	Hong Kong	3	0.03%
Jersey	96	0.95%	Malaysia	3	0.03%
Switzerland	72	0.71%	Israel	3	0.03%
Canada	56	0.56%	Anguilla	2	0.02%
Curacao	43	0.43%	United Arab Emirates	2	0.02%
Malta	36	0.36%	Saint Kitts And Nevis	2	0.02%
Mauritius	32	0.32%	China	2	0.02%
None	29	0.29%	Gibraltar	2	0.02%
France	26	0.26%	Japan	1	0.01%
Liechtenstein	23	0.23%	Saudi Arabia	1	0.01%
Isle of Man	17	0.17%	Virgin Islands (U.S.)	1	0.01%
United Kingdom	13	0.13%	Finland	1	0.01%

Appendix D. Distribution of base currency of the hedge funds in top 9 categories; presented in percent total. In this study we eliminate non-USD funds

There are total of 17,045 hedge funds in the TASS database. From this number, 9,648 are defunct funds and 7,397 are live funds. After eliminating non-USD funds, there are 3362 firms in the Live Funds database and 6250 in the Defunct Funds database.



2.12 Essay I Figures and Tables

Figure 1:

Time Trend of Failed Hedge Funds count and % Failed from.

Source: TASS Hedge Funds database

A total of 17,202 hedge funds reported monthly returns to TASS for the years between Jan-1996 and Jan-2012 in this database. From this number, 10,455 are defunct funds and 6,747 are live funds.

For each year from 2002 to 2011, the bar chart below reports the number of hedge funds that failed. The trend line shows the % failed, calculated as ratio of total failed HFs over total number of HFs in each year.

* Note that the distinction of the crisis period (as shown by the demarcation) starts from August 2007.

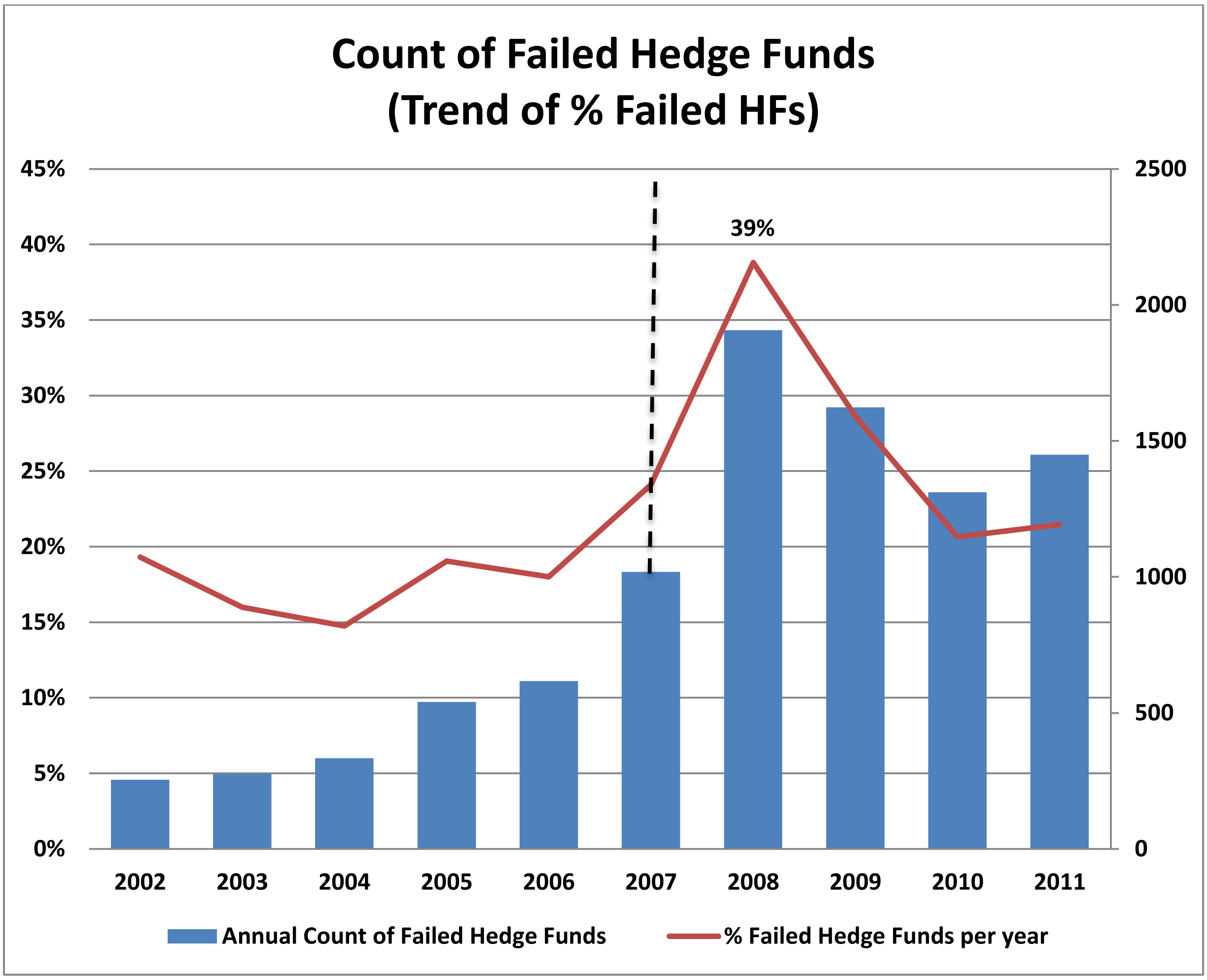


Figure 2:

Monthly Level on S&P 500 Composite Index over the period of Jan 2000 to 2011.
(Source: CRSP database)

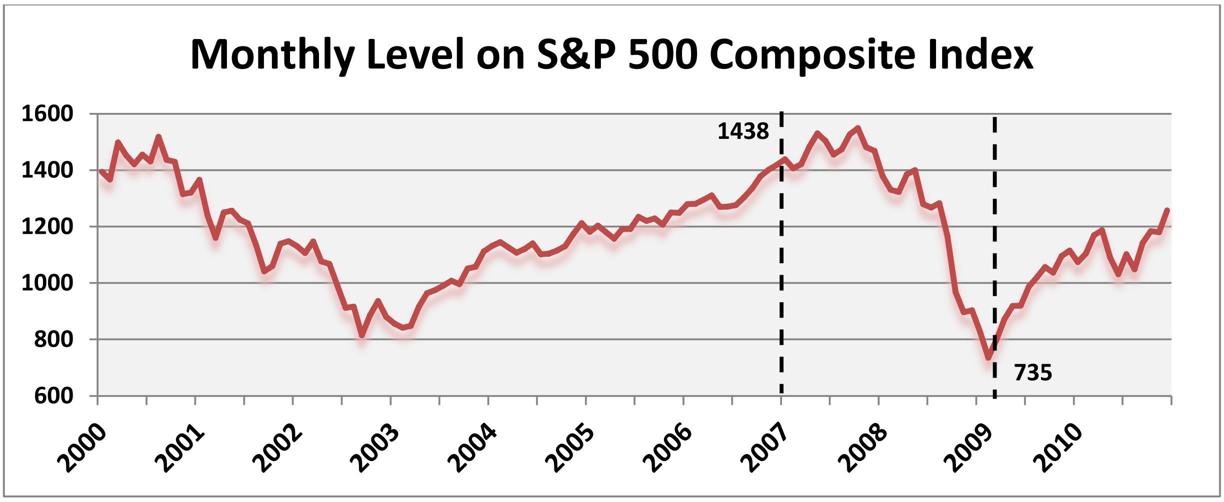
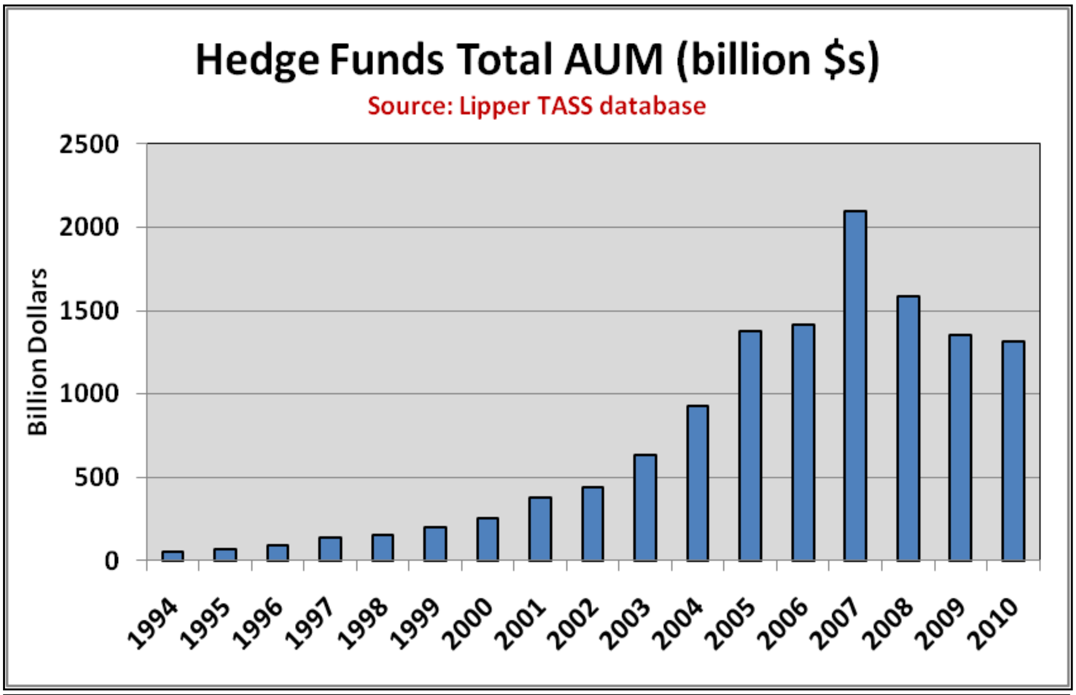


Figure 3:

3.A) Annual Asset Under Management (AUM) of Hedge Fund Industry over the period of 1994 to 2010.

(Source: Lipper TASS database)

3.B) Hedge Funds Average Weighted Leverage over the period of 2005 to 2012



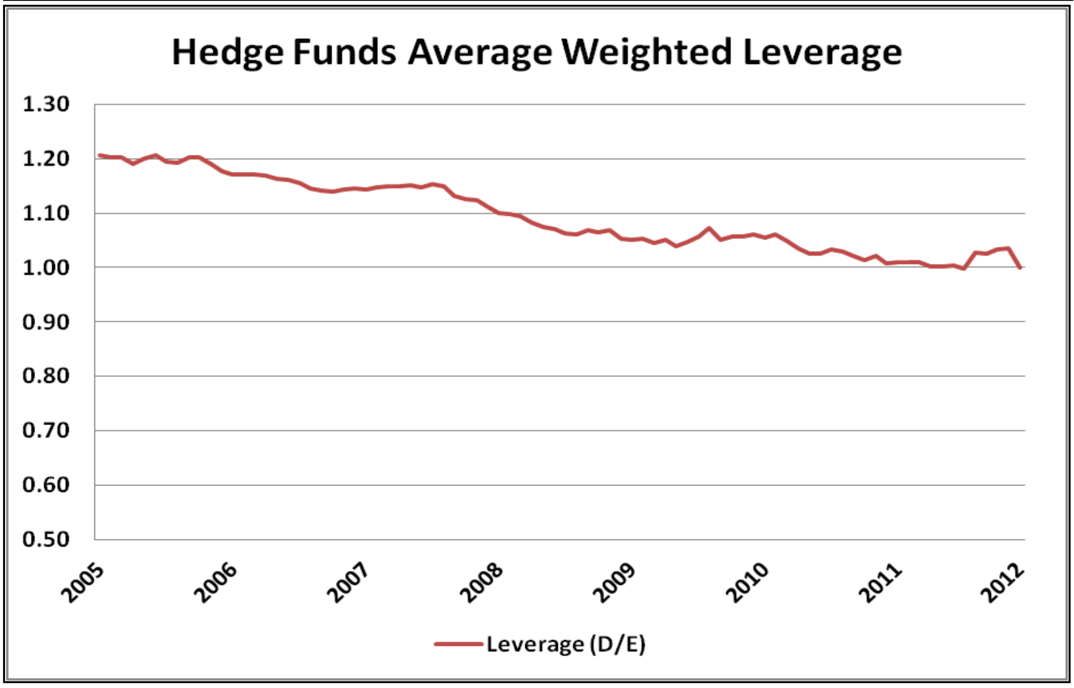


Figure 4: Timeline of Interval Selection for Regression Model:

We split our time period into two sub-periods of Jan 1996 to Dec 2006 and Jan 2007 to Dec 2010. (The set up of these sub-periods may modify based on stability tests.) The plot below displays this timeline. The two sub-periods form the test of stability to see whether the impact of the determinants changed during the crisis period.

(Failure of the hedge funds continues well after 2009 and prolongs into 2010. Hence, a rational conclusion is that the effect of the financial crisis on the failure rate of the hedge funds can be more prominent than the effect observed during normal periods.)

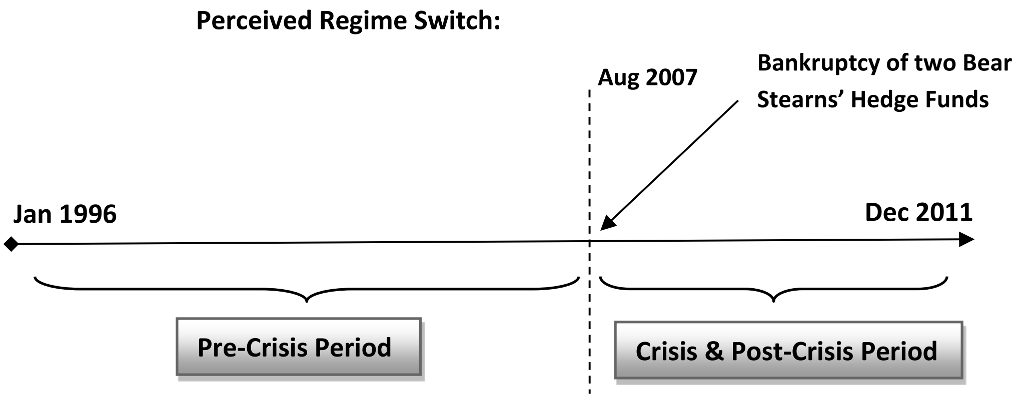


Table 1

Descriptive statistics: Cross-sectional

A total of 17,202 hedge funds reported monthly returns to TASS for the years between Jan-1996 and Dec-2011 in this database. From this number, 10,455 are defunct funds and 6,747 are live funds. After eliminating non-USD funds, we have 3156 Live Funds and 7076 Defunct funds. Table below reports the comprehensive characteristics of the funds for the two sub-data sets of Live and Grave yard (defunct) funds.

Significance level at the 1%, 5% and 10% is indicated as ***, **, and *, respectively. The test for difference in means is the independent t-test and the test for difference in median is the Wilcoxon test. The difference in mean (median) is defined as Live database minus Graveyard (Defunct) database.

	Live Funds			**Defunct Funds**			**Test for difference in Means**	**Test for difference in Medians**
	Mean	**Median**	**Std Dev.**	**Mean**	**Median**	**Std Dev.**	**t-stat**	**z-stat**
Funds Continuous Characteristics								
Average Actual Leverage (Debt/Equity)%	39.40	0.00	127.28	55.94	0.00	197.05	-21.67***	0
Max Leverage (Debt/Equity) %	85.69	0.00	204.78	101.25	0.00	293.30	-20.10***	0
Management Fees %	1.45	1.50	0.74	1.45	1.50	0.66	0	0
Incentive Fees %	14.62	20.00	7.93	15.67	20.00	7.64	-51.84***	0
Size (AUM)	17.63	17.71	1.79	17.68	18.37	1.74	-10.38***	-4.94***
Volatility (measured as return's Std. Dev.)	3.29	2.42	2.81	3.25	2.24	3.07	5.96***	32.68***
Rate of return (% Annual)	9.77	9.12	4.00	7.79	7.68	3.97	16.09***	21.73***

Table 2. Descriptive Statistics Continued

	Live Funds		Defunct Funds	
	Yes	**No**	**Yes**	**No**
Funds Dichotomous Characteristics				
Personal Capital (Do principals have money invested in the HF?)	27.78%	72.22%	31.43%	68.57%
Margin (Does the fund leverage using margin borrowing value?)	33.49%	66.51%	40.76%	59.24%
Derivatives (Does the fund leverage using derivatives value?)	18.87%	81.13%	19.65%	80.35%
Futures (Does the fund leverage using futures value?)	22.07%	77.93%	19.15%	80.85%
Leveraged (Does the fund use leverage as defined by Debt/Equity?)	54.13%	45.87%	56.90%	43.10%
HighWaterMark (Does the fund have High Water Mark structure?)	66.34%	33.66%	60.37%	39.63%
Off-shore (Is the fund registered off-shore?)	66.42%	33.58%	64.05%	35.95%

Note: The number of cross-sectional observations is 3362 firms in the Live Funds database and 6250 in the Defunct Funds database for the period 1996-2010.

Table 3

Descriptive statistics: Time Series

A total of 17,202 hedge funds reported monthly returns to TASS for the years between Jan-1996 and Dec-2011 in this database. From this number, 10,455 are defunct funds and 6,747 are live funds. For each year from 1996 to 2010, the table below reports the number of hedge funds that failed and the number of hedge funds that survived. It also reports the key characteristics of each group over time.

	Live Funds						Defunct Funds					
Year	**No. of Funds**	**No. of Obs.**	**Mean NAV ($Million)**	**Mean Annual Rate of Return (%)**	**Median Annual Rate of Return (%)**	**Std. Dev. Of Monthly Returns**	**Number of Funds Failed**	**Number of Obs.**	**Mean NAV ($Million)**	**Mean Annual Rate of Return (%)**	**Median Annual Rate of Return (%)**	**Std. Dev. Of Monthly Returns**
1996	284	3168	268.74	19.71	15.60	4.18	122	13740	132.48	16.38	14.64	4.41
1997	356	3979	273.50	18.49	14.45	4.42	94	16035	261.95	16.51	13.80	4.77
1998	437	4882	554.65	6.53	8.81	5.16	154	18424	445.96	6.36	8.65	5.47
1999	537	6055	683.73	24.62	15.12	4.85	175	20505	238.16	20.90	13.80	5.09
2000	660	7373	626.87	11.08	10.37	5.02	221	22518	230.72	7.99	9.12	5.45
2001	815	9108	613.71	9.18	8.00	4.02	242	24448	93.39	5.82	6.84	4.35
2002	988	11203	519.57	5.41	4.68	3.72	234	27324	229.97	2.34	3.96	3.65
2003	1212	13522	534.60	19.12	12.53	3.37	246	30484	137.63	14.85	10.13	3.10
2004	1485	16696	488.55	10.14	7.44	3.03	290	34350	124.99	7.28	6.00	2.78
2005	1795	20084	458.78	10.96	10.08	3.07	447	36728	150.71	7.96	7.79	2.83
2006	2091	23754	451.63	14.12	12.48	3.12	510	36873	176.00	10.94	10.32	2.86
2007	2427	27592	484.92	14.39	12.84	3.44	755	34747	411.92	9.62	9.84	3.23
2008	2750	31698	684.76	-15.14	-7.33	5.07	1209	27306	156.53	-15.90	-7.32	4.60
2009	3110	35764	827.02	18.69	12.60	4.14	799	15843	167.39	7.50	4.36	3.79
2010	3362	39470	913.65	9.46	7.99	3.59	752	8023	614.94	2.66	1.77	3.31

Note: The number of cross-sectional observations is 3156 firms in the Live Funds database and 7076 in the Defunct Funds database for the period 1996-2011

Table 4

Descriptive statistics: Investment Strategies

The table below shows the characteristics for each investment strategy category. AUM (Asset under management), Returns, Size, Leverage and Number of funds are provided as well as main statistics for each of them. Leverage is defined as debt-to-equity ratio.

		Annual Return			**Size (AUM), M$**			**Average Financial Leverage (D/E)**			**Number of Funds**		
		Mean	**Median**	**Std Dev.**	**Mean**	**Median**	**Std Dev.**	**Mean**	**Median**	**Std Dev.**	**Live Database**	**Defunct Database**	**Total**
	Funds Continuous Characteristics												
1	Convertible Arbitrage	8.20	9.12	2.76	166.38	119.22	399.81	144.66	100.0	187.45	47	175	222
2	Dedicated Short Bias	2.64	1.56	5.86	80.64	71.14	93.08	40.22	0.0	62.60	11	36	47
3	Emerging Markets	13.00	12.00	5.60	245.86	22.14	3581.21	19.13	0.0	50.76	386	391	777
4	Equity Market Neutral	7.26	6.58	2.89	689.40	103.96	10396.30	65.27	0.0	111.03	96	338	434
5	Event Driven	9.92	9.60	3.13	380.19	124.66	3457.26	41.51	0.0	86.68	185	481	666
6	Fixed Income Arbitrage	7.59	8.40	2.57	618.34	112.33	5149.51	302.43	3.5	770.18	52	204	256
7	Fund of Funds	5.62	7.26	2.79	114.51	19.36	408.49	21.32	0.0	92.02	1034	1592	2626
8	Global Macro	8.82	6.36	4.18	165.50	101.90	446.20	84.78	0.0	148.94	132	360	492
9	Long/Short Equity	10.58	9.36	4.75	711.40	94.96	15532.47	33.83	0.0	66.48	758	1728	2486
10	Managed Futures	8.83	6.14	5.39	430.09	110.67	3656.23	52.53	0.0	136.44	239	436	675
11	Multi-Strategy	8.60	8.76	3.48	302.88	92.36	1818.21	62.24	0.0	116.05	258	361	619
12	Options Strategy	8.34	7.92	3.15	109.33	115.47	75.65	58.24	0.0	86.57	16	8	24
13	Other/Undefined	10.60	9.36	3.16	1059.54	109.52	6644.73	96.05	0.0	155.41	148	140	288
	TOTAL	8.60	8.28	3.98	412.44	84.44	9498.74	49.44	0.0	173.22	3156	7076	10455

Table 5: Correlation Matrix

The table presents the pair-wise correlation coefficients. The sample includes 1996-2011 data from Lipper TASS database. Non-USD firms are excluded. Variable definitions are provided in Appendix A. Note that performance (return) and risk (volatility) are winsorized at 1% on each tail. * indicates statistical significance at the 95% level or higher.

	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	Leverage	1									
(2)	Margin	**0.47***	1								
(3)	Derivatives & Futures	**0.46***	0.28*	1							
(4)	$Size_{t-1}$	0.04*	0.05*	0.07*	1						
(5)	$Performance_{t-1}$	0.01*	0.01*	0	0.04*	1					
(6)	$Volatility_{t-1}$	0.1*	0.04*	0.1*	0	0	1				
(7)	Incentive fee	0.25*	0.13*	0.16*	0.15*	0.04*	0.19*	1			
(8)	Management fee	0.08*	-0.05*	0.13*	-0.04*	0	0.07*	0.09*	1		
(9)	HighWaterMark	0.06*	0	0	0.07*	0.01*	-0.02*	0.37*	0	1	
(10)	Offshore	-0.02*	0	0	-0.33*	-0.03*	-0.09*	-0.15*	0.09*	-0.09*	1

Table 6: Downside Correlation

Correlation Analysis for monthly returns of HFs with S&P 500 (Aggregate Market)
The table below shows the correlation between monthly returns on each Investment Strategy category with the aggregate market (here, S&P500).

The purpose is to identify the increased correlation during crisis, reflective of "downside correlation". [1]

	PrimaryCategory (Investment Strategy)	Returns Correlation with S&P500			
		All time	Post Crisis 2007-2010	Pre Crisis 1996-2006	Post-Pre Crisis
1	Convertible Arbitrage	0.26	0.41	0.16	0.25
2	Dedicated Short Bias	-0.50	-0.49	-0.51	0.02
3	Emerging Markets	0.38	0.45	0.32	0.13
4	Equity Market Neutral	-0.02	-0.03	0.03	-0.06
5	Event Driven	0.31	0.41	0.24	0.18
6	Fixed Income Arbitrage	0.11	0.24	0.01	0.23
7	Fund of Funds	0.27	0.28	0.23	0.05
8	Global Macro	0.11	0.11	0.11	-0.001
9	Long/Short Equity Hedge	0.09	0.06	0.30	-0.24
10	Managed Futures	-0.02	0.01	-0.04	0.05
11	Multi-Strategy	0.24	0.25	0.21	0.04
12	Options Strategy	0.06	0.03	0.17	-0.15

*All correlation values were found significant at 99% or 95% confidence level, unless indicated otherwise.

[1] See Ang and Chen (2002) for more details.

Table 7: Regression of probability of failure on fund's financial leverage for all hedge funds

$$\Pr ob(Failure_{i,t}) = \beta_0 + \beta_1 FinLeverage_{i,t-1} + \gamma' \mathbf{X} + \varepsilon_{i,t}$$

The dependent variable is the dichotomous variable "Failure" that takes values 1 and 0. The value is observed monthly, where *Failure =1* for the month that the hedge fund has failed. Actual Financial leverage (FinLev) is the average actual leverage ratio defined as total debt over total equity for each fund. **X** is a vector of control variables including size, prior performance, Incentive Fee, Offshore, and High water mark. Z-statistics are reported below as well as the estimated coefficients. Probit models (2) and (3) each take multiplicative factors into account. For complete detailed definitions of variables, refer to the Appendix A.
Significance levels are denoted as (*) if p < 0.10, (**) if p < 0.05, and (***) if p < 0.01.

Dependent Variable: Dichotomous for Failure	Probit (1)		Probit (2)		Probit (3)	
	Failure: All-Time (1996-2011)					
	Estimates	***Z-stat***	**Estimates**	***Z-stat***	**Estimates**	***Z-stat***
Financial Leverage	0.02***	*2.7*	0.02**	*2.2*		
(FinLev)*(Deriv/Future)					0.01	*0.5*
(FinLev)*(Margin)					0.05***	*5.1*
(Deriv/Futures)*(Margin)			0.01	*1.1*		
Size	-0.01**	*-2.1*	-0.01**	*-2.2*	-0.01***	*-3.5*
Performance ($Return_{t-1}$)	-0.03***	*-20.6*	-0.03***	*-20.6*	-0.02***	*-17.3*
Incentive Fee	0.01***	*8.5*	0.01***	*8.5*	0.01***	*8.9*
Highwatermark	-0.04***	*-4.6*	-0.04***	*-4.5*	-0.09***	*-8.3*
Offshore	0.02	*1.6*	0.02	*1.6*	0.01	*0.9*
Management Fee	-0.01	*-0.2*	0.00	*-0.2*	0.01	*1.2*
Crisis	0.09***	*9.9*	0.09***	*9.9*	0.09***	*8.5*
Intercept	-2.34***	*-45.6*	-2.34***	*-45.5*	-2.27***	*-39.5*
Year Effects & Cluster Effects	Yes		Yes		Yes	
Log pseudolikelihood	-35188		-35188		-19562	
Pseudo R^2	0.01		0.01		0.01	
Observations	637980		637980		258504	

Table 8: Regression of probability of failure on fund's financial leverage for all hedge funds: Pre-Crisis vs. Post-Crisis

$$\Pr ob(Failure_{i,t}) = \beta_0 + \beta_1 FinLeverage_{i,t-1} + \gamma' \mathbf{X} + \varepsilon_{i,t}$$

The dependent variable is the dichotomous variable "Failure" that takes values 1 and 0. The value is observed monthly, where *Failure =1* for the month that the hedge fund has failed. Actual Financial leverage (FinLev) is the average actual leverage ratio defined as total debt over total equity for each fund. **X** is a vector of control variables including size, prior performance, Incentive Fee, Offshore, and High water mark. Z-statistics are reported below as well as the estimated coefficients. Probit models (2) and (3) each take multiplicative factors into account. For more complete definitions of variables, refer to the Appendix A.

Significance levels are denoted as (*) if $p < 0.10$, (**) if $p < 0.05$, and (***) if $p < 0.01$.

Dependent Variable: Dichotomous for Failure	Probit (1)		Probit (2)		Probit (3)		Probit (1)		Probit (2)		Probit (3)	
	Failure: Non/Pre-Crisis						**Failure: During/Post-Crisis**					
	Estimates	***Z-stat***	**Estimates**	***Z-stat***	**Estimates**	***Z-stat***	**Estimates**	***Z-stat***	**Estimates**	***Z-stat***	**Estimates**	***Z-stat***
Financial Leverage	0.02	*1.3*	0.01	*0.9*			0.03**	*2.5*	0.03**	*2.1*		
(FinLev)*(Deriv/Future)					0.01	*0.5*					0.00	*0.3*
(FinLev)*(Margin)					0.04***	*2.9*					0.07***	*4.2*
(Deriv/Futures)*(Margin)			0.02	*0.9*					0.01	*0.7*		
Size	-0.01**	*-2.1*	-0.01**	*-2.1*	-0.01***	*-3.4*	-0.01	*-0.8*	0.00	*-0.9*	-0.01	*-1.3*
Performance (Return $_{t-1}$)	-0.02***	*-13.8*	-0.02***	*-13.8*	-0.02***	*-11.9*	-0.03***	*-15.6*	-0.03***	*-15.6*	-0.02***	*-12.9*
Incentive Fee	0.01***	*7.2*	0.01***	*7.2*	0.01***	*7.1*	0.01***	*4.6*	0.00***	*4.6*	0.01***	*5.4*
Highwatermark	-0.02	*-1.1*	-0.02	*-1.1*	-0.05***	*-3.6*	-0.08***	*-5.2*	-0.08***	*-5.2*	-0.13***	*-7.7*
Offshore	-0.01	*-0.6*	-0.01	*-0.7*	-0.02	*-1.1*	0.04***	*2.9*	0.04***	*2.9*	0.04**	*2.4*
Management Fee	0.00	*0.2*	0.00	*0.2*	0.02*	*1.9*	-0.01	*-0.4*	0.00	*-0.4*	0.00	*-0.2*
Crisis												
Intercept	-2.33***	*-31.9*	-2.32***	*-31.8*	-2.23***	*-27.7*	-2.27***	*-29.0*	-2.27***	*-29.0*	-2.23***	*-25.5*
Year Effects & Cluster Effects	Yes		Yes		Yes		Yes		Yes		Yes	
Log pseudolikelihood	-18587		-18587		-14443		-16589		-16589		-12650	
Pseudo R^2	0.01		0.01		0.08		0.01		0.01		0.01	
Observations	372542		372542		285800		265438		265438		202359	

Note: The R-squared for the Probit regression is the square of the correlation between the model's predicted values and the actual values

Table 9: Regression of probability of failure on fund's financial leverage for Significant Styles

$$\Pr ob(Failure_{i,t} = 1) = \beta_0 + \beta_1 FinLeverage_{i,t-1} + \gamma' \mathbf{X} + \varepsilon_{i,t}$$

The dependent variable is the dichotomous variable "Failure" that takes values 1 and 0. The value is observed monthly, where *Failure =1* for the month that the hedge fund has failed. Actual Financial leverage (FinLev) is the average actual leverage ratio defined as total debt over total equity for each fund. **X** is a vector of control variables including size, prior performance, Incentive Fee, Management Fee, Offshore, and High water mark. Z-statistics are reported below as well as the estimated coefficients. For more detailed definitions of variables, refer to the Appendix A. Significance levels are denoted as (*) if $p < 0.10$, (**) if $p < 0.05$, and (***) if $p < 0.01$.

The findings of the table below show that for the style groups **"Emerging Markets", "Equity Market", "Long/Short Equity", and "Multi-Strategy"**, the leverage is a significant factor for the failure of the fund during the entire time period. This can be due to the fact that these styles have larger values of market return component (i.e. *beta*) which is explained further in the paper.

Dependent Variable: Dichotomous for Failure	Emerging Markets 3		Equity Market Neutral 4		Long/Short Equity Hedge 9		Multi-Strategy 11	
	Estimates	***Z-Stat***	**Estimates**	***Z-Stat***	**Estimates**	***Z-Stat***	**Estimates**	***Z-Stat***
Financial Leverage	0.08**	*2.52*	0.09***	*2.79*	0.04***	*2.75*	0.07**	*2.03*
Size	*0.00*	*0.37*	0.01	*0.88*	0.00	*0.13*	-0.02**	*-2.12*
Performance (Return $_{t-1}$)	-0.02***	*-4.46*	-0.05***	*-6.51*	-0.02***	*-10.55*	-0.04***	*-7.80*
Incentive Fee	0.00	*0.48*	0.00	*0.75*	0.00	*0.79*	0.00	*-0.72*
Highwatermark	-0.06*	*-1.71*	0.01	*0.17*	0.00	*-0.25*	0.01	*0.25*
Offshore	-0.01	*-0.21*	0.06*	*1.65*	0.03**	*2.04*	-0.02	*-0.46*
Management Fee	0.11***	*2.94*	-0.01	*-0.27*	0.03*	*1.76*	-0.01	*-0.52*
Crisis	0.07*	*1.70*	0.02	*0.50*	0.04*	*1.93*	0.15***	*3.85*
Intercept	-2.39***	*-14.12*	-2.37***	*-12.36*	-2.25***	*-25.44*	-1.99***	*-9.73*
Year Effects & Cluster Effects	Yes		Yes		Yes		Yes	
Log pseudolikelihood	-2063		-1674		-9778		-2105	
Pseudo R^2	0.01		0.01		0.01		0.02	
Observations	22945		18184		169988		37935	

Note: The R-squared for the Probit regression is the square of the correlation between the model's predicted values and the actual values.

Table 10a: Regression of probability of failure on fund's financial leverage for Non-Significant Styles

$$\Pr ob(Failure_{i,t} = 1) = \beta_0 + \beta_1 FinLeverage_{i,t-1} + \gamma' \mathbf{X} + \varepsilon_{i,t}$$

The dependent variable is the dichotomous variable "Failure" that takes values 1 and 0. The value is observed monthly, where *Failure =1* for the month that the hedge fund has failed. Actual Financial leverage (FinLev) is the average actual leverage ratio defined as total debt over total equity for each fund. **X** is a vector of control variables including size, prior performance, Incentive Fee, Management Fee, Offshore, and High water mark. Z-statistics are reported below as well as the estimated coefficients. For more detailed definitions of variables, refer to the Appendix A. Significance levels are denoted as (*) if $p < 0.10$, (**) if $p < 0.05$, and (***) if $p < 0.01$.

The findings of table below show that for the style groups **"Convertible Arbitrage", "Dedicated Short Bias", "Event Driven",** and **"Fixed Income Arbitrage"**, the leverage is not a significant factor for the failure of the fund throughout the entire time period.

Dependent Variable: Dichotomous for Failure	Convertible Arbitrage 1		Dedicated Short Bias 2		Event Driven 5		Fixed Income Arbitrage 6	
	Estimates	***Z-Stat***	**Estimates**	***Z-Stat***	**Estimates**	***Z-Stat***	**Estimates**	***Z-Stat***
Financial Leverage	-0.06	*-1.14*	-0.12	*-1.54*	0.02	*0.94*	-0.02	*-0.4*
Size	-0.02	*-1.34*	-0.06	*-1.54*	-0.01	*-1.67*	0.01	*0.58*
Performance (Return $_{t-1}$)	-0.07***	*-5.81*	0.02**	*2.02*	-0.04***	*-6.46*	-0.06***	*-5.94*
Incentive Fee	0.00	*-0.83*	0.01	*1.6*	0.00	*-0.79*	0.00	*-0.15*
Highwatermark	0.07	*1.53*	-0.03	*-0.35*	0.05*	*1.78*	-0.07	*-1.33*
Offshore	0.01	*0.15*	0.10	*1.3*	0.02	*0.86*	-0.02	*-0.37*
Management Fee	0.08*	*1.94*	0.07	*0.99*	0.05*	*1.79*	0.04	*0.99*
Crisis	-0.01	*-0.1*	0.07	*0.43*	0.15***	*3.96*	0.32***	*5.53*
Intercept	-1.77***	*-5.05*	-1.41*	*-1.86*	-2.04***	*-12.65*	-2.40***	*-8.28*
Year Effects & Cluster Effects	Yes		Yes		Yes		Yes	
Log pseudolikelihood	-869		-181		-2401		-896	
Pseudo R^2	0.03		0.02		0.02		0.04	
Observations	11623		2590		30142		11549	

Note: The R-squared for the Probit regression is the square of the correlation between the model's predicted values and the actual values.

Table 10b: Regression of probability of failure on fund's financial leverage for Non-Significant Styles(Cont.)

$$\Pr ob(Failure_{i,t} = 1) = \beta_0 + \beta_1 FinLeverage_{i,t-1} + \gamma' \mathbf{X} + \varepsilon_{i,t}$$

The dependent variable is the dichotomous variable "Failure" that takes values 1 and 0. The value is observed monthly, where *Failure =1* for the month that the hedge fund has failed. Actual Financial leverage (FinLev) is the average actual leverage ratio defined as total debt over total equity for each fund. **X** is a vector of control variables including size, prior performance, Incentive Fee, Management Fee, Offshore, and High water mark. Z-statistics are reported below as well as the estimated coefficients. For more detailed definitions of variables, refer to the Appendix A.

The findings of table below show that for the style groups **"Fund of Funds", "Global Macro", "Managed Futures"**, and **"Option Strategy"**, the leverage is not a significant factor for the failure of the fund when observed throughout the entire time period.

Dependent Variable: Dichotomous for Failure	Funds of Funds 7		Global Macro 8		Managed Futures 10		Option Strategy 12	
	Estimates	***Z-Stat***	**Estimates**	***Z-Stat***	**Estimates**	***Z-Stat***	**Estimates**	***Z-Stat***
Financial Leverage	0.01	*0.85*	-0.06	*-1.31*	-0.01	*-0.27*	-0.25	*-1.13*
Size	-0.01	*-1.54*	-0.02	*-1.51*	-0.02***	*-2.6*	0.03	*0.29*
Performance (Return $_{t-1}$)	-0.05***	*-13.13*	-0.01**	*-2.49*	-0.01***	*-3.02*	0.12***	*3.79*
Incentive Fee	0.00***	*4.33*	0.00	*0.9*	0.00	*0.3*	0.02	*0.32*
Highwatermark	0.02	*1.58*	0.02	*0.54*	-0.03	*-1.07*	-1.80	*-1.46*
Offshore	0.01	*0.77*	0.02	*0.52*	0.05	*1.64*	0.24*	*1.78*
Management Fee	-0.05***	*-4.26*	0.00	*0.1*	-0.04***	*-2.99*	0.00	*-0.01*
Crisis	0.13***	*6.47*	0.11**	*2.15*	0.04	*1.01*	-1.00**	*-2.44*
Intercept	-2.10***	*-26.52*	-1.78***	*-7.53*	-1.64***	*-9.97*	-1.29	*-0.56*
Year Effects & Cluster Effects	Yes		Yes		Yes		Yes	
Log pseudolikelihood	-8365		-1426		-2147		-68	
Pseudo R^2	0.02		0.01		0.01		0.18	
Observations	105975		13834		23343		1041	

Note: The R-squared for the Probit regression is the square of the correlation between the model's predicted values and the actual values.

Chapter 3

DOES DEFAULT RISK IMPACT EQUITY OPTIONS?

Abstract

What is the impact of default risk on equity option pricing? We study this question in detail by empirically examining to what extent the firm-specific default risk matters in pricing individual equity options. Since credit default swaps (CDS) are similar to put options in that both offer a low cost and effective protection against downside risk, we use CDS spread as credit risk proxy to investigate the effects of default risk on put pricing. Recent financial crisis showed that for many financial firms equity options experienced high implied volatility (IV) when the underlying CDS spreads went up. By examining an exhaustive sample of US-listed firms with both CDS and put options data available over the period from 2002 to 2010, and studying the primary determinants of option IVs cross-sectionally and over time, the findings show that default risk is a significant factor in the prices of equity options. Moreover, the impact of default risk remains significant after controlling for firm-specific and macroeconomic factors. This study relates to recent literature that explains how default risk can get injected from the fixed income market to the equity options market and why default risk is important in the pricing of equity options and implied volatility.

Keywords: Option Pricing, Default Risk, Implied Volatility, CDS Spreads, Volatility Skew

3.1 Introduction

Does default risk matter in equity option pricing? If underlying firm is subject to higher default risk does that matter for option markets? In this essay we address this question in details, by empirically examining to what extent the firm-specific attributes and systematic variables matter in pricing individual equity options. The existence of an association between default risk and implied volatility, which is inverted from option prices, is a widely accepted notion. We study the primary determinants of option IVs cross-sectionally and over time, and measure the contribution of default risk on equity option pricing. Since credit default swaps (CDS) are similar to out-of-the-money put options in that both offer a low cost[1] and effective protection against downside risk, we use CDS spread as credit risk proxy to investigate the effects of credit risk on put option pricing.

In addition, we analyze whether the equity option skew is explained by the underlying firm's default risk or not. If there exists a default risk premium in stocks, the expected return on the stocks should include a compensation for default risk. This has been empirically tested in previous literature.[2] In other words, higher default risk is associated with closer distance to default of the underlying firm, and hence higher risk or implied volatility for that stock. In this paper we study if firm's default risk matters for option pricing: firms subject to higher default risk have greater firm-specific volatility increasing option value. Similarly, increase in default risk over time for a given firm can increase the delta-neutral cost of hedging and hence higher option prices.

[1] The cost can be relative.

[2] For example see Vassalou and Xing, (2004); Campbell, Hilscher, and Szilagyi (2008); Garlappi and Yan, (2011)

The overall objective of this study is to examine the association between default probability and implied volatility. This relationship is well known as the 'leverage effect'.[3] When stock prices go down the volatility goes up which increases the leverage component, and hence the firm becomes more defaultable. We investigate this association and study the interactions between CDS and option markets. Incorporating firm-specific default risk may be an alternative way of capturing additional risks. High default risk can enhance the implied volatility smiles/smirks, and hence potentially provide an alternative explanation as to why the underlying risk neutral distributions are skewed. [4]If default risk significantly explains the implicit variance risk premium, default risk premium may also be priced in equity options.

Further high aggregate default risk can also imply high systematic risk that can have a pronounced impact on option prices and IV skew.[5] If default risk matters and is not controlled for, existing pricing models could lead to pricing biases especially for OTM puts or ITM calls.[6] Profits from several buy-hold and arbitrage strategies involving individual options, straddles, put-call parity arbitrage trades can be spurious, as they could be driven by underlying credit risk. Our study intends to shed more light on these issues by highlighting the association between credit risk and options values.

Why does default risk matter for option prices? Extant literature shows that Implied volatility smiles and skews along the moneyness direction are direct results

---

[3]The term "Leverage effect" refers to the well-established relationship between stock returns and both implied and realized volatility: volatility increases when the stock price falls

[4]The variance risk premium can be one indication of this.

[5]Negative skewness is partly related to market beta (Dennis & Mayhew, 2002) and systematic risk is significant in the price structure of equity options (Duan & Wei, 2009).

[6]Previous theoretical framework has examined the pricing of vulnerable options (Johnson & Stulz, 1987), pricing options on defaultable stocks (Bayraktar 2008; Bayraktar and Yang 2011), and develop a robust link between deep OTM American puts on a company's stock and a credit insurance contract on the company's bond (Carr & Wu, 2011).

of conditional non-normality in the underlying stock returns under the risk-neutral measure. The negative slope of the implied volatility skew indicates negative skewness in the risk-neutral return distribution and presence of variance risk premium. As Carr and Wu (2009) and Bayraktar and Yang (2011) show, a possible source of such variance risk premium could be the underlying default risk of the firm.

Carr and Wu (2009) study the interaction between market risk (return variance) and credit risk (default arrival) in pricing stock options and credit default swaps. They model default intensity is a positive function of stock return variance (details below) . Higher variance leads to higher default arrival rate impacting option price; the default arrival rate itself contributes positively to the option implied volatility. The negative risk-neutral return skewness in their model therefore arises from three sources: (i) positive probability of default (ii) asymmetry in the high-frequency jump component and (iii) negative correlation between the return and variance processes. Their results show that when a company's credit spread widens, its implied variance skew becomes more negatively skewed. Shocks to the more persistent credit risk factor last longer across the term structure of options and credit spreads.

Further studies have examined the significance of firm-specific default risk in pricing equities (e.g. Vassalou and Xing, 2004; Campbell, Hilscher, and Szilagyi, 2008; Garlappi and Yan, 2011). However high default risk can also influence the risk of the underlying equity options. Variations in default risk can imply yet another source of diversifiable risk that needs to be dynamically hedged using a risk free bond, stock and underlying corporate bond in creating a delta-neutral portfolio. As a result, delta-neutral portfolios with excess default risk imply higher put option prices (or equivalently higher IVs), ceteris paribus. Existing option pricing literature shows that no-arbitrage portfolio consisting of stock and risk-free T-bills alone cannot successfully span option returns, and hence additional risk factors such as stochastic volatility and jumps are needed. [7]

High default risk can enhance the IV smiles/smirks, and hence provide an alternative explanation as to why the underlying risk neutral distributions are skewed. Further high aggregate default risk can also imply high systematic risk that can have a pronounced impact on option prices and IV skew. If default risk matters and is not controlled for, existing pricing models could lead to pricing biases especially for OTM puts or ITM calls. Profits from several buy-hold and arbitrage strategies involving individual options, straddles, put-call parity arbitrage trades can be spurious, as they could be driven by underlying credit risk. Our study intends to shed more light on these issues.

To the best of our knowledge, this is the first empirical study on impact of default risk on option markets.[8] Cao,Yu and Zhong (2010) investigate whether put option implied volatility is an important determinant of CDS spreads. Using firm level time series models of IV they find that individual firms' put option-implied volatility dominate historical volatility in explaining the time-series variation in CDS spreads. Their data is the closest to our empirical approach of this research. To better control for effects of option IV on CDS we employ instrumental variable approach for possible endogenity. If there is a default risk premium in stocks, the expected return includes a compensation for default risk. Higher default risk is associated with closer distance to default of the underlying firm relative to other peer firms, and hence higher idiosyncratic risk or implied volatility for that stock, thereby causing option prices to go up. Such a stock could also exhibit excess

[7] For example, it has been shown that deterministic volatility models are insufficient in spanning of the pricing kernel (Burasachi & Jackwerth, 2001).

[8] We use CDS data to compute default measures for individual firms and measure the effect of default risk on a comprehensive group of equity option prices.

skew in the risk neutral distribution compared to stock issued by a similar firm, but with lower default risk. So the additional question investigated in this paper is: Is the equity option skew associated with the underlying default risk of the reference firm?

Previously, there have been extensive theoretical work on modeling default risk for the purpose of valuing corporate debt and derivative products written on it, but little attention has been paid to the effects of default risk on equity options. See for example Bayraktar (2008), Bayraktar and Yang (2011), and Carr and Wu (2011).[9] Can we measure this impact empirically by identifying default risk as one of the determinants of the option-implied volatility after controlling for firm and issuer characteristics?

We measure default risk using CDS spreads on the underlying reference firm. We employ an exhaustive panel dataset of (a) individual equity options (Option Metrics) and (b) CDS (Markit) for the period 2002-2010, and undertake a comprehensive study by addressing these key issues: (1) Does credit risk matter for option pricing? How do credit markets influence individual option pricing? (2) Does firm specific default risk still matter for option pricing, even after controlling for option liquidity and aggregate default risk? (3) Does credit risk impact option skewness?

Since CDS, like put option, is mainly used for downside risk protection, we test our hypotheses on an extensive sample of put options. We find that after controlling for market variables and firm variables, the default risk is still a significant

[9] Bayraktar (2008) develops stock option price approximations for a model which takes both the risk of default and the stochastic volatility into account. The model for the first time also lets the volatility impact intensity of defaults and hence shows that an effective hazard rate from bonds issued by a company can be used to explain the implied volatility skew of the option prices issued by the same company.

factor in cross section of option implied volatility. This significance is prominent in far in-the-money and far out-of-the-money put options.[10]

3.2 Literature Background

This paper intersects several streams of literature on credit risk and option market. The following literature categories are summarized below: "default Risk and option markets", "Default risk in equity markets", and finally the latest studies on "Cross-sectional Pricing in Equity Options Markets".

3.2.1 Default Risk and Option Markets

The relationship between default risk and option prices have primarily been subject of a number of theoretical papers.

Probably one of the first studies is by Johnson and Stulz (1987) who study the pricing of options with third party risk, so called "vulnerable options". Many options and financial assets containing option-like payoffs (such as insurance contracts) are sold by firms that have limited assets. Such options are subject to default risk of the parties who sell the options.[11] Johnson and Stulz show that many of the well-established results of the option pricing literature do not hold for vulnerable options.

Hanke (2005) develop a new model for the (closed-form) pricing of options on leveraged equity, which allows us to study capital structure effects on equity

[10]The findings are also consistent with the "Leverage Effect" notion explained earlier.

[11]Because CDS trades OTC, there is counter party risk associated with that security so the default risk will be confounded with counter party risk.

option prices. Using the Leland and Toft (1996) model as an example, they show how classical capital structure models can be extended to allow for closed-form pricing of options on equity within these models. Their model features finite debt maturity, exponentially increasing coupon debt, taxes, intermediate bankruptcy, and deviations from the absolute priority rule.

Bayraktar (2008) develops stock option price models which take both the risk of default and the stochastic volatility into account. By building default risk into the pricing model, he shows that an effective hazard rate from bonds issued by a company can be used to explain the implied volatility skew of the option prices issued by the same company. This is another important finding that motivates our empirical question of testing whether default risk priced in equity options.

Chen and Kou (2009) propose a two-sided jump model for credit risk by extending the Leland–Toft endogenous default model based on the geometric Brownian motion. The model shows that jump risk and endogenous default can have significant impacts on credit spreads, optimal capital structure, and implied volatility of equity options:

For example, following similar approach of incorporating both default risk and option IV jumps, Carr and Wu (2009) propose a dynamically consistent framework that allows joint valuation and estimation of stock options and credit default swaps written on the same reference company. Carr and Wu assume that the stock price is strictly positive prior to default and falls to zero upon default. Prior to default, the stock price follows a jump-diffusion process and stochastic volatility. Further, stock return process has a drift determined by default rate, which is controlled by a Cox process with a stochastic arrival rate, i.e. the stock price converges to zero at time of default. The instantaneous default rate and variance rate follow a bivariate continuous process, with its joint dynamics specified to capture the observed behavior of stock option prices and credit default swap spreads. Their model estimation result (performed on eight companies across five sectors) shows proof for the interaction between market risk (return variance) and credit risk (default arrival) in pricing stock options and credit default swaps.

Under Carr and Wu's model specification, negative risk-neutral return skewness arises from three sources: (i) positive probability of default (ii) asymmetry in the high-frequency jump component and (iii) negative correlation between the return and variance processes. In their model, default intensity is a positive function of stock return variance. Higher variance leads to higher default arrival rate impacting option price; the default arrival rate itself contributes positively to the option implied volatility.

Their results show that when a company's credit spread widens, its implied variance skew becomes more negatively skewed. Shocks to the more persistent credit risk factor last longer across the term structure of options and credit spreads. Their estimation highlights the interaction between market risk (return variance) and credit risk (default arrival) in pricing stock options and credit default swaps.

Bayraktar and Yang (2011) demonstrate the importance of accounting for the default risk and stochastic interest rate in equity option pricing. Their pricing framework combines four building blocks i.e. a) Vasicek interest rate model, b) fast-mean reverting stochastic volatility model, c) defaultable stock price model, and d) multi-scale stochastic intensity model, which can be jointly calibrated to the corporate bond term structure and equity option volatility surface of the same company. Their purpose is to obtain explicit bond and equity option pricing formulas that can be calibrated to find a risk neutral model that matches a set of observed market prices. This risk neutral model can be in turn used to price more exotic, illiquid or over-the-counter derivatives. Bayraktar and Yang (2011) extends

Bayraktar (2008) model explained earlier, by taking the interest rate process to be stochastic.

Strong theoretical literature covered in this section supports the investigation of default risk impact on equity options pricing.

3.2.2 Default Risk in Equity Markets

The significance of firm-specific default risk in pricing equities has been documented in current literature. There are two major dimensions to this: First, high default risk is associated with closer distance-to-default of the underlying firm, and hence possible higher risk premium for that stock. The reason is that if there exists a default risk premium in stocks, the expected return includes an additional compensation for default risk. Empirical evidence on this issue has however been inconclusive. (e.g. Vassalou & Xing, 2004; Campbell, Hilscher, and Szilagyi, 2008; Garlappi & Yan, 2011). Second, high default risk can also influence the risk of the underlying equity options. Variations in default risk can imply yet another source of diversifiable risk that needs to be dynamically hedged in creating a delta-neutral portfolio. As a result, delta-neutral portfolios with excess default risk imply higher put option prices (or equivalently higher implied volatilities-IVs), ceteris paribus.

Burasachi & Jackwerth, 2001, have shown that deterministic volatility models are insufficient in spanning of the pricing kernel. Existing option pricing literature also shows that no-arbitrage portfolio consisting of stock and risk-free T-bills alone cannot successfully span option returns, and hence additional risk factors such as stochastic volatility and jumps are needed.

According to Vassalou and Xing (2004), default risk is systematic risk and therefore priced in cross section of returns. In their study they use Merton's

(1974) option pricing model for the first time to compute default measures for individual firms and assess the effect of default risk on equity returns. They find a risk-based interpretation for size and book to market (B/M) effects. The size effect is a default effect, and the same analogy is mostly true for the B/M effect. Both size and B/M effects exist only when there is high default risk. In other words, small firms have higher returns than large firms only if they have high default risk. Vassalou and Xing (2004) further show that the Fama French (FF) factors SMB and HML contain some default-related information, but this is not the main reason that the FF model can explain the cross section of equity returns. Based on FF (1996) argument, SMB and HML proxy for default risk. Vassalou (2003) and Li, Vassalou and Xing (2000) show that the risk based explanations for these factors and hence default by itself is a variable beyond size and B/M effects. This variable can be the risk premium for default risk. For example, distressed portfolios have low average returns, but high standard deviations, market betas, and loadings on FFs (1993) small-cap and value risk factors. These portfolios also tend to do poorly when market wide implied volatility increases. In other words, from the perspective of any of the leading empirical asset pricing models, these stocks have negative alphas. Based on this result, the conjecture that the value and size effects are proxies for a financial distress premium can be challenged. Moreover, in standard models of rational asset pricing in which the structure of the economy is stable and well understood by investors, this result is restricting.[12]

To analyze distress risk, Campbell, Hilscher and Szilagyi (2008) explore the determinants of corporate failure and the pricing of financially distressed stocks with high failure probability. Analyzing data from 1981, they show that financially distressed stocks have delivered anomalously low returns, but much higher standard

[12]Campbell, Hilscher, Szilagyi (2008)

deviations, market betas, and loadings on value and small-cap risk factors than stocks with low failure risk. Their patterns are stronger for stocks with possible informational or arbitrage-related frictions. Their findings hence are inconsistent with the FF conjecture that the value and size effects are compensation for the risk of financial distress. From methodology point of view, Campbell, Hilscher and Szilagyi (2008) estimate the failure probability from a dynamic logit model using accounting and market variables. This in essence is the same hazard model used by Shumway (2001) and Chava and Jarrow (2004). Distance to default measure of KMV, is used by Vassalou and Xing (2004) and Da and Gao (2010), however, Campbell, Hilscher and Szilagyi (2008) claim that their reduced form model more accurately measures the risk of failure at short and long horizons and hence can more accurately measure the premium that investors receive for holding distressed stocks.

Campbell, Hilscher and Szilagyi (2008) implement a reduced-form econometric model to predict corporate bankruptcies and failures at short and long horizons. Their model has greater explanatory power than the previous models estimated by Shumway (2001) and Chava and Jarrow (2004), and includes more variables with economic interpretation. They also show that failure risk cannot be adequately summarized by earlier measures of distance to default inspired by Merton's (1974) pioneering structural model. While not exactly the same measure as Crosbie and Bohn (2001) and Vassalou and Xing (2004), Campbell, Hilscher and Szilagyi (2008) claim their measure, similar to Bharath and Shumway (2008), is robust to alternative measures of distance to default.

In order to explain the risk-neutral skewness implied from option prices, Dennis and Mayhew (2002) empirically establish a link between the risk neutral skewness and the systematic risk of the underlying stock. They explain the structural difference in distributions by investigating the relative importance of several firm characteristics such as implied volatility, firm size, trading volume, leverage, and beta. They show that risk-neutral skewness tends to be more negative for stocks with larger betas. This is an evidence for the importance of market risk in option pricing. We ensure inclusion of systematic risk in our model by empirically controlling for the market risk and macroeconomic risk.

So far, the findings on the relationship of default risk to equity returns are not in consensus with the rational expectation of higher risk, higher return. A series of later literature try to tackle this anomaly. In the working paper by Anginer and Yldzhan (2010), they approach the same anomaly: while financial theory suggests a positive relationship between default risk and equity returns, the empirical papers find anomalously low returns for stocks with high probabilities of default. How can this be explained? Anginer and Yildizhan (2010) show that returns to distressed stocks in previous work, are in fact a combination of anomalies associated with three stock characteristics: leverage, volatility and profitability. In their work they use a market based measure, corporate credit spreads, to proxy for default risk. Their distinctive approach is that as opposed to using measures that proxy for a firm's real-world probability of default, they use credit spreads as proxy for a risk-adjusted (or a risk-neutral) probability of default and hence explicitly account for the systematic component of distress risk. Their results show that credit spreads predict corporate defaults better than previously used measures, such as, bond ratings, accounting variables and structural model parameters.

Finally, we need to incorporate the skew in asset returns in our study. Harvey and Siddique (2000) explore the role of systematic skewness in asset returns, by asking whether expected returns include premium for this risk. They form an asset pricing model that incorporates conditional skewness. Their results show that conditional skewness helps explain the cross-sectional variation of expected returns across assets and is significant even when factors based on size and book-to-market are included. Their finding proves that systematic skewness is economically important and requires a risk premium, on average, of 3.60 percent per year. Based on their results, the momentum effect is related to systematic skewness. In other words, the low expected return momentum portfolios have higher skewness than high expected return portfolios. In line with findings of Campbell, Hilscher and Szilagyi (2008) and unlike to Vassalou and Xing (2004), Anginer and Yldzhan (2010) do not find default risk to be significantly priced in the cross-section of equity returns. Yet, they also find no evidence of firms with high default risk delivering anomalously low returns.

3.2.3 Cross-sectional Pricing in Equity Options Markets

In light of the systematic risk and equity options, Duan and Wei (2009) show the impact of systematic risk on the prices of individual equity options. They show that option prices are characterized by the level and slope of implied volatility curves, and the systematic risk is measured as the proportion of systematic variance in the total variance. Duan and Wei (2009) use daily option quotes on the S&P 100 index and its 30 largest component stocks, and control for the underlying asset's total risk. They find that a higher amount of systematic risk leads to a higher level of implied volatility and a steeper slope of the implied volatility curve. Thus, systematic risk proportion can help differentiate the price structure across individual equity options. In our paper the study goes beyond the options listed in the index, and is performed on a broad cross section of US firms all US-listed firms with both options and CDS written on them.

Utilizing the theory of option market liquidity, Cao and Wei (2010) examine option market liquidity using Ivy DB's Option Metrics data. They empirically test for evidence of commonality for various liquidity measures based on the bid ask spread, volumes, and price impact. They show that commonality remains strong even after controlling for the underlying stock markets liquidity and other liquidity determinants such as volatility. In other words, smaller firms and firms with a higher volatility exhibit stronger commonalities in option liquidity. They also show some properties of the option market's liquidity such as role of information asymmetry vs. inventory risk, the linkage between market-wide option liquidity and the underlying stock market's movements and the options liquidity asymmetric response to upward and downward market movements (calls reacting more in up markets and puts reacting more in down markets).

The findings of Cao and Wei (2010) are important for our study and to ensure controlling for liquidity effects we build and include a measure of options liquidity in our study which will be explained in later sections.

Later, Chen, Lesmond and Wei (2007) find that liquidity is priced in corporate yield spreads. They use a wide range of liquidity measures covering over 4,000 corporate bonds and spanning both investment grade and speculative categories, and find that more illiquid bonds earn higher yield spreads. Hence an improvement in liquidity causes a significant reduction in yield spreads. In their results they control for common bond-specific, firm-specific, and macroeconomic variables, and find it robust to issuers' fixed effect and potential endogeneity bias. Their findings are in line with the default risk literature that neither the level nor the dynamic of yield spreads can be fully explained by current default risk determinants. Building on their findings, in our study we do control for liquidity on the fixed-income market through proxy for CDS liquidity.

In their recent working paper Han, Subrahmanyam and Zhou (WP 2015) study the link between credit and equity markets by focusing on the term structure of credit spreads. They find significant relation between the slope of the credit term structure and future stock returns They expect it to be mainly due to the limited attention and arbitrage costs, since the findings are more prominent for stocks with low institutional ownership, analyst coverage, and stock liquidity.

In this paper, we find evidence for the default risk impacting the equity option prices. Since default risk seems to significantly impact the implicit variance risk premium, default risk premium may be also priced in equity options.

3.3 Research Questions and Hypotheses

Recent financial crisis has shown that for many financial firms equity options experienced high implied volatilities (IVs) and wide option bid-ask spreads, when the underlying CDS spreads went up. Similar trends were perceived during the short-sale ban episode of 2008, when the short sellers migrated to option and CDS markets. Further the recent financial crisis also highlighted the predatory role of both CDS players and short sellers and possible interactions among credit and option markets.[13]

Previous literature has mainly examined how option IV and IV risk premia impact CDS pricing. These studies include Cremers, Driessen, Maenhout

[13]For example, according to a recent WSJ article (Nov 24, 2008) “Anatomy of the Morgan Stanley Panic,”, the CDS traders played a critical role in precipitating bearish sentiment on Morgan Stanley, in turn prompting traders to bet against the firm’s stock by selling it short.

(2008a), Cremers, Driessen, Maenhout and Weinbaum (2008b), Ericsson, Jacobs and Oviedo (2009), Cao and Wei (2010a), and Cao, Yu, and Zhong (2010b), Tang and Hong (2010). However, few empirical studies exist on the reverse relationship. Previous work on the impact of default risk on option pricing has been mainly theoretical, and has examined the pricing of vulnerable options (Johnson and Stulz, 1987), options on defaultable stocks (Bayraktar 2008; Bayraktar and Yang 2010), and theoretical linkages between deep OTM American puts and credit insurance contracts on a company's debt (Carr and Wu, 2011). Carr and Wu (2009) propose and test an internally-consistent joint valuation model for stock options and CDS traded on the same underlying firm, using data for eight firms for the 2002-06 period.

Current research also shows that systematic risk is significantly priced in equity options (e.g. Dennis and Mayhew, 2002; Duan and Wei, 2009). Since both options and CDS instruments help hedge downside firm risks, and hence share put-insurance features, there could exist bi-directional information flows between the two instruments, depending on the magnitude of aggregate shocks and relative liquidity of both markets.

In this paper, we undertake a comprehensive study by addressing several key issues: (1) What drives the pricing of equity options? (2) Are such option risk premia conditional on any default risk characteristics? How do credit markets influence individual option pricing? Does the slope of default term structure, proxying horizon related default risk, impact option pricing? Does firm specific default risk still matter for option pricing, even after controlling for option liquidity and aggregate default risk? (3) Does the magnitude of default-risk impact differ cross-sectionally (i.e. across option maturity and moneyness groups; financial vs. non-financial firms; high- vs. low- rated firms etc.), and over time (i.e. high- and low- credit and liquidity stress periods)? Default is proxied by underlying CDS spreads. We employ an exhaustive panel dataset of (a) individual equity options (Option Metrics) and (b) CDS (Markit) for the period 2002-2010.

From the empirical approach, Dennis & Mayhew (2002) linked Stock Market to Options Market by showing significant connection between market beta and negative skewness of risk-neutral distribution (RND). Vassalou & Xing (2004) connected the FF factors (size and B/M) to the default risk in equity markets. (Their research was focused in the equity market, with no discussion of the options markets). Campbell, Hilscher, and Szilagyi (2008) introduced better factors in the hazard model for default risk. Duan & Wei (2010) found systematic risk in the options equity market prices. The question remained untested is whether by using a better measure of default risk (i.e. derived from CDS market as opposed to DD[14]), we can better measure and explain the default risk in Options Market. In particular asking two key questions: (a) What is the impact of default risk on option implied volatility? (b) What is the impact of default risk on Option skewness? To do so, we set up the following hypotheses:

First, we focus on the impact of default risk on put option prices. This can be measured on two dimensions of cross-sectional and time series. When measured cross-sectionally, we expect the default risk to have a higher impact for distressed firms compared to similar firms.

H1: *Firms with higher credit risk have larger implied volatility compared to other firms controlling for other firm-specific variables.*

[14]Merton's Distance to Default measure

When compared on the time-series level, we expect that for a given firm the movement of the default risk through time can have an impact on its option price. Hence:

H2: *When default risk goes up for a given firm, it impacts underlying implied volatility, controlling for other risks such as aggregate and market variables.*

We can also empirically test whether such a distressed stock with high default risk could also exhibit excess skew in the risk neutral distribution compared to the stock issued by a similar firm, but with lower default risk. Hence, we have the following additional hypothesis:

H3: *When default risk goes up for a given firm, it impacts underlying option skewness, controlling for other risks such as aggregate and market variables.*

In other words, the hypotheses suggest that following the Black-Scholes (1973) option pricing theory, option prices do not depend on how much systematic risk (and default risk per Collin-Dufrense's) is contained in the underlying asset as long as its total risk is fixed. This means when the information content of the option prices are converted into implied volatilities, they should not be related to the default risk of the underlying stock. Therefore, we can build the following two null hypotheses.

Null Hypothesis 1: The implied volatility level of the options is unrelated to the default risk of the underlying asset as measured by CDS spread.

Null Hypothesis 2: The skewness of the options is unrelated to the default risk of the underlying asset.

As explained earlier, many empirical studies (e.g., Bates, 2000; Buraschi and Jackwerth, 2001; Bakshi and Kapadia, 2003; and Jones, 2006) indicate the existence of other risk factors (such as jumps and volatility risk) in option prices. These default risk factors become part of the pricing kernel, and how much they account for the total risk will obviously impacts the characteristics of the risk-neutral distribution. Therefore, our alternative hypotheses can be:

(i) both the level and the slope of the implied volatility curve will depend on the default risk of the underlying asset and, (ii) the amount of default risk will differentiate the price structures of individual equity options.

In addition to the hypotheses above, CDS spread level *and* changes both can explain changes in Option implied volatility. As such we can find the default risk as one of the determinants of the implied volatility of the firm. So we can augment the following robustness hypotheses to the first two hypotheses as follows:

H2: *Default risk (measured by corporate bond's CDS spread level and slope) can significantly explain cross-sectional variances in the implied volatility of the options written on the underlying stocks.*

3.4 Measure of Default Risk and Definitions

3.4.1 Credit Default Swap

CDS (Credit Default Swap) is a swap contract and agreement in which the credit protection buyer pays a fee, usually called a premium, to a credit protection provider in exchange for a payment in the event that a credit default event of a reference asset(s) occurs. The protection buyer is the seller of default risk. The protection seller (or provider) is a buyer of credit risk. The protection seller makes no payment unless a credit default event occurs.[15]

Purchasing a CDS is buying a contingent claim with payoffs that are based on the credit risk of a given entity. In essence, buying a CDS contract is similar to buying insurance against default where the premium payments are determined from the CDS spreads. See for example Das and Hanouna (2006) for detailed description of the CDS contracts.

CDS securities have enabled trading the credit risk of debt. For example fund managers wishing to hedge current credit risk exposures can invest in such securities. Why is CDS a good measure of default risk? In a fairly priced CDS contract, the expected present value of premium payments by the buyer to the seller will equal the expected present value of default loss payments from the seller to the buyer (under the risk-neutral probability measure). Also, since CDS contracts

[15] Tavakoli (2008); "Structured Finance and Collateralized Debt Obligations: New Developments in Cash and Synthetic Securitization", 2nd Edition.

are derivatives, pricing will be undertaken using the risk-neutral probability measure, which is then consistent with obtaining the no-arbitrage price of the security.

Further to the theoretical support of CDS as a viable measure of default risk, there exists empirical and practical support for this choice of measure: CDS is a more dynamic and forward-looking measure of default risk. Credit ratings (as traditionally used by investors, regulators, and managers) are mostly criticized for their slow response in predicting corporate defaults (e.g. cases of Enron, Worldcom), accuracy of their ratings and the conflicts of interest inherent in the agencies' business model (White (2010)). Chava, Ganduri, and Ornthanalai (WP, 2013) provide empirical evidence supporting the view that investors consider CDS markets as a viable alternative credit risk benchmark to credit ratings.

3.4.2 Equity Option Implied Volatility and Skewness

We are interested in measuring the impact of default risk on the prices of individual equity options. From the derivatives literature, we know that the option prices are characterized by the level and slope of implied volatility curves, and in order to numerically measure the impact of default risk on equity options, we should identify the change impact on the implied volatility. Moreover, the skewness will show significant difference for firms with higher default risk.

Empirical work in the derivatives literature have documented option prices fundamentals such as: (i) the Black-Scholes implied volatility is higher than the historical or realized volatility and (ii) the risk-neutral negative skewness is more pronounced than that in the physical distribution, and the index options have a more pronounced volatility smile/smirk than individual equity options (e.g., Jackwerth, 2000; Dennis and Mayhew, 2002; and Bakshi, Kapadia, and Madan, 2003).

These findings demonstrate structural differences between the risk-neutral and physical return distributions. Our research question is partly motivated by this documented structural difference. Using cross-section of options data, we show that a higher amount of default risk leads to a higher level of implied volatility and a steeper slope of the implied volatility curve. Thus, default risk proportion can help differentiate the price structure across individual equity options.

In order to correctly characterize the implied volatility curve, we follow BKM (2003) and Duan & Wei (2009). Duan & Wei argue that assuming a constant slope on the logarithmic scale for the curve (while simplifying the testing procedures) tends to cloud the complex features of the curve. As such, in order to uncover the expected different results for different moneyness regions, we categorize the implied volatility into four distinct moneyness buckets, i.e., $K/S = [0.8, 0.95), [0.95, 1.0), [1.0, 1.05)$ and $[1.05, 1.2]$, and conduct tests within each bucket. Throughout the empirical results, the findings are performed for each of these groups.

We will examine the relationship between CDS spread, as a measure of default risk, and equity option prices. In the next section I explain the data used and how the overlap of these markets gets selected.

3.5 Data and Summary Statistics

3.5.1 Sample Collection and Description

To perform this analysis, we employ data from several databases: Option metrics, Markit data (CDS), COMPUSTAT, CRSP, and Datastream.

The credit default swap data is available from January 2002. As such, Option Metrics data is employed from 2002 to 2010, in order to cover the matching time period with the CDS data. The CDS Markit database is used as the source for CDS data from January 2002 to December 2009. The COMPUSTAT annual files are used to retrieve the firms variables such as "Debt in One Year" and "Long-Term Debt" series, as well as the book value of equity information for all companies. In addition, the daily market values for firms from the CRSP daily files are retrieved. All fundamentals are calculated for firm-specific information.

[Insert Figure 1 here]

3.5.2 Variable Construction

Higher default risk is associated with closer distance to default of the underlying firm, and hence higher risk or implied volatility for that stock. Since such distressed stocks could also exhibit excess skew in the risk neutral distribution compared to stock issued by a similar firm, but with lower default risk, we expect to find that controlling for market risk, the equity option skew is impacted by the default risk of the the underlying reference firm.

3.5.2.1 Firm-specific Accounting-based Variables

We construct three firm-specific variables base on collected data from Compustat and CRSP. For every firm, we define: (1) **Financial Leverage:** defined as total liabilities divided by the sum of total liabilities and market capitalization. (2) **Firm's B/M:** as as a proxy for Tobin's Q. (3) **Firm's Size:** measured by Total assets value.

3.5.2.2 Options Variables

We collect options data from OptionMetrics, which provides daily closing prices, open interest, and trading volume on exchange-listed equity options in the United States. From Options Metrics dataset, we use the strike price, option price (average of bid and ask), as well as implied volatility. The implied volatility is calculated. We calculate the maturity and moneyness for every option written on the firms in the sample.

The implied volatility computed for each option, is calculated by taking the option price as the settlement price, or the last traded option, or to the midpoint of the closing bid and offer prices (in this order of availability). The underlying price then is synchronized at best with the option price. The implied volatility is computed using industry-standard equity pricing models: Black-Scholes for

European-style options, and Cox-Ross-Rubinstein binomial tree for American-style options. [16]

3.5.2.3 Default Risk Variables

Markit database reports the spread on the 1-year, 5-year and 10-year contracts. For the purpose of our research and in order to be dealing with the highest liquidity, we limit the data to 5-year spreads only. As well as the spread level, we need a proxy for controlling for the liquidity risk of the CDS, since these contracts are traded OTC and hence fairly illiquid. The variable "number of contributors" is also employed to proxy for liquidity of CDS trades. To remove any inconsistency we build a relative measure of CDS liquidity for our study in order to use this proxy effectively.[17]

3.5.2.4 Other Control Variables and Time Periods

In order to control for the systematic risk in the market, we use daily VIX data and account for the variation in market volatility. Also, return on S&P 500 index is used as the market return control variable or R_f when needed in the models. In addition, we use the aggregate measures such as 3-month T-bill, yield curve slope, aggregate default spread, and TED spread in our models.

[16] For each strike/exercise date and for put/call options, other variables such as the interest rate is calculated from a collection of continuously-compounded zero-coupon interest rates at various maturities. The zero curves used for the European options are derived from BBA LIBOR rates. For underlying paying dividents, they are estimated using a "constant divident yield" assumption based on the most recently announced divident payment. You can see "An introduction to OptionMetrics implied volatility data" from RiskMetrics for further details on volatility surface

[17] We demean the number of contributors for each firm by deducting the total average of all firm's number of contributors. Please see the Variables Glossary for more details.

Special attention is given to two sub-periods of financial distress and non-financial distress. It is argued that the recent financial distress period started in August 2007, and ended in March 2009. One of the objectives of this study is to empirically investigate the distinction of default risk impact on equity options market during financial crisis and severe distress times, since the linkage between the high default risk and systemic risk is also of great importance.

Systemic risk was greatly observed during the recent financial crisis period. It is also important to note that the correlations between stocks and aggregate market are much larger for downside moves, especially for extreme unexpected downside moves, than for upside moves. This is in-line with findings of Ang and Chen (2002) on "Asymmetric correlation". This fact was also empirically evident during the recent financial crisis. In this paper, we split the sample data with a perceived regime switch and try to identify the difference in the impact and magnitude of default risk in options market among these sub-samples.

Furthermore, we are also interested in the impact of industry on our findings. We hence test the difference between financial and non-financial firms. It is important to understand how different the impact of default risk on options would be when the underlying firm is a financial firm vs. a non-financial firm. For this purpose, we run the tests on both sub-samples.

3.5.3 Descriptive Statistics

Table 1 displays descriptive statistics for our final sample. There are total of 550 unique firms after removing firms that are not publicly traded, as well as applying minimum liquidity filters for CDS and removing options. We also remove any options for which put-call parity doesn't hold. On the 550 firms total of 67015 unique put options have been written which are getting fully tested in our model(s). Panel B shows the count and range of identifying variables. It is important to note that the cross-sectional variable for the regression tests would additionally include Option IDs so that we distinguish between various option contracts written on one firm (same issuer). We use this identification in order to control for the fixed effects of the same option contract across observations.

Panel C of Table 1 shows the number of unique options and CDS contracts for each firm in the final sample. The table shows the count as well as the summary statistics for each year in the time period studied. The average implied volatility ranges around 30%. The implied volatility (as expected) increases to double its average value during the financial crisis period in 2008 and 2009. Similar pattern is observed for average CDS, with the minimum average value of 0.69 in 2006 and rising to its maximum average of 1.93 in 2009.

In order to perform additional studies we also split the sample into financial firms and non-financial firms. In our final sample there are total of 81 financial and 469 non-financial firms. We repeat running each specification on each of these subgroups. Fixed effects for time and firms are also accounted for, as well as clustering issues.

Panel D of Table 1 reports the distribution of the options and CDS characteristics in various maturity and moneyness categories. We define three moneyness and three maturity groups and divide our sample into nine different bins.[18] Moneyness is defined as the ratio of spot price divided by strike price for calls, and the ratio of strike price divided by spot price for puts. Across put options covered by OptionMetrics, the distribution across moneyness and maturity appears to be fairly uniform. We note that at-the-money options (those with moneyness between 0.95 and 1.05) are heavily traded, and short maturity has relatively larger number of traded options.

3.6 Methodology

We explained the hypotheses to be tested and the details around the extensive merged dataset. Our dataset is a pooled time-series and cross-section unbalanced panel. A pooled unbalanced dataset requires extra caution when running regressions because some findings can be spurious and simply capturing the unbalanced variation of the observations. As such, we have paid extra attention to analyze such a panel by controlling for the time and fixed effects. We control for fixed effects on two dimensions: firm level, and option level. For firm level, we use cluster effect and for option level we use fixed effects.

Thompson (2006) and Petersen (2009), in their separate works explain this issue and provide detailed analysis and recommendation on the performance of various approaches. Following their suggestions we conduct several adjustments in our regressions including adjustment for firms' clustering, as well as controlling

[18] We define Short, Medium, and Long maturity buckets as "fewer than 70 days", "between 71 and 120 days, and "greater than 121 days" respectively.
We also define OTM put options as the options with moneyness less than 0.95, ATM with moneyness between 0.95 and 1.05, and ITM with moneyness greater than 1.05.

for the time and fixed effect.

As a result, the baseline specification we use in our regression analysis is provided below, which tests the first and second hypotheses:

$$\begin{aligned} OptionImpliedVol_{i,t} &= \alpha + \beta_0 \times OptionImpliedVol_{i,t-1} + \beta_1 \times CDS_{i,t} \\ &\quad + \beta_2 \times LiquidityProxy_{i,t} + \beta_3 \times MktVars_{i,t} \\ &\quad + \beta_4 \times FirmVars_{i,t} + \epsilon_{i,t} \end{aligned} \tag{3.1}$$

where $OptionImpliedVol_{i,t}$ is the implied volatility of the options written on the underlying firm i at time t, $CDS_{i,t}$ is the level of the CDS at each point in time. $MktVar$ includes market variables such as VIX, return on S&P500, Aggregate default risk, TED spread, yield curve slope, and finally $FirmVars$ includes firm-specific fundamental variables such as size, financial leverage, and B/M ratio.

In this paper, we focus on the findings for put options only. However, it would be interesting for future studies to see whether the relationship above is different for put and call options.

Finding positive significant coefficients for the CDS variables will support our hypothesis that default risk impacts equity options market.[19]

[19] We are trying to report the presence of default risk and not necessarily the magnitude of the default risk premia. Measuring the specific impact on pricing can only be established through two pass-test Fama-French type procedures. For example, economic significance using option strategies can be implemented by identifying firms during a credit event (e.g. rating downgrades) and predicting call /put prices and hence taking positions in the straddles accordingly.

To answer the second research question of the paper which discusses the impact of default risk on IV skew, we test the following model:

$$\begin{aligned} Skew_{i,t} &= \alpha + \beta_0 \times Skew_{i,t-1} + \beta_1 \times CDS_{i,t} \\ &\quad + \beta_2 \times LiquidityProxy_{i,t} + \beta_3 \times MktVars_{i,t} \\ &\quad + \beta_4 \times FirmVars_{i,t} + \epsilon_{i,t} \end{aligned} \tag{3.2}$$

where *Skew* is a measure of option implied volatility skew. At each point in time, volatility skew is defined as the difference between average OTM implied volatility and average ATM implied volatility for each issuer.

On Level Regressions vs. Change Regressions: As demonstrated in Eq. 3.1 and 3.2, we only perform level regressions and not changes. This is mostly because we are interested in cross-sectional option pricing and follow existing papers with similar data framework (e.g. Cao et al 2010). In addition, the irregular intervals or missing data for CDS observations further limits design of any change regressions. From a statistical perspective, first differencing is appropriate if the dependent variable and regressors are integrated, but this is difficult to determine for our irregularly spaced data. Note also that if one accepts the unit root hypothesis, differencing of the data may improve the efficiency of the resulting estimates, but the levels regressions yield consistent point estimates. This ensures that our findings are conservative with the levels regressions.

We also repeat the regression above for testing financial and non-financial firms.

To do so, we repeat the same specification with interactive terms. The corresponding regression hence will be:

$$
\begin{aligned}
OptionImpliedVol_{i,t} &= \alpha + \beta_0 \times OptionImpliedVol_{i,t-1} \times Fin + \beta_1 \times CDS_{i,t} \times Fin \\
&\quad + \beta_2 \times LiquidityProxy \times Fin + \beta_3 \times MktVars_{i,t} \\
&\quad + \beta_4 \times FirmVars_{i,t} + \epsilon_{i,t}
\end{aligned} \tag{3.3}
$$

where we define variable Fin to have a binary value of 1 if the underlying firm belongs to the financial industry and 0 if not. Significant results on the interactive terms will indicate significantly different impact of default risk in option markets for financial vs. non-financial firms. The results show that the impact of default risk is more prominent for financial forms, with higher levels of significance. This is consistent with the "Cascade Effect" notion in the Financial industry, where the risk of default can be contagious among financial firms and during distress periods.

3.7 Empirical Results

3.7.1 Correlation Analysis and Mitigating Multi-colinearity

We are interested in finding whether default risk impacts equity options or not.[20] To do so we test both contemporaneous as well as lagged variables of the parameters we are interested in. It is also important to consider the risk of colinearity

[20] We focus on the existence of the impact. In order to specifically measure the pricing impact, one should develop a more general option pricing model which nests the model devoid of such a factor and test for the significance of the extra factor. Alternatively, we can see if using the additional information included in CDS spreads, over and above the information incorporated in IV, would enable market participants to earn excess returns. At this point, we are not building a model of excess return to find the magnitude of impact.

when the pairwise correlation of the variables are high. [21]

As such, and with the colinearity effect in consideration, Table 2 displays the pairwise correlation amongst all potential variables in the regression specification. The first column shows the correlation values between the primary dependent variable (Option-implied volatility) and the rest of the variables. The significant values (both economically and statistically) are bolded. In this column the high correlation values are no issue since we are interested in the explanatory power of each of the independent variables for implied volatility.

It is, however, important to carefully note and care for the rest of the columns as they show the pairwise correlation between the independent variables. Two significant correlation coefficients of *CDS spread* with *lagged CDS spread* and *lagged implied volatility* lead us to conclude that if included in a specification together these values need further attention. As such, in the next sections, we run multiple types of regression analyses, *with* lagged variables, and *without* lagged variables. The lagged CDS will later be employed as the instrumental variable for CDS. The strong correlation shows that it meets preliminary requirements.

It is also important to observe the significant correlation between S&P 500 Index return and some of the macroeconomic variables such as VIX.

An additional observation as a result of the high correlation is to perform tests of endogeneity for two key variables of the research and the fact that the causality relationship between CDS spread and the implied volatility can go in both directions. This will be tested and discussed further in the robustness sections

[21] This is deemed to be more important when dealing with volatility and distress risk at daily frequency since these variables can be persistent.

due to the fact that IV and CDS can be endogenous variables, determined by a set of common explanatory variables. The need for this test is justified by the strong correlations of each of these variables with most of the firm level and market level variables.

3.7.2 Regression Results

The correlation results discussed in the previous section have already shown the existence of comovement of IV and CDS level for the underlying firms. In this section, we control for standard expected variables and explain each of the following tables presented for the regressions. It is important to note that we have a pooled panel data set with essentially two dimensions of cross-sectional variables: firms, and various options written on the same firm.

In the following subsections we continue the findings by testing for different hypotheses, sub-groups, stratifications, robustness (endogeneity), as well as options liquidity and financial leverage impact. We also explain the effect of firm-level variables included.

In order to be making comparative and consistent analysis we need to categorize the options into buckets of similar moneyness and maturity. As such we create three groups for maturity: "long", "medium", and "short"; and three groups for moneyness: deep in the money, deep out of the money, and, near or at the money. The result is nine groups on which we perform the regressions to record the special differences or similarities. Tables 3, 4, and 5 show results for "short maturity". In Table 3 we present four different regression models for "short maturity, ITM" put options. We start with excluding the default risk variables in model (1) and increment by adding CDS spread and CDS liquidity proxy in model (2). Then we augment by including lagged CDS in model (3). Finally, model (4) includes lagged IV, CDS, and CDS liquidity proxy. We shall refer to model (4) as the baseline model which will be tested across various buckets of maturity and moneyness as shown in hypothesis specification (3.5).

Table 3 shows that CDS is consistently a significant variable in explaining implied volatility. We also find the CDS liquidity proxy (measured through standardization of the number of contributors), is consistently significant with negative sign. This shows that more liquid CDS contracts signal smaller implied volatility while IV increases with as illiquidity risk increases. Intuitively, investors need to be compensated for the liquidity risk and hence a higher IV is expected when the option is less liquid, consistent with the findings presented.

Table 4 repeats the previous setup of four key regressions on "short maturity" and deep OTM put options. The interesting finding is that the explanatory power of CDS disappears when we move to deep OTM put options. Not CDS nor CDS liquidity can explain the IV variation based on Table 4 findings, while the aggregate variables and firm level variables gain more power in the IV regression for short deep OTM options.

Tables 5 to 7 perform the baseline regression (explained in equation 3.5) for short, medium, and long maturity groups. In each regression table we split data into the three moneyness groups and show the findings.

Table 5 shows that for short maturity, CDS is only significant for ITM options and loses significance for ATM and OTM groups. This finding is reversed for medium and long maturity groups where CDS is more significant for ATM and OTM groups. By looking at all the tables and focusing on the various bins tested, we find that for deep in the money and deep out of the money options the default risk as measured by CDS level is more economically significant.[22] This is in-line with our hypothesis that higher default risk is more significant for deep out of the money (subject to higher default risk) and deep in the money (more liquid) equity put options.

3.7.2.1 Default Risk and IV-Skew

We define option implied skewness besides the implied volatility measure. For each point in time we also compute an implied volatility skewness measure, which is the difference between the implied volatility of all out-of-the-money put option, and the implied volatility of an at-the-money put option with a strike-to-spot ratio closest to 1, for all contracts issued by the same firm. The implied volatility skew is closely related to the skewness of the risk-neutral equity return distribution and we expect it to be positively related to the CDS spread. In simple words, for our sample of put options:

$$Skew_{i,t} = AVERAGE_{OTM}(IV_{i,t}) - AVERAGE_{ATM}(IV_{i,t}) \tag{3.4}$$

Tables 8 to 10 show the results of applying the baseline regression to the IV-Skew. The results vary based on the option bins focused on which is expected given

[22]Calculation of the economic significance is provided in Table 11.

the definition of IV-Skew explained above. Among the 9 bins, skewness is best explained for by CDS in medium and long maturity groups and most prominent for OTM options. The signs and significance vary as well, resulting in no consistent finding across all 9 bins. As a result, the CDS level cannot conclusively explain the variation of skewness across all groups of moneyness and maturity; For all put options the findings vary bin by bin, with most statistically and economically significant values belonging to higher maturity and smaller moneyness.

3.7.2.2 Effect of Firm-level Variables

Reviewing the findings so far (for example, Table 3) we find that the firm size is proved to be a significant variable with negative coefficients which implies the size effect on the option implied volatility of the firm. As expected smaller firms have higher risk of default and higher implied volatility which is consistent with the theoretical understanding. Moreover, small firms also exhibit more volatility than large firms so the significance of this coefficient may also reflect this stylized fact.

Leverage is also a negative significant factor in the model for short-maturity options. However, the significance drops as the maturity increases and the results are not conclusive for medium and long maturity groups. Once the model includes lagged variable the significance of leverage is again dropped. To further uncover the impact of leverage we will run additional tests in Table 14.

Market to Book ratio is consistently significant with similar level and power as size impact which is consistent and expected. Firm's stock return is positive and significant in explaining IV changes in some of the bins and insignificant in the rest. This is also consistent with the options pricing theory as higher values of IV expect higher return on the underlying stocks (which can be due to compensation for the cost of higher default risk.)

3.7.2.3 Macroeconomic and Market Variables

We control for all standard macroeconomic variables and additionally include a complete model for default risk.

In the regression tables we control for VIX, return on S&P500, Yield curve slope, Treasury rate, Aggregate default Spread, and TED spread [23]. Once all included, market return becomes insignificant in explaining the IV. However, the CDS remains a prominent significant factor even after controlling for all macroeconomic and market variables as shown in Tables 3 and 4.

We later use the macroeconomic and market variables to test for endogeneity. Together with firm variables, they are the common set of explanatory variables that explain both CDS and implied volatility and hence potential for endogeneity existence.

3.7.3 Default Risk and Economic Significance

So far we showed that default risk is priced in equity options and the results are statistically significant. The very important resulting question is how much of an economic significance would default risk cause in the pricing of options as measured through option-implied volatility?

[23]The TED spread is the difference between the interest rates on interbank loans and on short-term U.S. government debt (T-bills). It is calculated as the difference between the three-month LIBOR and the three-month T-bill interest rate and is an indicator of perceived credit risk in the general economy

To test this, we run a sigma shock test on CDS spreads and show how much of difference would this cause in the level of IV and IV-skew. Table 11 presents the economic significance results for the applicable models and variables. We focus only on the bins with the significant default risk (i.e. the results shown in Tables 3, 4, and 8 as discussed in previous section). The findings show that on average the default risk contributes to 3% change in option-implied volatility, for one sigma shock in CDS spread. Next, we test and show that the findings are robust to endogeneity.

3.7.4 Tests of Endogeneity

Standard linear regression models assume that errors in the dependent variable are uncorrelated with the independent variable(s). When this is not the case (for example, in our case the relationships between variables can be bidirectional), linear regression using ordinary least squares (OLS) no longer provides optimal model estimates.

Two-stage least-squares regression uses instrumental variables that are uncorrelated with the error terms to compute estimated values of the problematic predictor(s) (the first stage), and then uses those computed values to estimate a linear regression model of the dependent variable (the second stage). Since the computed values are based on variables that are uncorrelated with the errors, the results of the two-stage model are optimal.

In our study, we can identify different possible explanations for default risk: Ideally, we should employ instrumental variables (for example, Faulkender and

Petersen (2006)) or natural experiments (for example, Leary (2009), Sufi (2009b), Lemmon and Roberts (2010)) to distinguish among the possible explanations and/or to establish causality between default risk and implied volatility. However, as we discuss below, it is difficult to identify one instrumental variable that can fully deal with endogeneity in our setting.

The instrumental variable procedure requires us to specify an instrument for implied volatility. Past CDS spread $(i, t-1)$ is a natural candidate for an instrument, since it is correlated with true time t CDS spread (i, t) but is possibly unrelated to the measurement error associated with CDS spread sampled one month later. We run the two stage regressions and archive the results in Tables 12a-12c.

The strong first regression results guarantee the choice of right instrument while the instrumental variables regressions (including control variables) shows the 2SLS regression results. [24]

In our baseline regression endogeneity arises because CDS potentially depends on the same firm level and market level variables as IV. Hence, to model endogeneity, we use the original IV regression and define CDS in terms of firm level and market level variables. In short,

$$CDS = f_1(LaggedCDS, FirmVars, MarketVars)$$
$$IV = f_2(CDS, CDSLiqProxy, LaggedIV, FirmVars, MarketVars) \qquad (3.5)$$

[24] We also perform a postestimation endogeneity test. We perform both Durbin and Wu-Hausman tests and the results are significant.

By introducing the common variables in the two stages above, we are now controlling for endogeneity. Table 12 shows the result of endogeneity tests for all moneyness, and three subgroups of ITM, ATM, and OTM.

Stage 1 (panel A) results show that lagged CDS is in fact a suitable variable for the instrumented variable CDS. Stage 2 results (shown in panel B) and the post-estimation tests (shown in panel C) provide two specific implications. The first immediate inference is based on the two post-estimation endogeneity tests of Durbin and Wu-Hausman. Based on the p-values of these tests the null hypothesis of exogenous variables is rejected indicating that IV and CDS are endogenous variables that are determined by a common set of explanatory variables. Hence, the findings of second stage regression (which is after controlling for endogeneity) is important.

The second important and incremental inference is driven from panel B results. We can see that CDS remains a significant value even after controlling for endogeneity. The economic significance, however, is relatively decreased. So, by showing the endogeneity tests and two stage least square regression results, we find that the default risk is impacting equity options and the finding is robust even after controlling for endogeneity of option implied volatility and CDS. Moreover, the endogeneity is tested across different moneyness bins. CDS spread remains a significant factor across bins after controlling for endogeneity.

To further explain why the results are important after controlling for endogeneity, note that CDS "impacts" option prices and may not necessarily " drives" them. So, what is added by default risk impact? The equity option prices are determined by risk premia attributed to volatility, jumps, default events, etc. The CDS spreads and term structure carry forward-looking default information that can enter option pricing through risk premia. It is however true that same information is embedded in options too. If completely efficient, CDS should have no information content. However, since CDS variables have significance after endogeneity [25], we believe that CDS measures are impacting risk premia that in turn determine option pricing.

3.7.5 Impact of Industry: Default risk for financial and non-financial firms

As part of the study, we are interested in observing the default risk in equity options of financial vs. non-financial firms. To do so, we split the data sample into financial and non-financial firms. We then repeat the main regressions in order to measure the difference in the economic and statistical significance of the default risk impact on options market. Table 13 shows the findings for this stratification.

Table 13A shows the results with dependent variable IV, and 13B records the results for dependent variable Skewness. By focusing on the CDS spread coefficient, we observe a significant increase in the CDS spread coefficient for financial firms compared to non-financials (the coefficient increases from 0.007 to 0.05). This shows that the impact of default risk on option prices are more prominent for financial firms than non-financial firms. The results for skewness, is reversed with showing significance only for non-financial firms. [26]

There are a few potential reasons attributed to why the impact of default risk on implied volatility is different for financial than for non-financial firms. For example, one reason can be the higher level of contagion effect across financial

[25] and later will show CDS is significant in explaining variance risk premium too

[26] Focusing on the characteristics of financial firms, the skewness results can partially be due to the fact that these firms are more exposed to higher market risk, limiting the skewness variation.

firms during distress periods. In addition, there can be a cascade effect resulting from a default event of one financial firm to the rest of the financial industry. Both result in stronger impact of default risk on equity options for the financial firms.

3.7.6 Impact of Financial Crisis

After the 2008 financial crisis, there is particular interest in finding the default risk impact and its significant difference for equity options during crisis and non-crisis periods. To do so, we split the data sample into two sub-samples of crisis and non-crisis periods. We then repeat the main regressions in order to observe the difference in the economic and statistical significance of the impact of default risk. Table 15 shows the findings for this sample stratification. Table 15A shows the results for implied volatility and Table 15B records the results for skewness. By focusing on the CDS spread coefficient, we observe that default risk loses its significance to explain IV during crisis period, however, it is very significant in explaining skewness of option implied volatility during the same period of crisis. The CDS liquidity remains significant in both periods for explaining both IV and skewness. This shows that during the crisis period the higher the default risk of an option the larger the skewness of the option implied volatility and hence a larger mispricing spectrum. This also is consistent with the actual observations during the 2008 financial crisis period.

3.7.7 Option Market Liquidity and Default Risk

One question remaining can be whether the risk observed is due to options liquidity and not purely default risk. To investigate this question further, in addition to the comprehensive regression tests performed, we run tests of options liquidity impact. Following the findings of Cao & Wei (JFM 2010) we investigate all the options liquidity measures introduced and tested and choose the best measure design for our model.

We define and use the best liquidity measure for options that is available based on our dataset. Working with daily and monthly options data, we follow Cao and Wei's laundry list of Options Liquidity definitions. We choose "Proportional bid-ask spread" which is available on a daily basis for each option trade, as the measure of Options Liquidity. It is defined as:

$$PBA = \frac{\Sigma_{k=0}^{n} VOL \times \frac{Ask-Bid}{(Ask+Bid)/2}}{\Sigma_{k=0}^{n} VOL}$$

where PBA is the proportional bid-ask spread, and VOL is the observed volume of the option on in that date. For put options sample, we expect the PBA to be around 0.13 with standard deviation of 0.04. We calculate the PBA and control for the illiquidity in the second round of tests.

The results of liquidity measures are provided in Table 16. We find the statistics on the PBA Option liquidity measure to be consistent with Cao & Wei's values.

We next add the options liquidity measure to the baseline regression model. The results are shown in Table 16A for IV and 16B for skewness. The findings affirm our hypothesis. Option liquidity is a significant factor but even after inclusion of option liquidity measure, the CDS spread is still a significant factor for the implied volatility and hence providing evidence that default risk is impacting the equity options.

3.7.8 Implied volatility risk premium

The findings so far show the impact of CDS on IV, using credit market information to explain the pricing of equity options. But can credit default swaps, this recently developed and rapidly growing segment of the derivatives market, also be measuring the deviation of option-implied volatility from realized volatility? The significance of CDS in option-implied volatility regressions motivates this since any mispricing in the equity options market can also be captured through the option-implied volatility risk premium.

To test this, we measure volatility risk premium as the difference between historical volatility and option-implied volatility and run the baseline panel regression. The findings are presented in Table 17 and show that CDS spread is a significant factor in explaining the implied volatility risk premium. The results are also consistent with findings of Cao, Yu, and Zhong (2010) where they find that implied volatility can explain the time-series behavior of CDS spreads both because it forecasts future volatility better, and because it captures a volatility risk premium. Our panel regression results show that CDS has information content that impacts the option-implied volatility risk premium as well.

3.7.9 Tests of Nonlinearity: Log Regressions

We showed that default risk is a significant factor in a linear regression model of implied volatility. What if the relationship observed is due to limitations of a linear modeling of this relationship? To test this, we run the non-linear regression by creating log variables of volatility, default risk, and return variables. The results provided in Table 18 show that the findings remain significant and prominent on a log regression model.

3.8 Conclusion

The recent financial crisis once again put credit risk (default risk) and the risk of contagion distress periods under the spotlight. This study relates to recent literature that explains how default risk can get injected from the fixed income market to the equity options market and why it is important to account for default risk in the pricing of equity option and implied volatility.

What is the impact of default risk on option pricing? We study this question in detail by empirically examining to what extent the firm-specific attributes and systematic variables matter in pricing individual equity options. This paper studies the primary determinants of option IVs cross-sectionally and over time, and measures the contribution of default risk on equity option pricing. Since Credit default swaps (CDS) are similar to out-of-the-money put options in that both offer a low cost and effective protection against downside risk, we use CDS spread as credit risk proxy to investigate the effects of credit risk on put pricing.

By merging an exhaustive dataset of CDS data (Markit), Options data (OptionMetrics), and equity data we test for the significance of default risk in option-implied volatility of the underlying firm. We find that after controlling for market variables and firm variables, the default risk is still a significant driver of option implied volatility. This significance is prominent in specific maturity and moneyness put option categories, and robust to endogeneity issues.

Furthermore, default risk significance is strongly evident in ITM bins for shorter maturities (more liquid) and OTM bins for longer maturity (subject to higher default risk impact) put options. The results are also tested and vary for segments of industry (larger impact for financial firms), and distress periods. We find that during crisis and distress periods, the significance of cross-sectional default risk is vanished, and mainly captured by the market level default risk variables. This can relate to the effect of systematic risk during distress periods of the whole economy.

Lastly, default risk is proved to be a significant factor in explaining the volatility risk premium which shows the deviation of implied volatility from realized volatility (here proxied by historical volatility).

These findings significantly contribute to the recent literature on options pricing and credit risk by empirically showing the flow of default risk from fixed income markets to equity options market and the need for inclusion of default risk in equity option pricing models.

3.9 Essay II Appendices

3.9.1 Appendix A: Complete Variables Glossary

In this section, the detailed definitions for each of the variables used in the baseline regression(s) are provided.

• **CDS spread:** 5-year maturity CDS spread, daily observations from Markit. When data is missing for one day in between two observations, it gets interpolated. Any consecutive missing for more than one day, is treated as missing value and automatically is dropped from the panel regressions.

• **Option-implied volatility (IV):** Put option IV, daily observations, daily observations, directly from Option Metrics. [27]

• **CDS Liquidity Proxy:** equals to $(No.ofContributors – AvgNo.ofContributors) \times CDSspread$, or "Demeaned number of contributors times CDS spread". The number of contributors are retrieved directly from Markit.

• **Firm Return (%):** For each underlying firm daily price is retrieved from CRSP database. $FirmRet = \frac{P_1 - P_0}{P_0}$.

• **Historical Volatility (%):** Standard deviation of a fund's daily returns over a rolling 6-month period.

• **Market to Book Value:** Tobin's Q calculated with data from Compustat quarterly files.

[27]IVs in OptionMetrics are retrieved from the Volatility Surface data file. The calculated interpolated implied volatility for each option on each day, uses a methodology based on a kernel smoothing algorithm. The data is first organized by the log of days to expiration and by "call-equivalent delta" (delta for a call, one plus delta for a put). A kernel smoother is then used to generate a smoothed volatility value at each of the specified interpolation grid points. At each grid point on the volatility surface, the smoothed volatility is calculated as a weighted sum of option implied volatilities. (See OptionMetrics Manual for additional details.)

• **Debt ratio:** = Total Debt/Total Asset; data from Compustat quarterly files; defined as total liabilities divided by the sum of total liabilities and market capitalization.

• **VIX:** Daily market volatility, $VOLATILITYS\&P500(^VIX)$, retrieved from CRSP, and available on OptionMetrics.

• **TED Spread:** The difference between the interest rates on interbank loans and on short-term U.S. government debt (T-bills). It is calculated as the difference between the three-month LIBOR and the three-month T-bill interest rate. Data files for each variable retrieved from Datastream.

• **IV Skew:** As formulated in the paper, IV skew is defined for each observation as "Avg OTM implied volatility – Avg ATM implied volatility":

$$IVSkew_{i,t} = AVERAGE_{OTM}(IV_{i,t}) - AVERAGE_{ATM}(IV_{i,t})$$

3.9.2 Appendix B: Merging CDS Markit and Option Metrics

There are total of 1833 Ticker/Comp Name combination in the Markit Database.Out of these 1833 , 34% are private or subsidiaries (See appendix A on this clean-up activity). Removing these , would leave a remainder of 1203 public to work with. The next step is to limit data to firms with Options. Matching the 1203 with Options Metric data, we are down to 1063 which both have CDS written on the underlying bond, and options written on their underlying company. The sample that we work with would have 1063 unique firm IDs. (see chart below)

3.9.3 Cleaning up Option Metrics Data

I downloaded all Option Metrics data from 1996 to 2011 inclusive. The Option Prices file (under Option Metrics Data files¿¿Options) gets downloaded with all possible variables (32 variables in total) in monthly files format from 1996 until 2010. (The average volume for every month is 1-2GB)
There are 32 listed variables downloaded, per below:

FILTERS:
1) DROP UNNECESSARY VARS ($ss-flag, index-flag, exchange-d, issue-type$) keep if $cp-flag ==$ "C" | $cp-flag ==$ "P" keep if $ss-flag == 0$ (121656 observations deleted) keep if index$_f lag$ == 0 (536045 observations deleted) drop if $exchange - d == 32768$ (0 observations deleted) keep if $issue - type ==$ "0" (1681856 observations deleted)
2)SORT: prior to saving the final file, I "sort secid optionid date" to organize the data; it makes it much easier in the data matching, & give the compress command one more time before saving the datafile.
3)LITERATURE FILTERS: Many in our sample do not have actively traded options. The choice of non-zero open interest emphasizes the information content of options that are currently in use by market participants. We also drop all zero volume options. I also exclude all options that violate Put-Call parity. For the first round of the analysis, I keep only put options.

3.9.4 Compustat Fundamentals Data Collection

We collect and compute the quarterly Company Variables (i.e. Size, Book/Market, and Leverage) for all of the in the sample. I also winsorize data and store statistics both before and after winsorization.

For merging with Option Metrics and CDS, we keep the pre-winsorization file.

3.9.5 Report on CDS Markit Data Clean-up for TICKERS & COMP NAMES

Markit Look-Up Dimensions: TICKER, Company Name, Date Period(s)

Issues:

1. Tickers may get recycled and as such multiple matches available when merged with other databases
2. Company names are in short form or miss spaces/dots in Markit so cannot automatically be matched. Need manual review.

Solution:

Review the complete Ticker/CompName list from Markit and verify their identification (i.e. match with the unique PERMNO from CRSP).

- Use Google/Yahoo Finance search engines.
- Confirm Private/Public Companies
- Find alternative names and tickers

Manual Correction Steps

1. Look up the ticker in CRSP full database
2. If the ticker and company name match fully, there should be one unique PERMNO. Record any alternative names.
3. If there are alternatives, provide correct match for the time period provided in Markit.
4. If there are no matches, use Google and Yahoo finance searches. – Is the company a Private company?

Results Statistics

There are total of 1833 unique Ticker/Comp Name combination in the CDS Markit Database. The final correction and inclusion provided final match of 1203 companies in total.

From the remaining, the provided correction notes indicate that:

a. 15 firms are subsidiaries with parent company not found before, so parent ticker is provided.
b. Total of 100 unmatched tickers are Private firms.

c. Total of 340 unmatched tickers are Subsidiaries with their parents present in the database.
d. Fewer than 10 firms are result of mergers or acquisitions that are still valid for inclusion. Details provided.
This is the result of individual scrubbing of each firm/Ticker.

3.10 References

Anginer, D., and Yildizhan C., 2010. Is there a Distress Risk Anomaly? Corporate Bond Spread as a Proxy for Default Risk, Working Paper.

Ang, A. and J. Chen, 2002. Asymmetric Correlation of Equity Portfolios, Journal of Financial Economics, 63 (3), 443-494.

Bayraktar, E., 2008. Pricing Options on Defaultable Stocks, Applied Mathematical Finance, 15(3), 277304.

Bayraktar, E., and Yang, B., 2011. A unified framework for pricing credit and equity derivatives. Mathematical Finance, 21(3), 493-517.

Bakshi, Gurdip and Nikunj Kapadia, 2003. Volatility risk premiums embedded in Individual equity options: Some new insights, Journal of Derivatives, 10, Fall, 45-54.

Bakshi, G., Kapadia N., Madan D., 2003. Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options, Review of Financial Studies, 16(1): 101-143.

Bates, D. S., 2000. Post-'87 crash fears in the S&P 500 futures option market. Journal of Econometrics, 94(1), 181-238.

Beber A., Brandt M.W., 2010. When It Cannot Get Better or Worse: The Asym¬metric Impact of Good and Bad News on Bond Returns in Expansions and Recessions, Review of Finance, 14(1), 119-155.

Bharath, S. T., and Shumway, T., 2008. Forecasting default with the Merton distance to default model. Review of Financial Studies, 21(3), 1339-1369.

Bollen, N and Robert E. Whaley, 2004. Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?, Journal of Finance 59, p.711-753.

Bongaerts, D., De Jong, F. And Driessen, J., 2011, Derivative Pricing with Liquidity Risk: Theory and Evidence from the Credit Default Swap Market. Journal of Finance, 66: 203240.

Buraschi, A. and J. Jackwerth, 2001. The price of a smile: hedging and spanning in option markets, Review of Financial Studies, 14(2), 495-527.

Câmara, A., Popova, I., and Simkins, B., 2012. A comparative study of the probability of default for global financial firms. Journal of Banking & Finance, 36(3), 717-732.

Campbell, John Y., and Glen B. Taksler, 2003. Equity Volatility and Corporate Bond Yields, Journal of Finance, 58, 23212350.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008. In Search of Distress Risk, Journal of Finance, 63, 2899-2939.

Carr, P., and Wu, L. 2009a. Variance risk premiums. Review of Financial Studies, 22(3), 1311-1341.

Carr, P., and Wu, L. 2009b. Stock options and credit default swaps: A joint framework for valuation and estimation. Journal of Financial Econometrics, nbp010.

Cao, M., Wei J., 2010. Option market liquidity: Commonality and other characteristics, Journal of Financial Markets, 13, 1, 20-48.

Cao, C.,Yu F., Zhong Z., 2010. The information content of option-implied volatility for credit default swap valuation, Journal of Financial Markets, 13, 3, 321-343.

Carr P. and L. Wu, 2011. A Simple Robust Link Between American Puts and Credit Protection, Review of Financial Studies, 24(2), 473-505.

Carr, P., Madan, D. B., 2010. Local volatility enhanced by a jump to default. SIAM Journal on Financial Mathematics, 1(1), 2-15.

Chava, S., Ganduri, R. and Ornthanalai, C., Are Credit Ratings Still Relevant? (WP April 5, 2015). Available at SSRN: http://ssrn.com/abstract=2023998

Chava, S., and Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. Review of Finance, 8(4), 537-569.

Chava S., Purnanandam A., 2010. Is Default Risk Negatively Related to Stock Returns? Review of Financial Studies, 23(6): 2523-2559.

Chen, L., Lesmond, D. A. and Wei, J, 2007. Corporate Yield Spreads and Bond Liquidity. Journal of Finance, 62: 119149.

Chen, N., Kou, S. G., 2009. Credit spreads, optimal capital structure, and implied volatility with endogenous default and jump risk. Mathematical Finance, 19(3), 343-378.

Collin-Dufresne, P., Goldstein, R. S., Martin, S., 2001. The determinants of credit spread changes, Journal of Finance, 56:2177207.

Coval, J. D., and T. Shumway, 2001. Expected option returns, Journal of Finance 56, 983-1009.

Cremers, M., Driessen, J., Maenhout, P., 2008a. Explaining the level of credit spreads: Option-implied jump risk premia in a firm value model. Review of Financial Studies. 21, 2209-2242.

Cremers, M., Driessen, J., Maenhout, P., Weinbaum, D., 2008b. Individual stock-option prices and credit spreads. Journal of Banking and Finance, 32, 27062715.

Crosbie, P. J., and Bohn, J. R., 2001. Modeling Default Risk, KMV LLC 2001.

Da, Z., and Gao, P., 2010. Clientele change, liquidity shock, and the return on financially distressed stocks.

Das, S. R., Hanouna, P., Sarin, A., 2009, Accounting-based versus market-based cross-sectional models of CDS spreads, Journal of Banking & Finance, 33, 3784266.

Das, S. R., Hanouna, P., 2009, Hedging credit: Equity liquidity matters, Journal of Financial Intermediation, 18, 1042-9573.

Dennis, P., Mayhew S., 2002. Risk-neutral skewness: Evidence from index options. Journal of Financial and Quantitative Analysis 37, 471 493.

Duan, J.C., Wei, J., 2009. Systematic risk and the price structure of individual equity options. Review of Financial Studies 22,19812006.

Ericsson, J., Jacobs, K., and Oviedo, R., 2009. The determinants of credit default swap premia. Journal of Financial and Quantitative Analysis, 44(01), 109-132.

Fama, Eugene F., 1998. Market efficiency, long-term returns, and behavioral finance, Journal of Financial Economics 49, 283306.

Fama, Eugene F., and Kenneth R. French, 1993. Common risk factors in the returns on stocks and bonds, Journal of Financial Economics 33, 356.

Fama, Eugene F., and Kenneth R. French, 1996. Multifactor explanations of asset pricing anomalies, Journal of Finance 51, 5584.

Fama, Eugene F., and Kenneth R. French, 2008. Dissecting anomalies, Journal of Finance 63, 16531678.

Fama, Eugene F., and James MacBeth, 1973. Risk, return, and equilibrium: Em¬pirical tests, Journal of Political Economy 71, 607636.

Faulkender, M., and Petersen, M. A., 2006. Does the source of capital affect capital structure?. Review of financial studies, 19(1), 45-79.

Garlappi, L., and Yan, H., 2011. Financial Distress and the Cross-section of Equity Returns. The Journal of Finance, 66(3), 789-822.

Han B., Zhou Y., 2011. Term Structure of Credit Default Swap Spreads and Cross Section of Stock Returns, Working Paper.

Han, B., Subrahmanyam A., and Zhou Y., The Term Structure of Credit Spreads and the Cross-Section of Stock Returns (WP: February 5, 2015). Available at SSRN: http://ssrn.com/abstract=2560693

Hanke, M., 2005. Pricing options on leveraged equity with default risk and exponentially increasing, finite maturity debt, Journal of Economic Dynamics and Control, 29, 3, 389-421.

Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 12631295.

Johnson, H., Stulz, R., 1987. The Pricing of Options with Default Risk, Journal of Finance, 42, 2, 267-280.

Jones, C. S., 2006. A nonlinear factor analysis of S&P 500 index option returns. The Journal of Finance, 61(5), 2325-2363.

Leary, M. T., 2009. Bank loan supply, lender choice, and corporate capital structure. The Journal of Finance, 64(3), 1143-1185.

Lemmon, M., and M. R. Roberts. The response of corporate financing and investment to changes in the supply of credit." 2010: 555-587.

Longstaff, Francis A., Sanjay Mithal, and Eric Neis, 2005. Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market, Journal of Finance 60, 22132253.

Petersen, M. A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. Review of financial studies, 22(1), 435-480.

Shumway, Tyler, 2001. Forecasting bankruptcy more accurately: A simple hazard model, Journal of Business 74, 101124.

Sufi, A., 2009. Bank lines of credit in corporate finance: An empirical analysis. Review of Financial Studies, 22(3), 1057-1088.

Tang, D.Y., Yan, H., 2010. Market conditions, default risk and credit spreads, Journal of Banking & Finance 34 (2010) 743–753.

Thompson, S. B. (2006 received; 2011 published). Simple formulas for standard errors that cluster by both firm and time. Journal of Financial Economics, 99(1), 1-10.

Vassalou, M., 2003. News related to future GDP growth as a risk factor in equity returns. Journal of financial economics, 68(1), 47-73.

Vassalou, Maria, and Yuhang Xing, 2004. Default risk in equity returns, Journal of Finance 59, 831868.

White, L. J., 2010. Markets: The credit rating agencies. The Journal of Economic Perspectives, 24(2), 211-226.

Zhang, B.Y., Zhou, H. , and Zhu, H., 2009. Explaining credit default swap spreads with the equity volatility and jump risks of individual firms, Review of Financial Studies, 22, 12.

3.11 Essay II Figures and Tables

Figure 1. Default risk in Options Market - Visualization

- Default/distress risk impacts **debt valuation**
  - Merton model (1974): pioneering structural model Distance to Default (DD) that measures distress risk on corporate bankruptcy:

    High default risk = closer Distance to Default

- Default risk also impacts **equity valuation**
  - Vassalou-Xing (2004): existence of default risk premium in stocks valuation, shown in cross section of equity returns.

    High default risk = larger equity risk premium

- If default risk impacts firm value, the prices of corresponding (exchange traded) equity options too must be subject to distress shocks/default risk
  - Question: Will default risk create a **premium** in pricing of equity options?

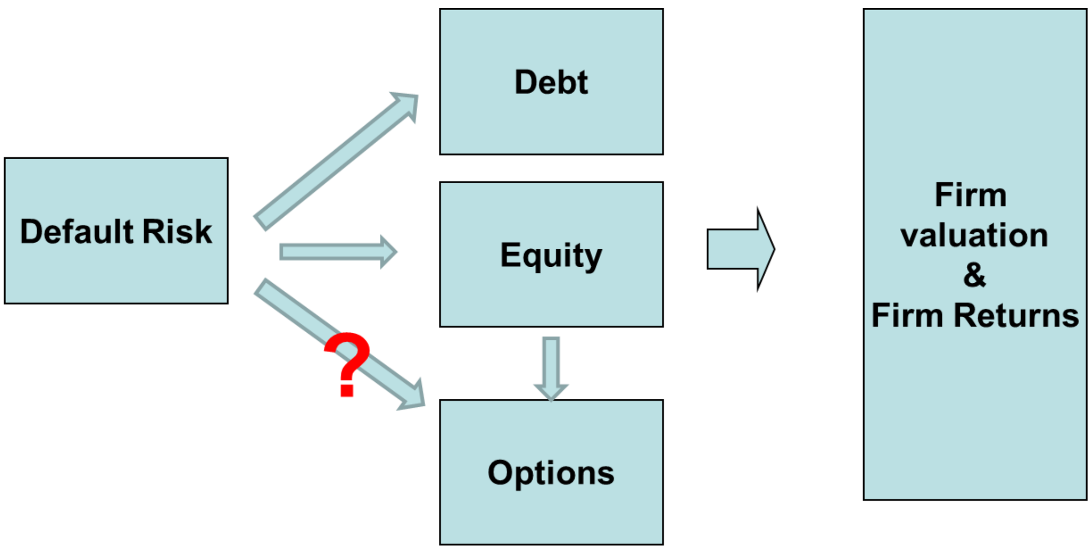
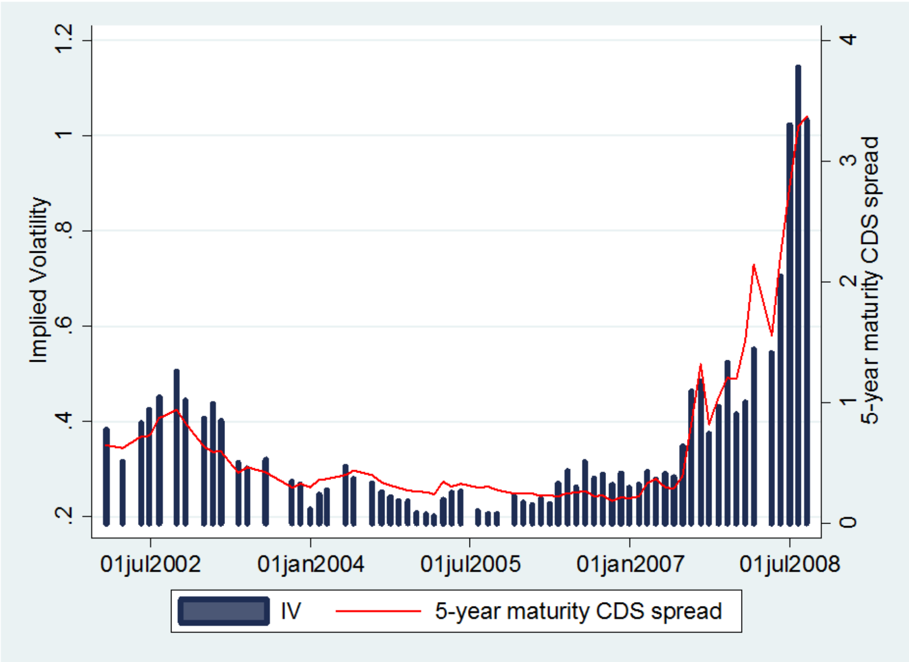


Figure 2. Default risk and implied volatility - Comovement

The graphs below shows the comovement of IV for put options and 5-year CDS spread for the same underlying firm. Two firms are shown (Lehman Brothers: **LEH,** Merrill Lynch: **MER**) during the period Jan 2002 to Jan 2009, for most liquid option bins.

2a) Lehman Brothers **(LEH)**



2b) Merrill Lynch **(MER)**

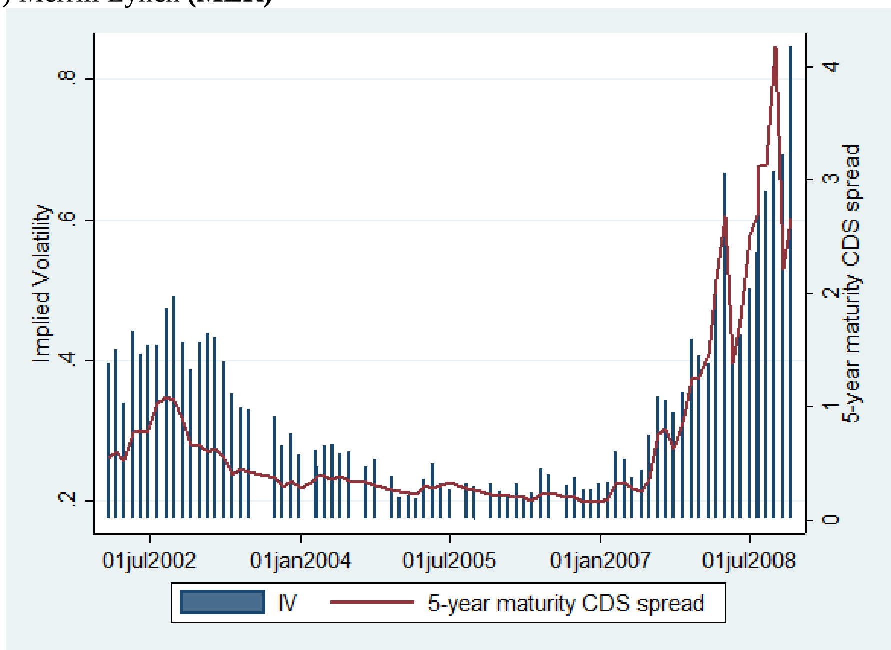
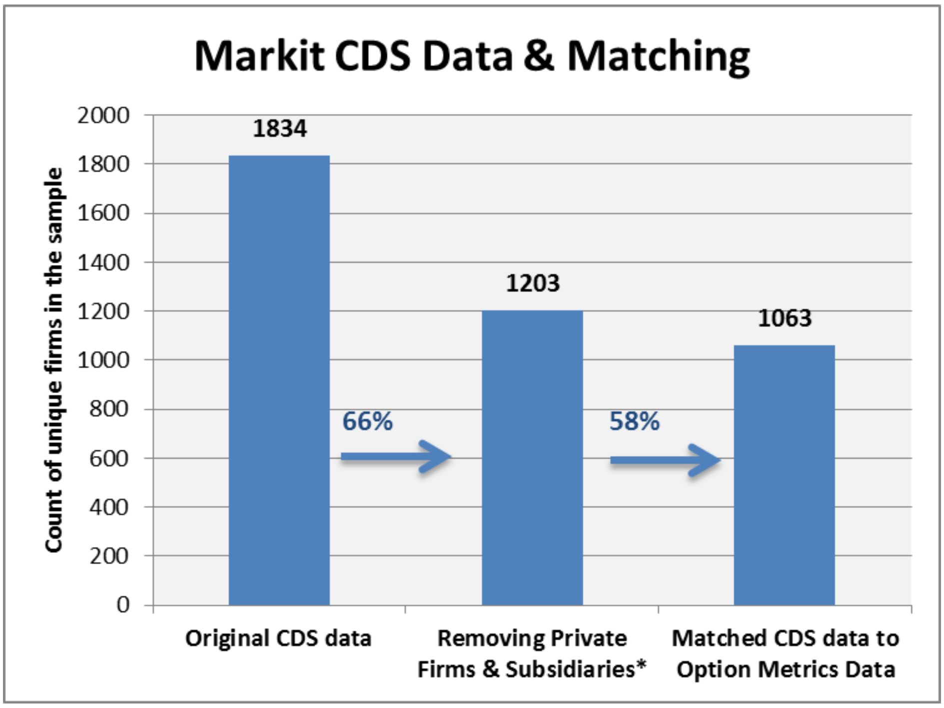


Figure 3. Data Visualization of count of firms at each clean-up and merging stage.

There are total of 1834 firms with CDS from the Mark-it database. After removing private firms and subsidiaries* there will be 1203 unique firms in the sample. We then match these firms with the Options data derived from Option Metrics database. As a result, there are 1063 unique firms remaining. We also remove all Non-USD based firms by retrieving and matching equity data from CRSP. The base sample before applying test-based filters has 1063 unique firms.



**Refer to Appendix A for details on scrubbing subsidiaries.*

Table 1. Data description and Summary Statistics

Panel A displays the basic statistics for the full sample on the observed variables to be tested. Panel B shows the unique count. In the final filtered sample we have 550 unique firms and 67015 unique Options written on these firms. Data ranges from Jan2002 to Dec 2009.

Panel A. Observed variables:

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**	**25th Percentile**	**Median**	**75th Percentile**
CDS Spread	75361	1.11	2.14	0.01	74.81	0.25	0.49	1.10
# of Contributors	75361	9	6	2	29	4	9	13
CDS Return	29292	0.03	0.48	-10.63	14.73	-0.04	0.00	0.24
Strike price	108120	56.09	61.89	2.50	840.00	30.00	45.00	65.00
Volume	108120	167	728	1	63151	10	26	93
Maturity	108120	92	45	44	177	50	80	138
Price	108120	57.59	62.55	2.09	707	29.17	46.12	66.37
Moneyness	108120	0.98	0.10	0.80	1.20	0.90	0.97	1.05
Implied Volatility	107201	0.38	0.18	0.01	3.79	0.26	0.33	0.43
VIX	108120	19.28	8.92	10.42	59.89	12.96	17.00	23.01
Market to Book Value	69297	1.99	1.06	0.39	9.19	1.21	1.64	2.46
Leverage	69238	0.27	0.17	0.00	1.34	0.15	0.24	0.35
Size	69238	9.79	1.48	3.26	14.67	8.79	9.60	10.53

Panel B. Identifying variables:

Time series Variables	**Number of unique values**	**Range**
Years	8	2002-2009
Months	96	Jan2002-Dec2009
Cross-sectional Variables	**Number of unique values**	
Firms (permno)	550	
Options (OptionID)	67015	

Table 1. Data description and Summary Statistics (Cont.)
Panel C displays the basic statistics for the full sample on the observed variables to be tested on an annual basis. In the final filtered sample we have 550 unique firms and 67015 unique Options written on these firms. Data ranges from Jan2002 to Dec 2009.
Panel D shows the distribution of average characteristics across the 9 bins.

Panel C. Annual Data:

Year	**No of Obs.**	**No. of firms with both Option and CDS**	**CDS Spread Mean**	**Implied Volatility Mean**
2002	17451	428	1.52	0.47
2003	20493	427	1.11	0.37
2004	25397	449	0.97	0.32
2005	29155	459	0.95	0.30
2006	35666	461	0.69	0.31
2007	41387	459	0.85	0.33
2008	36943	444	1.85	0.51
2009	37691	453	1.93	0.62

Paned D. Bin distributions
The table below shows the distribution of each key variable across 9 bins with different maturity and moneyness values.

	Maturity/ Moneyness	**Short** <=70d	**Medium** 70< & <120	**Long** >=120
Avg IV	**ITM** >=1.05	0.4027	0.3596	0.3318
Avg IV skew		0.0423	0.0425	0.0416
Avg CDS spread		1.2874	1.2762	1.2031
Avg CDS contributors		9.0469	9.2802	9.6075
Avg IV	**ATM** 0.95< & <1.05	0.3388	0.3345	0.3218
Avg IV skew		0.0392	0.0391	0.0390
Avg CDS spread		0.9764	0.9933	0.9642
Avg CDS contributors		8.9628	9.2768	9.7416
Avg IV	**OTM** =<0.95	0.4394	0.3842	0.3504
Avg IV skew		0.0500	0.0426	0.0371
Avg CDS spread		1.2066	1.1013	1.0435
Avg CDS contributors		9.0044	9.2431	9.6926

Table 2a. Complete Correlation Analysis

Table below displays all the pairwise correlation coefficients between the variables to be included in the regression specifications. Correlations that are both economically and statistically significant are bolded in the table below. Special attention is noted to these variables when running multivariate regressions so that the issues of multi-colinearity is addressed. The dependent variable is *Implied Volatility* to be tested in the following regression tables.

	Implied Volatility	CDS Spread	Lagged CDS spread	Lagged Implied volatility	Leverage	M/B	Size	Firm Return	Historical Volatility	CDS Liquidity proxy	VIX	S&P500 Return	Yield Curve Slope	Treasury Rate	Agg. Default Spread	TED Spread
Implied Volatility	1.00															
CDS Spread	**0.61**	1.00														
Lagged CDS spread	**0.58**	**0.97**	1.00													
Lagged Implied volatility	**0.93**	**0.60**	0.61	1.00												
Leverage	0.12	0.31	0.29	0.11	1.00											
Market-to-Book ratio	-0.16	-0.23	-0.24	-0.16	-0.25	1.00										
Size	-0.15	-0.14	-0.18	-0.23	0.02	-0.17	1.00									
Firm Return	-0.10	-0.05	-0.01	-0.09	-0.01	0.00	-0.01	1.00								
Historical Volatility	0.33	0.15	0.16	0.33	0.00	-0.06	-0.14	-0.17	1.00							
CDS Liquidity proxy	-0.15	-0.13	-0.14	-0.17	0.04	-0.26	0.28	0.01	-0.04	1.00						
VIX	**0.54**	0.19	0.14	0.43	0.03	-0.03	0.09	-0.19	0.09	-0.17	1.00					
S&P500 Return	-0.34	-0.11	-0.04	-0.15	-0.02	0.02	-0.04	0.31	-0.05	0.07	**-0.65**	1.00				
Yield Curve Slope	0.20	0.09	0.09	0.24	0.03	0.04	0.07	0.00	0.06	-0.18	0.35	-0.06	1.00			
Treasury Rate	-0.30	-0.14	-0.13	-0.31	-0.04	-0.01	-0.07	0.03	-0.07	0.13	-0.48	0.17	**-0.90**	1.00		
Agg. Default Spread	0.46	0.19	0.15	0.46	0.03	-0.04	0.08	-0.09	0.09	-0.15	**0.73**	-0.30	0.23	-0.45	1.00	
TED Spread	0.34	0.12	0.10	0.27	0.02	-0.04	0.06	-0.09	0.05	-0.02	**0.58**	-0.34	-0.17	-0.06	**0.51**	1.00

Table 2b. Complete Univariate Regression Analysis

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
CDS spread	0.0307***															
	(74.72)															
CDSLiqproxy		0.00940***														
		(22.48)														
Lagged IV			0.354***													
			(75.40)													
Lagged CDS				0.0364***												
				(117.0)												
Leverage					0.121***											
					(32.63)											
M/B Value Ratio						-0.0152***										
						(-26.39)										
Size							-0.00709***									
							(-14.42)									
Firm Return								-0.0581***								
								(-27.94)								
HV									0.210***							
									(31.36)							
Size										0.00821***						
										(283.8)						
S&P 500 Return											-0.625***					
											(-132.5)					
YieldCurveSlope												0.0370***				
												(81.00)				
TreasuryRate													-0.0290***			
													(-117.8)			
AggDefSpread														0.114***		
														(151.0)		
TEDSpread															0.0565***	
															(118.0)	
OptionLiq																0.0216***
																(26.36)
Constant	0.317***	0.347***	0.232***	0.318***	0.334***	0.396***	0.435***	0.374***	0.347***	0.226***	0.387***	0.349***	0.469***	0.265***	0.359***	0.387***
	(675.1)	(1,704)	(142.4)	(411.4)	(271.1)	(300.6)	(91.09)	(508.7)	(322.1)	(292.6)	(674.7)	(439.9)	(529.9)	(267.0)	(566.7)	(614.7)
Observations	140,647	140,647	105,602	76,258	129,056	129,169	129,056	105,868	58,456	201,571	201,572	201,572	201,572	201,572	201,572	201,572
R-squared	0.069	0.007	0.089													
Number of optionid	65,133	65,133	47,312	33,581	59,583	59,661	59,583	47,452	20,956	95,845	95,845	95,845	95,845	95,845	95,845	95,845

t-statistics in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3. Regression on Option Implied Volatility for “Short Maturity” and “In The Money”
We run the pooled regression on dependent variable *Option-Implied Volatility*.

	Model 1	Model 2	Model 3	Model 4
Default Risk Variables				
CDS spread		0.0520***	0.0675***	0.0521***
		(2.637)	(3.211)	(2.641)
CDS Liquidity Proxy		-0.0231*	-0.0244**	-0.0222*
		(-1.913)	(-2.023)	(-1.832)
Lagged CDS			0.0351**	
			(2.589)	
Lagged Implied Volatility				-0.0928
				(-1.011)
Firm-level Variables				
Total Debt to Total Assets Ratio	0.312	0.341	0.347	0.315
	(1.240)	(1.359)	(1.396)	(1.248)
Market-to-Book Value Ratio	-0.122***	-0.119***	-0.116***	-0.120***
	(-3.915)	(-3.831)	(-3.789)	(-3.853)
Size as Total Assets	-0.143*	-0.135	-0.137	-0.138
	(-1.654)	(-1.564)	(-1.605)	(-1.605)
FirmRet	0.0360	0.0342	0.0114	0.0410
	(0.864)	(0.828)	(0.271)	(0.979)
HV	0.0869	0.105	0.111	0.111
	(0.653)	(0.789)	(0.839)	(0.835)
Aggregate Variables				
VIX	0.0125***	0.0120***	0.0114***	0.0122***
	(8.195)	(7.787)	(7.344)	(7.838)
Return on the S&P 500 Index	0.146	0.168	0.0903	0.200
	(0.932)	(1.080)	(0.578)	(1.262)
Yield Curve Slope	-0.00547	-0.0109	-0.0131	-0.0104
	(-0.225)	(-0.450)	(-0.547)	(-0.428)
Treasury Rate	0.0360*	0.0375*	0.0464**	0.0382*
	(1.707)	(1.784)	(2.196)	(1.816)
Agg Default Spread	0.0239	0.0194	0.0139	0.0253
	(0.773)	(0.630)	(0.455)	(0.806)
TED Spread	0.0208	0.0200	0.0211*	0.0201
	(1.619)	(1.561)	(1.652)	(1.568)
Intercept	1.559*	1.418	1.371	1.480*
	(1.754)	(1.605)	(1.561)	(1.672)
No. of Observations	2,835	2,835	2,773	2,788
R-squared	0.358	0.372	0.386	0.374
Fixed Effects	Yes	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 4. Regression on Option Implied Volatility for "Short Maturity" and "Out of the Money". We run the pooled regression on dependent variable *Option-Implied Volatility*.

	Model 1	Model 2	Model 3	Model 4
Default Risk Variables				
CDS spread		0.00399	0.00394	0.00230
		(0.403)	(0.394)	(0.231)
CDS Liquidity Proxy		0.000517	0.000186	-0.000327
		(0.151)	(0.0524)	(-0.0948)
Lagged CDS			0.00296	
			(0.368)	
Lagged Implied Volatility				-0.0656*
				(-1.828)
Firm-level Variables				
Total Debt to Total Assets Ratio	-0.0610	-0.0620	-0.0724	-0.0484
	(-0.721)	(-0.730)	(-0.847)	(-0.569)
Market-to-Book Value Ratio	0.0312**	0.0318**	0.0300**	0.0315**
	(2.466)	(2.498)	(2.341)	(2.480)
Size as Total Assets	0.0804***	0.0817***	0.0800***	0.0811***
	(2.849)	(2.877)	(2.802)	(2.860)
FirmRet	0.0101	0.00991	0.00921	0.00182
	(0.613)	(0.603)	(0.559)	(0.107)
HV	-0.0333	-0.0332	-0.0389	-0.0406
	(-0.605)	(-0.602)	(-0.701)	(-0.736)
Aggregate Variables				
VIX	0.00932***	0.00930***	0.00929***	0.00930***
	(13.54)	(13.45)	(13.33)	(13.45)
Return on the S&P 500 Index	0.107	0.109	0.103	0.119*
	(1.579)	(1.602)	(1.507)	(1.744)
Yield Curve Slope	-0.0209*	-0.0211*	-0.0210*	-0.0234*
	(-1.725)	(-1.742)	(-1.729)	(-1.926)
Treasury Rate	0.0256***	0.0258***	0.0263***	0.0230**
	(2.609)	(2.617)	(2.644)	(2.307)
Agg Default Spread	0.0313**	0.0307*	0.0294*	0.0401**
	(2.005)	(1.960)	(1.832)	(2.434)
TED Spread	0.0239***	0.0241***	0.0242***	0.0231***
	(3.752)	(3.770)	(3.772)	(3.599)
Intercept	-0.736**	-0.752***	-0.734**	-0.723**
	(-2.572)	(-2.608)	(-2.519)	(-2.505)
No. of Observations	5,875	5,875	5,792	5,875
R-squared	0.257	0.257	0.258	0.259
Fixed Effects	Yes	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 5. Regression on Option Implied Volatility for “Short Maturity”. We run the pooled regression on dependent variable *Option-Implied Volatility* on all moneyness groups.

	ITM: Moneyness >=1.05	ATM: 0.95< Moneyness <1.05	OTM: Moneyness >0.95
Default Risk Variables			
CDS spread	0.0521***	0.00683	0.00230
	(2.641)	(0.669)	(0.231)
CDS Liquidity Proxy	-0.0222*	-0.00201	-0.000327
	(-1.832)	(-0.415)	(-0.0948)
Lagged Implied Volatility	-0.0928	-0.235***	-0.0656*
	(-1.011)	(-6.948)	(-1.828)
Firm-level Variables			
Total Debt to Total Assets Ratio	0.315	0.135	-0.0484
	(1.248)	(1.629)	(-0.569)
Market-to-Book Value Ratio	-0.120***	-0.00932	0.0315**
	(-3.853)	(-1.002)	(2.480)
Size as Total Assets	-0.138	-0.0409**	0.0811***
	(-1.605)	(-2.032)	(2.860)
FirmRet	0.0410	-0.00554	0.00182
	(0.979)	(-0.392)	(0.107)
HV	0.111	0.173***	-0.0406
	(0.835)	(4.150)	(-0.736)
Aggregate Variables			
VIX	0.0122***	0.00796***	0.00930***
	(7.838)	(16.59)	(13.45)
Return on the S&P 500 Index	0.200	0.193***	0.119*
	(1.262)	(4.244)	(1.744)
Yield Curve Slope	-0.0104	0.0321***	-0.0234*
	(-0.428)	(3.752)	(-1.926)
Treasury Rate	0.0382*	0.0227***	0.0230**
	(1.816)	(3.294)	(2.307)
Agg Default Spread	0.0253	-0.0142	0.0401**
	(0.806)	(-1.322)	(2.434)
TED Spread	0.0201	0.00821*	0.0231***
	(1.568)	(1.896)	(3.599)
Intercept	1.480*	0.518**	-0.723**
	(1.672)	(2.514)	(-2.505)
No. of Observations	2,788	5,659	5,875
R-squared	0.374	0.408	0.259
Fixed Effects	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6. Regression on Option Implied Volatility for "Medium Maturity". We run the pooled regression on dependent variable *Option-Implied Volatility* on all moneyness groups.

	ITM: Moneyness >=1.05	ATM: 0.95< Moneyness <1.05	OTM: Moneyness >0.95
Default Risk Variables			
CDS spread	0.00201	0.0233***	0.0127***
	(0.716)	(6.897)	(6.413)
CDS Liquidity Proxy	0.00300	0.00174	0.00252*
	(0.844)	(0.936)	(1.881)
Lagged Implied Volatility	0.0376	0.0796***	0.0652***
	(0.911)	(3.450)	(3.496)
Firm-level Variables			
Total Debt to Total Assets Ratio	0.122*	-0.00333	0.00224
	(1.914)	(-0.109)	(0.0975)
Market-to-Book Value Ratio	0.00326	0.0125***	0.0240***
	(0.415)	(2.756)	(5.716)
Size as Total Assets	-0.0419	0.00252	0.0244***
	(-1.047)	(0.148)	(2.671)
FirmRet	0.000131	-0.00841	0.0143**
	(0.00930)	(-0.980)	(2.430)
HV	0.0811**	0.00180	0.0159
	(2.221)	(0.105)	(1.067)
Aggregate Variables			
VIX	0.00445***	0.00485***	0.00540***
	(11.76)	(22.45)	(27.71)
Return on the S&P 500 Index	-0.0763*	0.0146	0.0223
	(-1.696)	(0.616)	(1.004)
Yield Curve Slope	0.0158	0.00376	-0.0160***
	(1.600)	(0.790)	(-3.784)
Treasury Rate	-0.00330	-0.00682**	-0.0159***
	(-0.538)	(-2.220)	(-5.535)
Agg Default Spread	0.00481	-0.0119*	-0.0128**
	(0.493)	(-1.743)	(-2.238)
TED Spread	0.00878*	0.00575**	0.00278
	(1.954)	(2.337)	(1.274)
Intercept	0.598	0.143	-0.0206
	(1.450)	(0.797)	(-0.210)
No. of Observations	3,251	5,699	7,267
R-squared	0.455	0.451	0.464
Fixed Effects	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 7. Regression on Option Implied Volatility for "Long Maturity". We run the pooled regression on dependent variable *Option-Implied Volatility* on all moneyness groups.

	ITM: Moneyness >=1.05	ATM: 0.95< Moneyness <1.05	OTM: Moneyness >0.95
Default Risk Variables			
CDS spread	-0.00692***	0.00956***	0.0154***
	(-4.201)	(3.954)	(13.97)
CDS Liquidity Proxy	0.00295*	-0.00120	5.51e-05
	(1.820)	(-0.753)	(0.0728)
Lagged Implied Volatility	0.277***	0.449***	0.285***
	(11.00)	(23.69)	(21.87)
Firm-level Variables			
Total Debt to Total Assets Ratio	0.00352	0.00828	0.000763
	(0.163)	(0.602)	(0.0567)
Market-to-Book Value Ratio	0.0150***	0.00982***	0.0136***
	(4.617)	(3.930)	(6.047)
Size as Total Assets	0.00517	-0.00315	0.00106
	(0.451)	(-0.533)	(0.202)
FirmRet	-0.00321	-0.00854	0.00954*
	(-0.351)	(-1.351)	(1.950)
HV	0.0316*	0.0457***	0.0500***
	(1.699)	(3.959)	(5.064)
Aggregate Variables			
VIX	0.00358***	0.00303***	0.00368***
	(14.50)	(16.71)	(22.18)
Return on the S&P 500 Index	-0.0589**	-0.0487**	-0.0754***
	(-2.275)	(-2.543)	(-4.538)
Yield Curve Slope	0.00913**	0.00560**	0.00365*
	(2.491)	(2.193)	(1.679)
Treasury Rate	-0.00339*	2.76e-05	-8.04e-05
	(-1.679)	(0.0199)	(-0.0649)
Agg Default Spread	0.00819*	-0.00404	0.00173
	(1.854)	(-1.174)	(0.562)
TED Spread	-0.00925***	-0.00183	0.00661***
	(-3.542)	(-0.976)	(3.961)
Intercept	0.0587	0.100	0.0883
	(0.468)	(1.510)	(1.520)
No. of Observations	2,160	2,976	4,395
R-squared	0.525	0.631	0.584
Fixed Effects	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 8. Regression on IV Skew for "Short Maturity". We run the pooled regression on dependent variable *Option-Implied Volatility* on all moneyness groups. We define IV-skew as the difference between average OTM and ATM option implied volatilities for each issuer.

	ITM: Moneyness >=1.05	ATM: 0.95< Moneyness <1.05	OTM: Moneyness >0.95
Default Risk Variables			
CDS spread	0.000622	-0.00644	0.0126**
	(0.0669)	(-1.201)	(2.110)
CDS Liquidity Proxy	0.0133**	0.0117***	0.00803***
	(2.348)	(3.063)	(3.925)
Lagged Implied Volatility	-0.113**	0.183***	-0.137***
	(-2.014)	(6.890)	(-6.325)
Firm-level Variables			
Total Debt to Total Assets Ratio	0.315	0.135	-0.0484
	(1.248)	(1.629)	(-0.569)
Market-to-Book Value Ratio	-0.120***	-0.00932	0.0315**
	(-3.853)	(-1.002)	(2.480)
Size as Total Assets	-0.138	-0.0409**	0.0811***
	(-1.605)	(-2.032)	(2.860)
FirmRet	0.0410	-0.00554	0.00182
	(0.979)	(-0.392)	(0.107)
HV	0.111	0.173***	-0.0406
	(0.835)	(4.150)	(-0.736)
Aggregate Variables			
VIX	0.00103	-0.000729*	0.00230***
	(1.064)	(-1.931)	(5.407)
Return on the S&P 500 Index	0.0903	0.0621*	0.0246
	(0.941)	(1.730)	(0.591)
Yield Curve Slope	0.0116	0.00703	-0.0234***
	(0.767)	(1.044)	(-3.109)
Treasury Rate	0.0261**	-0.000408	0.00198
	(2.044)	(-0.0749)	(0.321)
Agg Default Spread	0.0120	-0.0313***	0.0173*
	(0.600)	(-3.673)	(1.714)
TED Spread	0.0162**	-0.000136	-0.00103
	(2.032)	(-0.0398)	(-0.263)
Intercept	0.0310	0.150	0.144
	(0.0568)	(0.925)	(0.972)
No. of Observations	3,212	5,688	6,115
R-squared	0.074	0.105	0.118
Fixed Effects	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 9. Regression on IV Skew for "Medium Maturity". We run the pooled regression on dependent variable *Option-Implied Volatility* on all moneyness groups. We define IV-skew as the difference between average OTM and ATM option implied volatilities for each issuer.

	ITM: Moneyness >=1.05	ATM: 0.95< Moneyness <1.05	OTM: Moneyness >0.95
Default Risk Variables			
CDS spread	-0.00830**	-0.00631*	0.00863***
	(-2.017)	(-1.791)	(4.308)
CDS Liquidity Proxy	0.00542	0.0105***	-0.000125
	(1.042)	(5.420)	(-0.0889)
Lagged Implied Volatility	-0.139**	0.149***	-0.105***
	(-2.334)	(6.177)	(-5.399)
Firm-level Variables			
Total Debt to Total Assets Ratio	0.0110	-0.0161	-0.0273
	(0.119)	(-0.505)	(-1.133)
Market-to-Book Value Ratio	0.00721	-0.00402	0.00487
	(0.637)	(-0.845)	(1.105)
Size as Total Assets	-0.00318	-0.0375**	0.00433
	(-0.0553)	(-2.110)	(0.452)
FirmRet	0.00657	0.0495***	-0.00824
	(0.320)	(5.528)	(-1.335)
HV	-0.0164	-0.00580	-0.0161
	(-0.308)	(-0.324)	(-1.029)
Aggregate Variables			
VIX	-0.000231	0.000175	0.00165***
	(-0.419)	(0.777)	(8.097)
Return on the S&P 500 Index	-0.0523	0.110***	-0.0267
	(-0.798)	(4.442)	(-1.144)
Yield Curve Slope	-0.00213	0.00222	-0.0146***
	(-0.149)	(0.446)	(-3.291)
Treasury Rate	0.0125	0.00397	-0.00910***
	(1.404)	(1.240)	(-3.016)
Agg Default Spread	0.0374***	0.0104	0.00312
	(2.628)	(1.454)	(0.518)
TED Spread	0.0196***	0.00137	0.000766
	(3.000)	(0.534)	(0.335)
Intercept	0.0275	0.366*	0.0269
	(0.0463)	(1.959)	(0.262)
No. of Observations	3,302	5,700	7,270
R-squared	0.036	0.111	0.103
Fixed Effects	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 10. Regression on IV-Skew for "Long Maturity". We run the pooled regression on dependent variable *Option-Implied Volatility* on all moneyness groups. We define IV-skew as the difference between average OTM and ATM option implied volatilities for each issuer.

	ITM: Moneyness >=1.05	ATM: 0.95< Moneyness <1.05	OTM: Moneyness >0.95
Default Risk Variables			
CDS spread	-0.00165	0.0149***	-0.00979***
	(-0.708)	(5.468)	(-7.625)
CDS Liquidity Proxy	-0.00245	0.00286	0.00474***
	(-1.067)	(1.594)	(5.376)
Lagged Implied Volatility	-0.0611*	0.0837***	-0.145***
	(-1.711)	(3.909)	(-9.558)
Firm-level Variables			
Total Debt to Total Assets Ratio	-0.0517*	-0.0230	0.00814
	(-1.697)	(-1.479)	(0.518)
Market-to-Book Value Ratio	0.00379	-0.00481*	0.00307
	(0.823)	(-1.705)	(1.166)
Size as Total Assets	0.0290*	-0.00341	-0.00107
	(1.821)	(-0.511)	(-0.174)
FirmRet	0.0186	0.0243***	0.0161***
	(1.434)	(3.411)	(2.824)
HV	-0.0211	0.00475	0.00732
	(-0.803)	(0.364)	(0.635)
Aggregate Variables			
VIX	0.000925***	0.000669***	0.00122***
	(2.641)	(3.267)	(6.310)
Return on the S&P 500 Index	-0.0197	0.0595***	-0.0508***
	(-0.538)	(2.755)	(-2.621)
Yield Curve Slope	-0.000144	-0.00207	-0.00638**
	(-0.0278)	(-0.718)	(-2.513)
Treasury Rate	0.00170	0.000696	-0.00482***
	(0.601)	(0.444)	(-3.334)
Agg Default Spread	0.0266***	-0.00229	0.00617*
	(4.257)	(-0.590)	(1.713)
TED Spread	0.0113***	0.00478**	0.00369*
	(3.052)	(2.256)	(1.897)
Intercept	-0.296*	0.0453	0.0785
	(-1.700)	(0.604)	(1.157)
No. of Observations	2,172	2,975	4,395
R-squared	0.066	0.091	0.128
Fixed Effects	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 11. Economic Significance Comparison
What is the economic significance of CDS change on IV? We can answer the question by measuring the impact of one-sigma shock of CDS on option-implied volatility.

We calculate and show the economic significance of the key regression models in the following table. The results are based on the pooled regression of "Short Maturity" and "In The Money" bins. The *reference regression model* for each of the following economic significance tests displayed in panels A and B, are provided in Tables 3 and 8, respectively.

Panel A records the economic significance for the *Option-Implied Volatility.* The calculated economic significance of one sigma shock in CDS on IV, is bolded and shows an average of 3% deviation for implied volatility for one level of CDS sigma shock. Panel B displays the economic significance for *Implied Volatility Skew*.

Panel A: **Dependent Variable: Implied Volatility**	Model 2	Model 3	Model 4
Default Risk Variables			
CDS spread	**2.559%**	**3.322%**	**2.564%**
CDS Liquidity Proxy	-2.380%	-2.514%	-2.287%
Lagged CDS		1.65%	
Lagged IV			N/A

Firm level and Market level variables are controlled for;
All coefficients tested above are significant. Detailed significance levels recorded in Table 3.

	Short Maturity		
Panel B: **Dependent Variable: IV Skew**	ITM	ATM	OTM
Default Risk Variables			
CDS spread	N/A	N/A	0.573%
CDS Liquidity Proxy			0.886%

Firm level and Market level variables are controlled for;
All coefficients tested above are significant. Detailed significance levels recorded in Table 8.

Table 12. Endogeneity Tests for Option-implied volatility and CDS spread

Panel A. Two-Stage Lease Squares Estimation
To control for the potential endogeneity problems arising from the contemporaneous measurement of the CDS spread, and the implied volatility, we employ a simultaneous equation model using two equations that represent each of the potentially endogenous variables. The system of equations is as follows:

$$CDS_{it} = \beta_0 + \beta_1 L\ aggedCDS_{it}, +\beta_2\ FirmLevelVars_i, +\beta_3\ MarketLevelVars_{it} + \varepsilon_{it}$$

$$IV_{it} = \beta_0 + \beta_1\ CDS_{it} + \beta_2\ CDSLiquidityProxy_i + \beta_3\ LaggedIV_{it}, + \beta_4 FirmLevelVars_{it}, + \beta_5 MarketLevelVars_{it} + \varepsilon_{it}$$

Panel A records the first stage regression results of CDS. The results show that the lagged CDS is in fact a suitable instrument for CDS spread. The adjusted R^2 is 93.53%.

	Coefficient	**t-statistic**
Lagged CDS	0.9830***	*539.21*
CDS Liquidity Proxy	0.0821***	*23.94*
Lagged Implied Volatility	-0.0901***	*-3*
Total Debt to Total Assets Ratio	0.1046***	*6.84*
Market-to-Book Value Ratio	-0.0026	*-1.08*
Size as Total Assets	-0.0040*	*-2.25*
HV	0.0149	*0.31*
VIX	0.0067***	*9.09*
Return on the S&P 500 Index	-1.5057***	*-15.57*
Yield Curve Slope	-0.0435***	*-4.81*
Treasury Rate	-0.0040	*-0.83*
Agg Default Spread	-0.0102	*-0.85*
TED Spread	-0.0105	*-1.69*
Intercept	0.0097	*0.29*
No. of Observations	39436	
Adjusted R-squared	0.9353	

The t-statistics are recorded in italics; *** p<0.01, ** p<0.05, * p<0.1

Table 12. (Cont.)

Panel B. Instrumented Variable (Second Stage) regression

	All bins	ITM bins	ATM bins	OTM bins
Default Risk Variables				
CDS spread	0.00216***	0.00622***	0.00250***	0.000629**
	(10.98)	(14.77)	(8.276)	(2.031)
CDS Liquidity Proxy	0.00546***	0.00559***	0.00498***	0.00712***
	(14.95)	(7.157)	(9.065)	(12.41)
Lagged Implied Volatility	0.857***	0.785***	0.854***	0.857***
	(270.1)	(101.3)	(190.5)	(173.6)
Firm-level Variables				
Total Debt to Total Assets Ratio	-0.00131	-0.00815**	-0.00568**	0.00342
	(-0.808)	(-1.982)	(-2.548)	(1.384)
Market-to-Book Value Ratio	-0.00261***	-0.00355***	-0.00147***	-0.00371***
	(-10.38)	(-5.517)	(-4.250)	(-9.708)
Size as Total Assets	-0.00328***	-0.00416***	-0.00210***	-0.00461***
	(-17.46)	(-8.868)	(-8.099)	(-15.98)
FirmRet	0.0962***	0.0994***	0.0875***	0.101***
	(18.77)	(7.239)	(11.61)	(13.84)
HV	-0.00131	-0.00815**	-0.00568**	0.00342
	(-0.808)	(-1.982)	(-2.548)	(1.384)
Aggregate Variables				
VIX	0.00401***	0.00402***	0.00376***	0.00443***
	(51.74)	(22.67)	(34.28)	(35.99)
Return on the S&P 500 Index	-0.185***	-0.223***	-0.213***	-0.176***
	(-18.03)	(-9.026)	(-14.88)	(-11.15)
Yield Curve Slope	-0.0151***	-0.0230***	-0.0107***	-0.0156***
	(-15.75)	(-9.433)	(-8.080)	(-10.71)
Treasury Rate	-0.00365***	-0.00755***	-0.000840	-0.00394***
	(-7.073)	(-5.798)	(-1.186)	(-4.984)
Agg Default Spread	-0.0462***	-0.0450***	-0.0426***	-0.0454***
	(-36.30)	(-15.53)	(-23.85)	(-22.30)
TED Spread	-0.00425***	-0.00705***	-0.00445***	-0.00290***
	(-6.469)	(-4.507)	(-4.862)	(-2.841)
Intercept	0.0849***	0.131***	0.0505***	0.103***
	(23.78)	(14.52)	(10.29)	(18.92)
No. of Observations	39,436	8,041	14,110	17,285
R-squared	0.860	0.857	0.890	0.849
Fixed Effects	Yes	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 12. (Cont.)

Panel C. As robustness, we perform two post-regression tests of Durbin and Wu-Hausman on the 2SLS regression model below, where CDS is instrumented.

Stage 1: CDS = f_1 *(Lagged_CDS, firm_level_vars, market_level_vars)*
Stage 2: IV = f_2 *(CDS, CDSLiquidityProxy, lagged_IV, firm_level_vars, market_level_vars)*

Panel C records the findings and shows that the null hypothesis is rejected by both tests.* After controlling for endogeneity, CDS remains a significant variable in the regression model.

Panel C. Post-regression Tests of endogeneity

H0: variables are exogeneous	**All moneyness**	**ITM bins**	**ATM bins**	**OTM bins**
Durbin (score) chi2(1)	542.965	67.8885	82.0269	431.386
Wu-Hausman F(1,39421)	550.336	68.3388	82.4188	442.044
*p-value	(p = 0.00)	(p = 0.00)	(p = 0.00)	(p = 0.00)

**Based on the p-values of the two post-estimation tests, the null hypothesis of exogenous variables is rejected. Thus, IV and CDS are endogenous variables, determined by the set of common explanatory variables at firm level and market level.*

Table 13. Impact of Industry. We run the regression for two subsamples of Financial and Non-financial firms (from total of 550 firms in this study, 469 are non-financial firms and 81 are financial firms). Panel A reports IV regressions and panel B reports IV-Skew regressions.

	Panel A) Dependent: IV		**Panel B) Dep.: IV-Skew**	
	NonFinancial	Financial	NonFinancial	Financial
Default Risk Variables				
CDS spread	0.00730***	0.0498***	0.00219***	0.00167
	(10.46)	(15.11)	(3.904)	(0.576)
CDS Liquidity Proxy	0.00102*	0.00438*	0.00441***	0.00603***
	(1.954)	(1.895)	(10.58)	(2.965)
Lagged Implied Volatility	0.200***	0.146***	-0.0237***	0.0195
	(27.32)	(6.908)	(-3.912)	(1.053)
Firm-level Variables				
Total Debt to Total Assets Ratio	-0.0220**	0.0189	0.000944	-0.0351**
	(-2.070)	(0.987)	(0.106)	(-2.117)
Market-to-Book Value Ratio	0.0201***	-0.00527	-0.000277	-0.0369
	(14.75)	(-0.189)	(-0.244)	(-1.517)
Size as Total Assets	0.0587***	0.0424***	-0.00375	0.00378
	(15.08)	(3.428)	(-1.166)	(0.348)
FirmRet	0.0255***	0.0789***	0.0113***	0.0343***
	(8.863)	(8.489)	(4.667)	(4.203)
HV	0.0615***	0.0491**	-0.00903	-0.0441**
	(8.941)	(2.304)	(-1.565)	(-2.357)
Aggregate Variables				
VIX	0.00573***	0.00661***	0.000988***	0.000350
	(66.02)	(21.50)	(13.56)	(1.293)
Return on the S&P 500 Index	-0.00726	-0.124***	0.00314	0.0149
	(-0.681)	(-3.828)	(0.352)	(0.525)
Yield Curve Slope	-0.00869***	-0.00596	-0.00747***	0.00182
	(-5.624)	(-1.186)	(-5.759)	(0.414)
Treasury Rate	-0.00721***	-0.00568**	-0.00274***	-3.16e-05
	(-8.356)	(-2.063)	(-3.792)	(-0.0131)
Agg Default Spread	-0.0106***	-0.0211***	0.00538***	-0.00996
	(-6.221)	(-2.967)	(3.788)	(-1.594)
TED Spread	0.00144	0.00732**	0.00188**	-0.00278
	(1.616)	(2.497)	(2.520)	(-1.079)
Intercept	-0.436***	-0.367**	0.0714**	0.0496
	(-10.83)	(-2.273)	(2.148)	(0.350)
No. of Observations	35,444	4,626	36,154	4,675
R-squared	0.454	0.516	0.038	0.021
Fixed Effects	Yes	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 14. Impact of Financial Leverage We run the regression for two subsamples of Low Leverage and High Leverage firms (defined based on median leverage). Panel A reports IV regressions and panel B reports IV-Skew regressions.

	Panel A) Dependent: IV		**Panel B) Dep.: IV-Skew**	
	Low Lev	High Lev	Low Lev	High Lev
Default Risk Variables				
CDS spread	0.0201***	0.0105***	0.00599***	0.00185***
	(11.24)	(13.44)	(4.288)	(2.797)
CDS Liquidity Proxy	0.00675***	0.00219***	0.00137	0.00516***
	(5.966)	(3.659)	(1.541)	(10.32)
Lagged Implied Volatility	0.193***	0.164***	-0.0249***	-0.0196**
	(20.17)	(16.06)	(-3.306)	(-2.199)
Firm-level Variables				
Total Debt to Total Assets Ratio	-0.132***	-0.0214	-0.0491***	0.00388
	(-6.942)	(-1.180)	(-3.280)	(0.241)
Market-to-Book Value Ratio	0.0215***	0.0357***	-0.00208*	0.00469*
	(14.00)	(11.44)	(-1.733)	(1.698)
Size as Total Assets	0.0991***	0.0765***	-0.00348	-0.00674
	(16.40)	(11.27)	(-0.732)	(-1.124)
FirmRet	0.0273***	0.0330***	0.0132***	0.0141***
	(7.617)	(7.667)	(4.627)	(3.714)
HV	0.0555***	0.0786***	-0.00436	-0.0208**
	(6.750)	(7.258)	(-0.670)	(-2.165)
Aggregate Variables				
VIX	0.00571***	0.00598***	0.000613***	0.00114***
	(50.79)	(46.81)	(6.879)	(10.06)
Return on the S&P 500 Index	-0.0336**	0.0188	0.00520	0.000526
	(-2.470)	(1.214)	(0.483)	(0.0383)
Yield Curve Slope	-0.00323*	-0.0148***	-0.00506***	0.00701***
	(-1.688)	(-6.254)	(-3.342)	(-3.334)
Treasury Rate	-0.00297***	-0.0110***	-0.00247***	-0.00112
	(-2.774)	(-8.213)	(-2.923)	(-0.946)
Agg Default Spread	-0.00257	-0.0195***	0.00663***	0.00495**
	(-1.106)	(-8.043)	(3.629)	(2.317)
TED Spread	0.00553***	0.00513***	-8.18e-05	0.00436***
	(4.929)	(-3.856)	(-0.0925)	(3.701)
Intercept	-0.892***	-0.617***	0.0822*	0.0861
	(-14.11)	(-8.441)	(1.654)	(1.333)
No. of Observations	20,941	19,129	21,304	19,525
R-squared	0.491	0.411	0.022	0.044
Fixed Effects	Yes	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 15. Impact of Financial Crisis We repeat the main regression for two subsamples of Crisis and Non-crisis periods referring to the 2008 financial crisis (details in paper). Panel A reports IV regressions and panel B reports IV-Skew regressions.

	Panel A) Dependent: IV		**Panel B) Dep.: IV-Skew**	
	NonCrisis	Crisis	NonCrisis	Crisis
Default Risk Variables				
CDS spread	0.0146***	-0.000168	0.000269	0.00616***
	(19.46)	(-0.101)	(0.461)	(3.973)
CDS Liquidity Proxy	0.00228***	0.00644***	0.00479***	0.00330***
	(4.181)	(4.992)	(11.44)	(2.741)
Lagged Implied Volatility	0.285***	0.00831	-0.00895	-0.0420***
	(36.76)	(0.502)	(-1.423)	(-2.720)
Firm-level Variables				
Total Debt to Total Assets Ratio	0.01000	-0.285***	2.19e-05	-0.0331
	(1.177)	(-5.451)	(0.00313)	(-0.683)
Market-to-Book Value Ratio	0.0198***	0.0113**	-0.000437	-0.00643
	(15.18)	(2.181)	(-0.410)	(-1.334)
Size as Total Assets	0.0391***	0.161***	0.00311	-0.0681***
	(11.29)	(7.211)	(1.102)	(-3.289)
FirmRet	0.0272***	0.0264***	0.0125***	0.0171**
	(10.15)	(2.987)	(5.631)	(2.068)
HV	0.0729***	-0.00638	-0.00364	-0.0848***
	(11.64)	(-0.264)	(-0.702)	(-3.748)
Aggregate Variables				
VIX	0.00428***	0.00819***	0.000737***	0.00143***
	(41.18)	(36.49)	(8.568)	(6.775)
Return on the S&P 500 Index	-0.0522***	0.179***	0.00375	0.0526*
	(-5.195)	(5.351)	(0.451)	(1.681)
Yield Curve Slope	-4.92e-05	-0.0126***	-0.00765***	-0.00300
	(-0.0306)	(-3.105)	(-5.743)	(-0.787)
Treasury Rate	0.00121	-0.0101***	-0.00471***	0.00411*
	(1.220)	(-4.217)	(-5.747)	(1.827)
Agg Default Spread	-0.0128***	-0.00881**	-0.000556	0.0105***
	(-4.335)	(-2.310)	(-0.229)	(2.948)
TED Spread	0.0130***	-0.0246***	0.000960	-0.000176
	(12.70)	(-11.18)	(1.137)	(-0.0850)
Intercept	-0.308***	-1.340***	0.0188	0.729***
	(-8.430)	(-5.773)	(0.629)	(3.382)
No. of Observations	32,743	7,327	33,414	7,415
R-squared	0.298	0.588	0.015	0.073
Fixed Effects	Yes	Yes	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 16. Impact of Option Market Liquidity We run the pooled regression on dependent variable *Option-Implied Volatility* and IV-*Skew*, and add the new variable Option Market Liquidity measured by ***"Proportional bid-ask spread"***. (See PBA equation in paper)

	Panel A) Dependent: IV	Panel B) Dep.: IV-Skew
Default Risk Variables		
CDS spread	0.0117***	0.00223***
	(17.70)	(4.122)
CDS Liquidity Proxy	0.00345***	0.00446***
	(6.992)	(11.16)
Lagged Implied Volatility	0.184***	-0.0231***
	(27.14)	(-4.049)
OptionLiquidity	0.0343***	0.00407***
	(32.39)	(5.126)
Firm-level Variables		
Total Debt to Total Assets Ratio	-0.0192**	-0.00665
	(-2.095)	(-0.858)
Market-to-Book Value Ratio	0.0174***	-0.000783
	(12.99)	(-0.692)
Size as Total Assets	0.0452***	-0.00492
	(12.29)	(-1.593)
FirmRet	0.0214***	0.0123***
	(7.850)	(5.302)
HV	0.0678***	-0.0116**
	(10.47)	(-2.106)
Aggregate Variables		
VIX	0.00589***	0.000969***
	(71.63)	(13.83)
Return on the S&P 500 Index	-0.0112	0.00587
	(-1.120)	(0.689)
Yield Curve Slope	-0.00608***	-0.00633***
	(-4.159)	(-5.079)
Treasury Rate	-0.00744***	-0.00255***
	(-9.138)	(-3.688)
Agg Default Spread	-0.0131***	0.00504***
	(-8.057)	(3.664)
TED Spread	-0.000570	0.00155**
	(-0.679)	(2.170)
Intercept	-0.313***	0.0869***
	(-8.035)	(2.655)
No. of Observations	40,070	40,829
R-squared	0.472	0.035
Fixed Effects	Yes	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 17. Implied Volatility Risk Premium

We define "Volatility Risk premium = Implied volatility – Historical Volatility". For historical volatility we use the 6-months rolling standard deviation. Table below test the role of CDS in explaining changes in option-implied volatility risk premium. This table shows the impact including the control variables.

Vol risk premium = f(CDS spread, control variables)

Dependant Variable: ***Implied Vol risk premium***	
Default Risk Variables	
CDS spread	0.0116***
	(17.24)
CDS Liquidity Proxy	0.00355***
	(7.038)
Firm-level Variables	
Total Debt to Total Assets Ratio	-0.0223**
	(-2.394)
Market-to-Book Value Ratio	0.0205***
	(15.02)
Size as Total Assets	0.0587***
	(15.75)
FirmRet	0.0297***
	(10.72)
Aggregate Variables	
VIX	0.00583***
	(69.59)
Return on the S&P 500 Index	-0.0145
	(-1.416)
Yield Curve Slope	-0.00858***
	(-5.755)
Treasury Rate	-0.00717***
	(-8.632)
Agg Default Spread	-0.0122***
	(-7.364)
TED Spread	0.000703
	(0.821)
Intercept	-0.453***
	(-11.48)
No. of Observations	40,070
R-squared	0.602
Fixed Effects	Yes

t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 18. Test of Nonlinearity (Log regressions) for IV and CDS relationship

What if the relationship observed is due to limitations of a linear modeling of IV? To test this, we run the non-linear regression by creating "Log" variables of volatility, default risk, and return variables. The results are shown below.

Dependant Variable: ***Ln (IV) or "Log implied volatility"***	
Default Risk Variables	
Ln (CDS spread)	0.0398***
	(6.647)
Ln (CDS Liquidity Proxy)	0.0244***
	(5.302)
Ln (Lagged IV)	0.447***
	(32.73)
Firm-level Variables	
Total Debt to Total Assets Ratio	0.132***
	(2.636)
Market-to-Book Value Ratio	0.0895***
	(11.60)
Ln (Size)	0.848***
	(5.388)
Ln (FirmRet)	0.00859***
	(7.483)
Aggregate Variables	
Ln (VIX)	0.145***
	(12.98)
Ln (Return on the S&P 500 Index)	-0.00211*
	(-1.863)
Yield Curve Slope	-0.00449
	(-0.585)
Treasury Rate	0.00134
	(0.302)
Agg Default Spread	0.00932
	(1.046)
TED Spread	0.0323***
	(5.838)
Intercept	-3.241***
	(-8.696)
No. of Observations	14,796
R-squared	0.462
Fixed Effects	Yes

t-statistics in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Chapter 4

FORECASTING OPTION-IMPLIED VOLATILITY USING CREDIT RISK

Abstract

We study the time series properties of credit default swap spread in enhancing the forecasting power of future implied volatility and hence future option prices. We run a series of contemporaneous and forward-looking models as encompassing regressions and test whether we can improve the predictability power of implied volatility by extending the forecast model that includes historical volatility (HV) and current implied volatility (IV) to enhance credit default swap. Both in-sample and out-of-sample estimation errors show that CDS can be a significant factor in improving forecasting performance of implied volatility.

Keywords: CDS Spread, Implied Volatility, Encompassing Regressions, Out-of-sample performance

4.1 Introduction

The need for enabling users to manage credit risk has made credit derivatives one of the most important financial innovations of the last quarter of a century. [1]. The volume of the credit default swap market skyrocketed in the last ten years after the first CDS contract was issued by JP Morgan in mid 1990s, making it one of the fastest growing over the counter (OTC) derivatives markets during this period. The CDS contract is a mean of protection against bond default. The seller of the contract is committed to remunerate the buyer with the face value of the bond upon default of the debtor. The buyer in turn has to pay the certain amount as an insurance premium to the seller. This insurance premium is determined as a percentage of the bond's face value and usually named as CDS spread.

Similar to put options, CDS is used for downside risk protection. From the hedging point of view, the insurance and risk protection similarities of the put options and CDS is further shown by the comovement of CDS spread and option implied volatility (See Figure 1). Building on the results found in the previous chapter, this study will contribute to the existing literature by investigating the time-series properties of the CDS when augmented as an enhancement to the future implied volatility forecast models.

[Insert Figure 1 here]

Forecasting option Implied Volatility (IV) is of interest to option market participants, who routinely formulate volatility and option price forecasts for trading and hedging purposes. Given that the IV is a re-parameterization of the market

[1] See Greenspan [2004]

option price, forecasting IV falls within the vast literature on the predictability of asset prices. In addition, to the extent IV yields a forward-looking measure of firm's total risk it has applications predicting firm's beta, credit risk, etc.

The extant literature have explored better estimations of future implied volatility using encompassing regressions on index options. In this paper, we have the advantage of testing forecasting power of on the cross-section of extensive set of equity options. We combine the approach of Chan, Jha, Kalimipalli (2009) and Cao, Yu, and Zhong (2010) by (1) testing on an extensive panel of cross-sectional firms and options data, (2) forecasting the future implied volatility, and (3) finding how incorporating for CDS in the time series model can improve the forecasting performance of the implied volatility of the underlying asset, particularly out-of-sample.

In other words, we use credit risk to forecast out-of-sample option prices. As discussed in previous chapter, credit risk matters for option pricing as options are valued on firms which are not free from default. So we ask that if credit risk matters for option prices, can we develop better out-of-sample forecasts for IV using lagged credit-risk measures?

Previous literature is mostly focused on predicting IVs at aggregate or index level. We, however, provide a characterization of IVs forecasts at individual firm level and use a robust measure of credit risk i.e. CDS spreads of the underlying firm written for senior sub-ordinated debt.

We provide a first comprehensive study of role of credit risk on IV forecasting using an exhaustive sample of cross-sectional firms.

4.2 Literature Review

Previous literature on forecasting IV falls broadly under four main categories: (1) Modeling IV evolution at the index level (SPX: options on S&P 500 index), where a time series model for the changes in IV is used. (2) Modeling moments of risk neutral distribution (RND), where variance, skewness and kurtosis of the RNDs are modeled through time. (3) Encompassing regressions, where past historical volatility, and IVs are used to forecast realized volatility (RV) or option IV. (4) Implied volatility functions, where cross-sectional IV surface is defined as a polynomial function of moneyness and maturity to capture the pricing biases and option prices are obtained using numerical methods.

4.2.1 Modeling IV Evolution at the Index Level

Modeling volatility indices is an important part of IV predictability studies. Konstantinidi, Skiadopoulos, and Tzagkaraki (2008) investigate whether the evolution of of implied volatility can be forecasted by focusing on a number of European and US implied volatility indices. They use an economic model that estimates the change in lagged IV based on a number of lagged default risk and firm level variables. They test for both point and interval forecasts and study the statistical and economic significance of these forecasts. They further use principal component, ARIMA, ARFIMA, and VAR models and assess the performance by trading strategies in the volatility futures markets. Although they find statistical support for predictability, they show that the findings are not economically significant.

In our paper, we focus on cross section of the firms and are interested to be able to forecast IV for each option ID. But how do we test the forecast performance of the results?

Goncalves and Guidolin (2006) provide a clear framework for calculating estimation for options with the specification in mind that any averaging should be performed on count of traded options only. We use their approach in designing MSE, RMSE, and MAE.[2] Motivated by empirical evidence on lack of stability of the parameters characterizing the implied volatility surface (IVS) in option prices, Goncalves and Guidolin (2006) investigate the predictability patterns of IVS. They employ a two-stage model that first estimates the surface along cross-sectional moneyness and maturity dimensions. Then in the second stage they model the dynamics of the first-stage coefficients. They find that the movements of the S&P 500 IVS are in fact very predictable. As a result, a set of profitable delta-hedged trading rules can be created, however the spreads will disappear due to higher transaction costs and on a wider sample of IVs.

Although performed at the index level, Goncalves and Guidolin's paper is particularly interesting as the performance measures introduced in the paper are able to identify the predictability power of IV forecasts across different models. We employ the same measurement errors in comparing forecast performance of our models and show impact of CDS inclusion on the wide cross section of options.

4.2.2 Modeling Moments of Risk-neutral Distribution (RND)

A parametric approach for unveiling implied volatility is to model the moments of the underlying RND. This can be done both on the cross-section and the index

[2] MSE stands for Mean Square Error, RMSE for Root Mean Square Error, and MAE for Mean Absolute Error. The definitions of these error terms will be discussed in the regression section.

level. By focusing on data extracted from the market prices of Standard & Poor's (S&P) 500 index options, Neumann and Skiadopoulos (2013) investigate whether there are predictable patterns in the dynamics of higher-order risk-neutral moments (RNMs). They conduct a horse race among alternative forecasting models within an out-of-sample context over various forecasting horizons (daily, weekly, monthly). They find that higher RNMs can be statistically forecasted. However, only the 1-day-ahead skewness forecasts can be economically exploited since the economic significance vanishes once the transaction costs is incorporated.

Comparing the index options vs. individual equity options, and in order to explain the risk-neutral skewness implied from option prices, the findings of Dennis and Mayhew (2002) empirically establish a link between the risk neutral skewness and the systematic risk of the underlying stock. They explain the structural difference in distributions by investigating the relative importance of several firm characteristics such as implied volatility, firm size, trading volume, leverage, and beta. They show that risk-neutral skewness tends to be more negative for stocks with larger betas. This is an evidence for the importance of market risk in option pricing. Their findings show that the index options have a more pronounced volatility smile/smirk than individual equity options.

If the moments are successfully modeled, this means the dynamics of implied volatility surfaces can also benefit from the modeling.

In the same line of parametric approach to index level forecast, Panigirtzoglou and Skiadopoulos (2004) try to model the dynamics of implied distributions by obtaining a parsimonious description of the dynamics of the S&P 500 implied cumulative distribution functions by applying PCA[3]. After identifying the factors,

[3] Principal Component Analysis

they employ arbitrage-free Monte Carlo simulation methods that model the evolution of the whole distribution as a diffusion process. Traditionally, modeling only the first two moments as diffusion processes is deemed to be sufficient to span the whole IV. Their findings have important implications for "smile-consistent" option pricing and for risk management, and they also test their model out-of-sample.

4.2.3 Encompassing Regressions

Encompassing regressions are an additional line of studies with the goal of better modeling of IV and RV. In their papers Christensen and Prabhala (1998) model the relationship between IV and RV. Chan, Jha, and Kalimipalli (2009) and Becker, Clements, and White (2007), both use the S&P 500 implied volatility index (VIX) when running encompassing regressions. VIX modeling through encompassing regressions shows better out-of-sample performance. Our paper's methods are closely related to Chan, Jha, and Kalimipalli (2009), in that we also use the encompassing regressions to model IV.

However, an important contribution of this paper is that we use a wide cross section of individual options, as opposed to limiting the tests to the options index.

4.2.4 Modeling through Implied Volatility Functions and IV Surface

Implied volatility functions are deterministic approach to modeling IV. Derman and Kani (1994), Dupire (1994), and Rubinstein (1994) hypothesize that asset return volatility is a deterministic function of asset price and time, and develop a deterministic volatility function (DVF) option valuation model that has the potential of fitting the observed cross section of option prices exactly. Dumas, Fleming, and Whaley (1998) use S&P 500 options from to examine the predictive and hedging performance of the DVF option valuation model and find it that is no better than an ad hoc procedure that merely smooths Black–Scholes (1973) implied volatilities across exercise prices and times to expiration.

Although deterministic volatility models allow for more flexible volatility surfaces, these models refrain from introducing additional risk factors. This means we need stochastic models to introduce additional risk factors, and options are then needed for spanning of the pricing kernel. Buraschi and Jackwerth (2001) develop a statistical test based on this difference in spanning. Again, they use index level options data and show that both in- and out-of-the-money options are needed for spanning which supports stochastic volatility, interest rates, or jumps models.

Throughout the four different lines of literature discussed, it is shown that most forecasting attempts of IV have been empirically tested on index level data or a limited selection of underlying firms, and not the cross section of options. In our paper, we test for this by testing on an extensive collection of options data. All tests will be run in a time series pattern and then averaged cross-sectionally.

4.3 Data and Summary Statistics

To perform this analysis, we employ cross-section and time series observations from the following databases: Option Metrics, Markit data (CDS), and CRSP.

The credit default swap data is available from January 2002. As such, we collect matching data from January 2002 to December 2009, in order to cover the matching time period with the CDS data.

Markit database reports the CDS spread on the 1-year, 5-year and 10-year contracts. For the purpose of our research and in order to be dealing with the highest liquidity, we limit the data to 5-year spreads only.

Table 1 displays the Summary statistics of the complete dataset used for the forecasting. The three variables of interest are historical volatility, CDS, and IV since we are focused on time-series forecast of implied volatility. Panels A and B show the characteristics of these three variables across moneyness and maturity bins as well as through time. Based on the distribution of observations, we can see that across all moneyness groups, most liquid bins lie under short maturity, and across all maturity groups, most frequent observations belong to at-the-money and out-of-the money put options.

Panel C of Table 1 shows that during crisis and distress periods, as expected, IV increases to high of 51% on average with the median of 45%.

Are lagged variables significant cross-sectionally? Given the purpose of this paper is the time series analysis, we shall soon start building and performing tests on a time series basis. However, in order to highlight the significance of the variables chosen for the model, we run one set of univariate pooled panel regressions on the three key lagged variables. Table 2 shows the result of univariate regression of lagged variables on implied volatility. As indicated all three lagged variables of CDS_{t-1}, HV_{t-1}, and IV_{t-1} are significant variables for the cross section of implied volatility.[4] Note that the R-squared of the regressions are not very high due to simplicity of the model at this stage. Future research will focus on improving the liquidity as well as variables.

In the next section we design and test the time series forecast models.

4.4 Methodology

In this section we explain the methodology used, mainly the design of the forecasting model of future implied volatility.

4.4.1 Volatility Measures Background

By definition volatility aims to capture the strength of the (unexpected) return variation over a given period of time. However, two distinct features importantly differentiate the construction of all (reasonable) volatility measures. First, given a set of actual return observations, how is the realized volatility computed? This means that the emphasis is explicitly on "ex-post" measurement of the volatility. Second, decision making processes often require forecasts of future return volatility. The focus this way is on "ex-ante" expected volatility.

The ex-ante concept requires a model that can effectively map the current information set into a volatility forecast. In contrast, the (ex-post) realized volatility

[4]Lagged CDS is the value of CDS for one period before for the same firm, i.e. last trading day's value. Note that this results in many missing values that have to be dropped when the time series regressions run. No interpolation is performed for the lagged variables. Same logic applies to lagged IV value. See Cao and Yu (2010) for a similar data approach for time series forecast.

may be computed (or approximated) without reference to any specific model, thus rendering the task of volatility measurement essentially a nonparametric procedure. An example is a rolling historical volatility measure.

In addition to these model classes, the **implied volatility** approaches are also significantly covered in the literature. The implied volatilities are based on a parametric volatility model for the returns, as defined above, along with an asset pricing model and an augmented information set consisting of options prices and/or term structure variables.

Various design of IV and HV have been used in conjunction with default risk studies. Campbell and Taksler (2003) used 180-day rolling standard deviation of equity return as an explanatory variable for the credit spread. Collin-Dufresne, Goldstein and Martin (2001) suggest the stock option implied volatility to be one of the important determinants of the changes in CDS spreads. We chose to use the 120 days rolling historical volatility because it corresponds to the maturity of the put options observed in our sample.

In our research we retrieve IV from the Optionmetrics database, and hence it is important to explain the estimation method employed. The calculated interpolated implied volatility for each option on each day, uses a methodology based on a kernel smoothing algorithm. The data is first organized by the log of days to expiration and by "call-equivalent delta" (delta for a call, one plus delta for a put). A kernel smoother is then used to generate a smoothed volatility value at each of the specified interpolation grid points. At each grid point on the volatility surface, the smoothed volatility is calculated as a weighted sum of option implied volatilities.[5]

The implied volatility can also be seen as the volatility prognosis given the information available at the current state and hence it will be state-sensitive. Nonetheless the estimation of the implied volatility by CDS can be complicated due to the lack of data on CDS trades on a consistent daily basis. We do deal with missing variables which is due to the illiquidity of CDS compared to IV observations. Currently, interpolation is done for a single day missing (with trade data available prior and after). Future research can benefit from the augmented dataset (time period) and advanced interpolation (two or more consecutive days missing) in order to deal with this issue.

4.4.2 Out-of-sample Evaluation background

What is the purpose of out-of-sample forecasts? Cross-validation is the process of assessing how the results of a statistical analysis will generalize to an independent data set. Note that IV is a reparamaterized value. If the model has been estimated over some, but not all, of the available data, then the model using the estimated parameters can be used to predict the held-back data. We test the performance by measuring the out-of-sample mean squared error (MSE), also known as the mean squared prediction error, in order to test the strength of the model in forecasting the desired variable. In addition, we measure the mean absolute error (MAE) and compare these statistics for each forecast model.

[5] See OptionMetrics Manual for additional details. Also, limited additional details provided in Appendix.

With the data and measures description above, we continue to build the baseline regression model of forecasting implied volatility in the next section.

4.5 Regression Model Specification and Empirical Results

Based on the descriptions above, we can now specify the regression model to test our hypotheses. We split the analysis into two sets of contemporaneous and lagged models, as well as out-of-sample performance models.

Because the average maturity for the options used in our implied volatility estimation is about four months, both historical volatility and future realized volatility are computed over 120 trading days in this exercise in order to match the horizon of option-implied volatility. We use daily data for the regression with the Newey and West (1987) correction to the standard errors for autocorrelation and heteroscedasticity.

We report two tests for each of contemporaneous and forward looking in-sample models, one "with" and one "without" inclusion of the CDS measure. The results can be found in Tables 3 and 4. We first explain the model and then elaborate on the findings of these two tables.

The contemporaneous regression specification, hence, can be explained per below:

$$IV_t = \beta_0 + \beta_1 HV_t + \epsilon_t \tag{4.1}$$

$$IV_t = \beta_0 + \beta_1 HV_t + \beta_3 CDS_t + \epsilon_t \quad (4.2)$$

Table 3 shows the results of our findings. The specification above is run on each option ID and then averaged through all option IDs and reported. HV tend to be an insignificant factor in determining the contemporaneous implied volatility. CDS, on the other hand, shows to be statistically and economically significant, and the inclusion improves the R^2 of the forecast model to 27.6%.

On Forecasting Levels vs. Changes: We perform level regressions and not changes. This is mostly because the irregular intervals or missing data for CDS observations limits the design of any change regressions. From a statistical perspective, first differencing is appropriate if the dependent variable and regressors are integrated, but this is difficult to determine for our irregularly spaced data. The potential extension of the data set and additional interpolation techniques will allow us to explore the change forecast in future studies.

We then move on to in-sample forecasts with 1-day ahead. Table 4 shows the results for this forecast.

$$IV_t = \beta_0 + \beta_1 HV_{t-1} + \beta_2 IV_{t-1} + \epsilon_t \quad (4.3)$$

$$IV_t = \beta_0 + \beta_1 HV_{t-1} + \beta_2 IV_{t-1} + \beta_3 CDS_{t-1} + \epsilon_t \quad (4.4)$$

where IV_t is the daily volatility on day t, HV_t is the historical volatility on day t, and CDS_t is the 5-year observed CDS spread on day t.

We estimate each of the above models for the sample period. (To report the results, we run this time series regression on each cross-sectional firm and then report the average betas, the cross-sectional mean of the regression coefficients, and the Newey-West adjusted t-stats, as well average R-squared values.)

To test multiple horizons, we generate multistep forecasts on a given day for each option. We examine 1-day-ahead (accounting for daily forecast),and 5-day-ahead (accounting for weekly forecast). The results show that CDS is a significant estimator of the forecasted IV.

In the next section we investigate out-of-sample performance of the designed forecast models.

4.5.1 Out-of-sample Tests

The reasoning for looking at the out-of-sample forecasting performance in addition to the in-sample fit comes from the objective of the analysis. In forecasting it is not necessarily the model that provides the best in-sample fit that produces the best out-of-sample volatility forecast, which is the main objective of the analysis;[6] Hence it is common to use the out-of-sample forecast to aid the selection of which model is best suited for the series under study;[7]. The out-of-sample forecast refers to that the data used in the fitting of the model is different from the data used for evaluating the forecasts. Typically the data in divided into two subsets, one in which the model parameters are fitted (estimation subsample), and another subset used to evaluate the forecasting performance of the models (forecasting subsample).

[6] Shamiri and Isa (2009).

[7] Andersen and Bollerslev (1998), Hansen and Lunde (2001) and Brandt and Jones (2006)

In Tables 3 and 4 we tested in-sample performance. Tables 5 and 6 investigate out-of-sample performance for two different step sizes: Table 5 shows the result for $k = 1$, which corresponds to daily forecast and Table 6 shows the results for weekly forecast, i.e. $k = 5$.

To calculate the out-of-sample results we estimate each model based on a rolling window on the sample period, and based on the number of days to maturity, then we generate multistep out-of-sample forecasts for each day. To feed sufficient data points, out-of-sample is not performed on the first year of the data. In summary, we can explain the process as below:

- Consider an option with a specific option ID (that belongs to a moneyness and maturity group). We derive the observed IV for each day from OptionMetrics from the in-sample period.
- We use the underlying firm's 5 year CDS spread, and the recorded historical volatility for each day. (Note that lagged CDS only exists if there exists a trade observation for the previous period. Otherwise, need to drop the observation)
- We split the lifetime of the observed option into two periods: estimation subsample (time= 0 to t) and forecasting subsample (t to T).
- Using the fitted values until time t, we then forecast IV for $T - t$.
- The reported errors measure performance on the forecasting subsample only and not the fitted values in the estimation subsample.

For both models we calculate three error values: MSE, RMSE, and MAE values across all option IDs. [8] Based on comparison of all three measures across bins, we can see that: First, the model with CDS enhancement outperforms the rest of the forecast models since the error values are lowest. Observing across each of the bins, we can see that overall the forecasting error drops when the model is enhanced by CDS (i.e. last row). In addition, for the complete sample, (including all maturity and moneyness groups in one sample), the RMSE drops by 4 basis points between the first and last models.

Second, we can also conclude that Long option bins in general have the best out-of-sample performance of IV forecast.The forecasting error for these bins (Long ITM, Long ATM, and Long OTM), are consistently measured lowest in their groups.

Overall, inclusion of CDS can decrease the out-of-sample forecasting measurement error. Next, we investigate whether the predictability power of CDS acts differently during crisis periods and across industry.

4.5.2 Impact of Crisis and Industry

We further like to compare the out-of-sample performance of IV forecasting with the enhanced CDS, on the key sub-samples. Table 7 investigates impact of financial crisis on model's out-of-sample performance. The results show that the estimation error is consistently lower for non-crisis periods. The test of differences in mean residual of the two subsamples are also statistically significant. This shows

[8](i)The root mean squared prediction error in implied volatilities (RMSE) is the square root of the average squared deviations of BS implied volatilities from the model's forecast implied volatilities, averaged over the number of options traded. (ii) The mean absolute prediction error in implied volatilities (MAE) is the average of the absolute differences between the BS implied volatility and the model's forecast implied volatility across traded options.

that although we can improve the out-of-sample performance of IV forecast, our progress is somewhat restricted during crisis periods.

Finally, we like to see whether financial firms do better in out-of-sample forecasting or not. Following the same approach as above we can test this and see that non-financial and financial firms do not show any "economically" significant difference in out-of-sample performance of their IVs, despite the statistically significant results in test of difference in mean. Table 8 shows that the values are very similar and the difference is economically minor.

4.6 Conclusion

In previous chapter we showed how credit risk impacts option pricing. This association is an interesting component that can further be explored in the time series context. Specifically, we can use the forward-looking information content of CDS in forecasting option Implied Volatility.

In addition, forecasting option Implied Volatility (IV) is of interest to option market participants, who routinely formulate volatility and option price forecasts for trading and hedging purposes.

Previous literature is mostly about predicting IVs at aggregate or index level. In this paper, however, we provide a characterization of IVs forecasts at individual firm level. We use a robust measure of credit risk i.e. CDS spreads of the underlying firm written for senior sub-ordinated debt and therefore provide a first comprehensive study of role of credit risk on IV forecasting using an exhaustive sample of 550 firms for the 2002-2009 period.

Inclusion of CDS improves the precision of implied volatility forecast, specifically for out-of-sample. We further show that the forecasting error of implied volatility will vary across moneyness and maturity indicating deep OTM and deep ITM options can be subject to higher pricing error. Longer maturity options have lowest estimation errors and show that

4.7 References

Becker, Ralf, Adam E. Clements, and Scott I. White. “Does implied volatility provide any information beyond that captured in model-based volatility forecasts?.” Journal of Banking & Finance 31.8 (2007): 2535-2549.

Berndt, Antje, and Iulian Obreja. “Decomposing European CDS Returns. Review of Finance 14, no. 2 (2010): 189-233.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31, 307–326.

Bongaerts, Dion, Frank De Jong, and Joost Driessen. “Liquidity and liquidity risk premia in the CDS market.” Working paper, University of Amsterdam, 2008.

Brigo, Damiano, and Kyriakos Chourdakis. “Counterparty Risk for Credit Default Swaps: Impact of spread volatility and default correlation.” International Journal of Theoretical and Applied Finance 12, no. 07 (2009): 1007-1026.

Buraschi, Andrea, and Jens Jackwerth. “The price of a smile: Hedging and spanning in option markets.” Review of Financial Studies 14.2 (2001): 495-527.

Campbell, J. T., and G. B. Taksler. “Equity Volatility and Corporate Bond Yields.” Journal of Finance, 58 (2003), 2321–2349.

Cao, Charles, Fan Yu, and Zhaodong Zhong. “The information content of option-implied volatility for credit default swap valuation.” Journal of financial markets 13, no. 3 (2010): 321-343.

Cao, Charles, Fan Yu, and Zhodong Zhong. “How Important is Option-Implied Volatility for Pricing Credit Default Swaps?.” Working Paper, UC Irvine, 2006.

Chan, Wing Hong, Ranjini Jha, and Madhu Kalimipalli. “The economic value of using realized volatility in forecasting future implied volatility.” Journal of Financial Research 32.3 (2009): 231-259.

Christensen, Bent J., and Nagpurnanand R. Prabhala. “The relation between implied and realized volatility.” Journal of Financial Economics 50.2 (1998): 125-150.

Collin-Dufresne, P.; R. Goldstein; and S. Martin. “The Determinants of Credit Spread Changes.” Journal of Finance, 56 (2001), 2177–2207.

Das, Sanjiv R., Paul Hanouna, and Atulya Sarin. “Accounting-based versus market-based cross-sectional models of CDS spreads.” Journal of Banking & Finance 33, no. 4 (2009): 719-730.

Dennis, P., Mayhew S., “Risk-neutral skewness: Evidence from index options.” Journal of Financial and Quantitative Analysis (2002) 37, 471 493.

Derman, E. and I. Kani, 1994, “The Volatility Smile and Its Implied Tree”, Goldman Sachs Quantitative Strategies Research Notes, and “Riding on a Smile”, RISK 1994.

Dumas, Bernard, Jeff Fleming, and Robert E. Whaley. “Implied volatility functions: Empirical tests.” The Journal of Finance 53.6 (1998): 2059-2106.

Dupire, B. “Pricing with a Samile”, RISK, 1994.

Ericsson, Jan, Jacobs, Kris and Oviedo, Rodolfo, The Determinants of Credit Default Swap Premia (January 2005). Journal of Financial and Quantitative Analysis (JFQA).

Greenspan, A. “Risk and uncertainty in monetary policy.” American Economic Review (2004), 33-40.

Goncalves, Silvia, and Massimo Guidolin. “Predictable dynamics in the S&P 500 index options implied volatility surface.” The Journal of Business 79.3 (2006): 1591-1635.

Konstantinidi, Eirini, George Skiadopoulos, and Emilia Tzagkaraki. “Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices.” Journal of Banking & Finance 32.11 (2008): 2401-2411.

Neumann, Michael, and George Skiadopoulos. “Predictable dynamics in higher-order risk-neutral moments: Evidence from the S&P 500 options.” Journal of Financial and Quantitative Analysis 48.03 (2013): 947-977.

Newey, W. K., and K. D. West. “Hypothesis testing with efficient method of moments estimation.” International Economic Review (1987): 777-787.

O’Kane, Dominic, and Stuart Turnbull. “Valuation of credit default swaps.” Lehman Brothers Quantitative Credit Research Quarterly 2003 (2003): Q1-Q2.

Panigirtzoglou, Nikolaos, and George Skiadopoulos. ”A new approach to modeling the dynamics of implied distributions: Theory and evidence from the S&P 500 options.” Journal of Banking & Finance 28.7 (2004): 1499-1520.

Rubinstein, M. “Implied Binomial Trees”, Journal of Finance, 1994.

Tang, Dragon Yongjun, and Hong Yan. “What moves CDS spreads.” Working paper, University of Hong Kong, University of South Carolina and Shanghai Advanced Institute of Finance, 2011.

Zhang, Benjamin Yibin, Hao Zhou, and Haibin Zhu. “Explaining credit default swap spreads with the equity volatility and jump risks of individual firms.” Review of Financial Studies 22, no. 12 (2009): 5099-5131.

4.8 Appendices

4.8.1 Appendix A: Complete Variables Glossary

In this section, the detailed definitions for each of the variables used in the baseline regression(s) are provided. (Please note that the variables below are similar to previous chapter's glossary.)

• **CDS spread:** 5-year maturity CDS spread, daily observations from Markit. When data is missing for one day in between two observations, it gets interpolated. Any consecutive missing for more than one day, is treated as missing value and automatically is dropped from the panel regressions.

• **Option-implied volatility (IV):** Put option IV, daily observations, daily observations, directly from Option Metrics. [9]

• **CDS Liquidity Proxy:** equals to $(No.ofContributors{-}AvgNo.ofContributors)\times CDSspread$, or "Demeaned number of contributors times CDS spread". The number of contributors are retrieved directly from Markit.

• **Firm Return (%):** For each underlying firm daily price is retrieved from CRSP database. $FirmRet = \frac{P_1 - P_0}{P_0}$.

• **Historical Volatility (%):** Standard deviation of a fund's daily returns over a rolling 6-month period.

• **Market to Book Value:** Tobin's Q calculated with data from Compustat quarterly files.

[9] IVs in OptionMetrics are retrieved from the Volatility Surface data file. The calculated interpolated implied volatility for each option on each day, uses a methodology based on a kernel smoothing algorithm. The data is first organized by the log of days to expiration and by "call-equivalent delta" (delta for a call, one plus delta for a put). A kernel smoother is then used to generate a smoothed volatility value at each of the specified interpolation grid points. At each grid point on the volatility surface, the smoothed volatility is calculated as a weighted sum of option implied volatilities. (See OptionMetrics Manual for additional details.)

• **Debt ratio:** = Total Debt/Total Asset; data from Compustat quarterly files; defined as total liabilities divided by the sum of total liabilities and market capitalization.

• **VIX:** Daily market volatility, $VOLATILITYS\&P500(^{V}IX)$, retrieved from CRSP, and available on OptionMetrics.

• **TED Spread:** The difference between the interest rates on interbank loans and on short-term U.S. government debt (T-bills). It is calculated as the difference between the three-month LIBOR and the three-month T-bill interest rate. Data files for each variable retrieved from Datastream.

• **IV Skew:** As formulated in the paper, IV skew is defined for each observation as "Avg OTM implied volatility – Avg ATM implied volatility":

$$IVSkew_{i,t} = AVERAGE_{OTM}(IV_{i,t}) - AVERAGE_{ATM}(IV_{i,t})$$

Note: This section of the Appendix is identical to previous essay's data organization steps, since the underlying datasets for the last two essays are the same.

4.8.2 Merging CDS Markit and Option Metrics

There are total of 1833 Ticker/Comp Name combination in the Markit Database.Out of these 1833 , 34% are private or subsidiaries (See appendix A on this clean-up activity). Removing these , would leave a remainder of 1203 public to work with. The next step is to limit data to firms with Options. Matching the 1203 with Options Metric data, we are down to 1063 which both have CDS written on the underlying bond, and options written on their underlying company. The sample that we work with would have 1063 unique firm IDs. (see chart below)

4.8.3 Cleaning up Option Metrics Data

I downloaded all Option Metrics data from 1996 to 2011 inclusive. The Option Prices file (under Option Metrics Data files¿¿Options) gets downloaded with all possible variables (32 variables in total) in monthly files format from 1996 until 2010. (The average volume for every month is 1-2GB)
There are 32 listed variables downloaded, per below:

FILTERS:
1) DROP UNNECESSARY VARS ($ss_flag, index_flag, exchange_d, issue_type$) keep if cp_flag == "C" | cp_flag == "P" keep if ss_flag == 0 (121656 observations deleted) keep if index$_flag$ == 0 (536045 observations deleted) drop if $exchange_d$ == 32768 (0 observations deleted) keep if $issue_type$ == "0" (1681856 observations deleted)

2)prior to saving the final file, I "sort secid optionid date" to organize the data; it makes it much easier in the data matching, & give the compress command one more time before saving the datafile.
3)LITERATURE FILTERS: Many in our sample do not have actively traded options. The choice of non-zero open interest emphasizes the information content of options that are currently in use by market participants. We also drop all zero volume options. I also exclude all options that violate Put-Call parity. For the first round of the analysis, I keep only put options.

4.8.4 Compustat Fundamentals Data Collection

I collect and compute the quarterly Company Variables (i.e. Size, Book/Market, and Leverage) for all of the in the sample. I also winsorize data and store statistics both before and after winsorization.
For merging with Option Metrics and CDS, I keep the pre-winsorization file.

4.8.5 Report on CDS Markit Data Clean-up for TICKERS & COMP NAMES

Markit Look-Up Dimensions: TICKER, Company Name, Date Period(s)

Issues:
1. Tickers may get recycled and as such multiple matches available when merged with other databases
2. Company names are in short form or miss spaces/dots in Markit so cannot automatically be matched. Need manual review.

Solution:
Review the complete Ticker/CompName list from Markit and verify their identification (i.e. match with the unique PERMNO from CRSP).
• Use Google/Yahoo Finance search engines.
• Confirm Private/Public Companies
• Find alternative names and tickers

Manual Correction Steps
1. Look up the ticker in CRSP full database
2. If the ticker and company name match fully, there should be one unique PERMNO. Record any alternative names.
3. If there are alternatives, provide correct match for the time period provided in Markit.
4. If there are no matches, use Google and Yahoo finance searches. – Is the company a Private company?

Results Statistics

There are total of 1833 unique Ticker/Comp Name combination in the CDS Markit Database. The final correction and inclusion provided final match of 1203 companies in total.

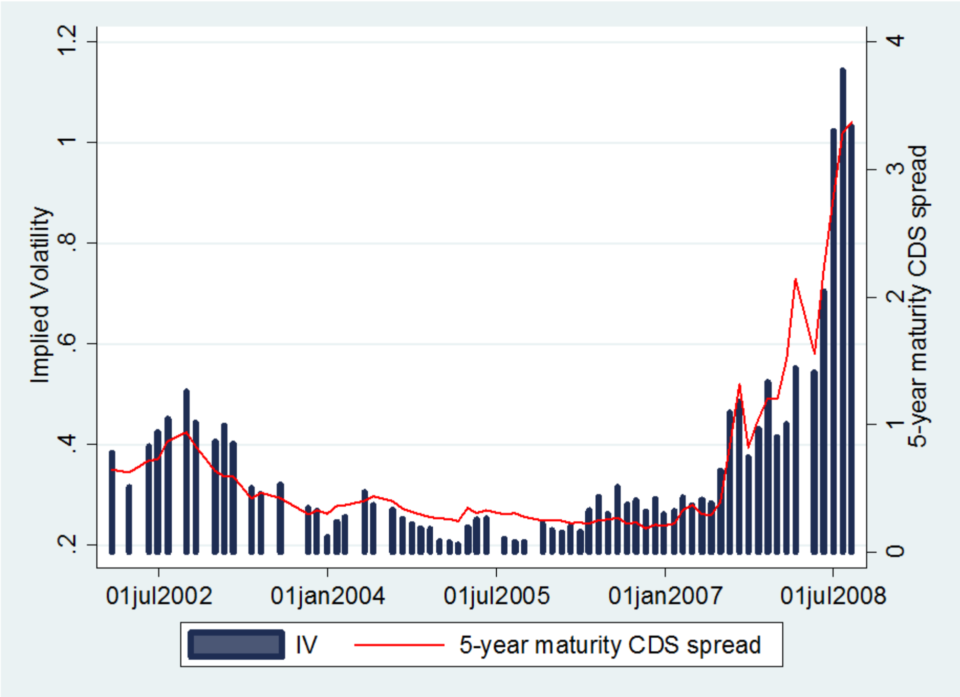
From the remaining, the provided correction notes indicate that:

a. 15 firms are subsidiaries with parent company not found before, so parent ticker is provided.
b. Total of 100 unmatched tickers are Private firms.
c. Total of 340 unmatched tickers are Subsidiaries with their parents present in the database.
d. Fewer than 10 firms are result of mergers or acquisitions that are still valid for inclusion. Details provided.
This is the result of individual scrubbing of each firm/ticker.

4.9 Essay III Figures and Tables

Figure 1. Time-series comovement of CDS spread and implied volatility

This chart shows the IV and CDS dynamics for Lehman Brothers (**LEH**) and Merrill Lynch (**MER**) over the Jan2002-Jan2009 sample. Consistent with the hypotheses of the paper on the forecasting power of CDS, we can see the comovement of CDS and implied volatility. Our goal here is to trace the time-series relationship of CDS in estimating future implied volatility and use it in the IV forecast model.



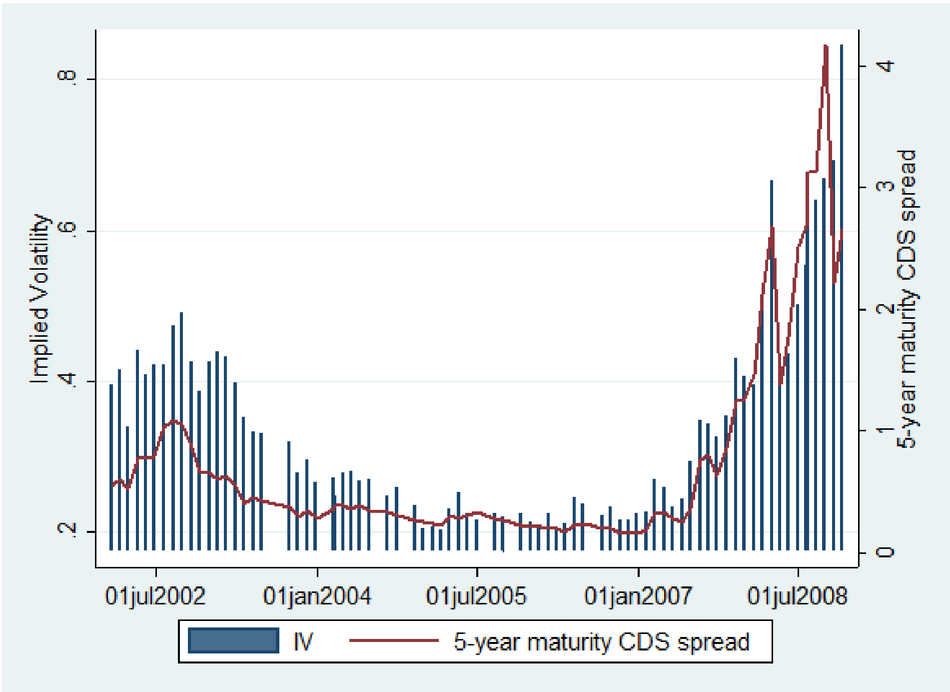


Table 1. Summary Statistics

Panel A reports the overall summary statistics of the time-series means of 550 sample firms for each of the key variables. Panel B reports the summary statistics across bins. CDS Spread is the five-year composite credit default swap spread; historical volatility is the 252-day historical volatility; implied volatility is the volatility inferred from put options with nonzero open interests; The sample period extends from January 2002 through 2009.

Panel A: All data					
	Mean	**Q1**	**Median**	**Q3**	**Std.Dev.**
CDS Spread	1.10	0.24	0.47	1.09	2.17
Historical Volatility	0.38	0.03	0.25	0.58	0.05
Implied Volatility	0.36	0.26	0.33	0.43	0.17
Nobs	205441				

Panel B: Across Bins				
		Short	**Medium**	**Long**
CDS Spread	**ITM**	1.28	1.28	1.20
Historical Volatility		0.39	0.43	0.45
Implied Volatility		0.40	0.36	0.33
Percent Obs.		11.45%	7.00%	5.51%
CDS Spread	**ATM**	9.76	0.99	0.96
Historical Volatility		0.33	0.34	0.41
Implied Volatility		0.34	0.33	0.32
Percent Obs.		18.75%	10.12%	6.83%
CDS Spread	**OTM**	1.20	1.10	1.04
Historical Volatility		0.38	0.38	0.45
Implied Volatility		0.43	0.38	0.35
Percent Obs.		17.55%	12.96%	9.83%

Panel C: Implied volatility across years		
	IV Mean	**IV Median**
2002	0.4707201	0.428546
2003	0.3727789	0.348667
2004	0.3205156	0.3027075
2005	0.2985788	0.282648
2006	0.3101929	0.297454
2007	0.3316317	0.312099
2008	0.5067201	0.450139

Table 2. Pooled Regression of Implied volatility

The table presents univariate pooled regression of Implied volatility on the set of key lagged variables to be employed in the forecasting model. The pooled regression is aimed to highlight the comovement of each with the implied volatility.

Lagged CDS = previous observation day's CDS for the issuer firm.
Lagged IV = previous observation day's IV for the same Option ID.
Lagged HV = previous observation day's HV for the issuer firm.

Dependent Var: IV	(1) Lagged IV	(2) Lagged HV	(3) Lagged CDS
IV			
Lagged CDS	0.0118***		
	(21.08)		
Lagged HV		0.0640***	
		(7.729)	
Lagged IV			0.354***
			(75.40)
Intercept	0.327***	0.325***	0.232***
	(585.9)	(510.0)	(142.4)
Observations	76,258	117,519	105,602
R-squared	0.010	0.002	0.089

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3. Contemporaneous regression of implied volatility (run on each option)

This table estimates the following three models of variations with and without CDS:

$IV_t = \beta_0 + \beta_1 \times HV_t$

$IV_t = \beta_0 + \beta_2 \times CDS_t$

$IV_t = \beta_0 + \beta_1 \times HVt + \beta_2 \times CDS_t$

Where CDS is the observed credit default swap spread, HV is the rolling historical volatility, and IV is the option implied volatility.

We run this time series regression on each cross-sectional firm and then report the average betas. Both cross-sectional mean of the regression coefficients, the (Newey-West adjusted) t-stats, as well as standard errors and average R-squared are presented.
(*Italic values show the average t-stat of cross-sectional values.)

	HV without CDS	CDS without HV	HV and CDS
Average Beta_0 (β0)	0.3347***	8.1225***	0.4293***
	30.27	*29.09*	*14.09*
Average Beta_1 (β1)	0.0449258		2.577973
	0.51		*.6215886*
Average Beta_2 (β2)		3.5876***	4.8797***
		3.43	*2.69*
N	13429	33081	9921
Average R-squared	8.51%	18.28%	27.62%

Table 4. Time-series regression of implied volatility (run on each option)

This table estimates the forecasting power of the CDS spreads when augmented to the IV forecasting model. Both regression coefficients and Newey-West adjusted t-stats are presented.

This table estimates the following two models with and without CDS:

$$IV_t = \beta 0 + \beta 1 \times HV_{t-1} + \beta 2 \times IV_{t-1}$$

$$IV_t = \beta 0 + \beta 1 \times HV_{t-1} + \beta 2 \times IV_{t-1} + \beta 3 \times CDS_{t-1}$$

Where HV is the rolling historical volatility, and IV is the option implied volatility. CDS is the 5-year maturity credit default swap for the underlying firm. We run this time series regression on each cross-sectional firm and then report the average betas. Both cross-sectional mean of the regression coefficients, the (Newey-West adjusted) t-stats, as well as average R-squared are presented.

	Without CDS	**With CDS**
Average Beta_0 (β0)	-0.0256***	-0.0873***
	-4.20	*-5.25*
Average Beta_1 (β1)	0.1092***	0.0561***
	12.29	*11.50*
Average Beta_2 (β2)	0.3289***	0.0750***
	14.38	*10.55*
Average Beta_3 (β3)		0.0027***
		8.30
N	8379	6284
Average R2	0.21	0.29

Table 5. Daily Out-of-sample performance of the model specifications for each of the maturity and moneyness bins (k= 1 day)

Out-of-sample performance of the model specifications for each one of the implied volatility forecast. We do this estimation for all three models, and calculate the errors for nine different bins as well as all sample. The mean square error (MSE), the root mean squared prediction error (**RMSE**), and the mean absolute prediction error (**MAE**) are provided.

(i)The root mean squared prediction error in implied volatilities (RMSE) is the *square root of the average squared deviations of BS implied volatilities from the model's forecast implied volatilities*, averaged over the number of options traded. (ii) The mean absolute prediction error in implied volatilities (MAE) is the average of the *absolute differences between the BS implied volatility and the model's forecast implied volatility* across traded options.

Table 5. Out-of-sample performance of IV prediction: k= 1 day

		ITM			ATM			OTM		
	All sample	Short	Medium	Long	Short	Medium	Long	Short	Medium	Long
Panel A: HV_{t-1}										
MSE	0.0182	0.0316	0.0202	0.0142	0.0202	0.0147	0.0085	0.0246	0.0166	0.0092
RMSE	0.1349	0.1779	0.1420	0.1192	0.1423	0.1211	0.0920	0.1568	0.1289	0.0959
MAE	0.0182	0.0316	0.0202	0.0142	0.0202	0.0147	0.0085	0.0246	0.0166	0.0092
Panel B: CDS_{t-1}										
MSE	0.0201	0.0364	0.0215	0.0122	0.0200	0.0162	0.0117	0.0285	0.0163	0.0099
RMSE	0.1419	0.1907	0.1467	0.1105	0.1414	0.1274	0.1081	0.1688	0.1277	0.0995
MAE	0.0972	0.1277	0.0999	0.0821	0.0994	0.0913	0.0778	0.1174	0.0871	0.0710
Panel C: CDS_{t-1} and HV_{t-1}										
MSE	0.0137	0.0250	0.0166	0.0076	0.0140	0.0116	0.0062	0.0200	0.0122	0.0065
RMSE	0.1171	0.1582	0.1288	0.0870	0.1185	0.1077	0.0789	0.1415	0.1103	0.0803
MAE	0.0810	0.1103	0.0878	0.0653	0.0841	0.0758	0.0606	0.1011	0.0742	0.0589

Table 6. Weekly Out-of-sample performance of the model specifications for each of the maturity and moneyness bins (k=5 days)
Out-of-sample performance of the model specifications for each one of the implied volatility forecast. We do this estimation for all three models, and calculate the errors for nine different bins as well as all sample. The mean square error (MSE), the root mean squared prediction error (**RMSE**), and the mean absolute prediction error (**MAE**) are provided.
(i)The root mean squared prediction error in implied volatilities (RMSE) is the *square root of the average squared deviations of BS implied volatilities from the model's forecast implied volatilities*, averaged over the number of options traded. (ii) The mean absolute prediction error in implied volatilities (MAE) is the average of the *absolute differences between the BS implied volatility and the model's forecast implied volatility* across traded options.

Table 6. Out-of-sample performance of IV prediction: k= 5 days

		ITM			ATM			OTM		
	All sample	Short	Medium	Long	Short	Medium	Long	Short	Medium	Long
Panel A: HV_{t-1}										
MSE	0.0122	0.0235	0.0195	0.0054	0.0208	0.0154	0.0045	0.0234	0.0135	0.0044
RMSE	0.1104	0.1532	0.1398	0.0732	0.1443	0.1241	0.0669	0.1529	0.1160	0.0662
MAE	0.0809	0.1215	0.1005	0.0615	0.1009	0.0914	0.0619	0.1160	0.0761	0.0600
Panel B: CDS_{t-1}										
MSE	0.0141	0.0242	0.0238	0.0073	0.0153	0.0148	0.0058	0.0204	0.0153	0.0052
RMSE	0.1187	0.1555	0.1543	0.0857	0.1237	0.1215	0.0762	0.1428	0.1238	0.0721
MAE	0.0822	0.1165	0.1019	0.0638	0.0905	0.0853	0.0605	0.1009	0.0806	0.0539
Panel C: CDS_{t-1} and HV_{t-1}										
MSE	0.0141	0.0255	0.0209	0.0064	0.0278	0.0171	0.0060	0.0258	0.0138	0.0062
RMSE	0.1186	0.1596	0.1445	0.0797	0.1668	0.1306	0.0778	0.1607	0.1176	0.0785
MAE	0.0752	0.1196	0.0935	0.0558	0.0971	0.0831	0.0527	0.1163	0.0745	0.0511

Table 7. Impact of Financial Crisis: Out-of-sample performance during crisis vs. non-crisis periods

The table presents results of out-of sample performance of Implied volatility forecast for two subperiods of Crisis (2007-2009), and Non-crisis (2002-2006) periods, on the set of key lagged variables. We show the **MAE (Mean Absolute Error)** for each of these regressions and also test for the difference between the two set of residuals.
The results show that during the crisis period the out of sample performance power consistently and significantly drops, as measured by a larger value of MAE.
Also, across all four models, the model with lagged CDS, lagged IV and lagged HV, possesses the smallest level of MAE.

Lagged CDS = previous observation day's CDS for the issuer firm.
Lagged IV = previous observation day's IV for the same Option ID.
Lagged HV = previous observation day's HV for the issuer firm.

	(1)	(2)	(3)
Forecasted Var: IV	Non Crisis	Crisis	Test of difference
Model 1: IV_t on $\{ IV_{t-1}\}$	0.0356	0.0508	0.0152***
Model 2: IV_t on $\{ IV_{t-1}, HV_{t-1}\}$	0.0300	0.0422	0.0121***
Model 3: IV_t on $\{ IV_{t-1}, CDS_{t-1}\}$	0.0328	0.0487	0.0159***
Model 4: IV_t on $\{ IV_{t-1}, HV_{t-1}, CDS_{t-1}\}$	0.0286	0.0414	0.0129***

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8. Impact of Industry: Out-of-sample performance for Financial vs. NonFinancial firms

The table presents results of out-of sample performance of Implied volatility forecast for two subsections of financial and non-financial firms, on the set of key lagged variables. We show the **MAE (Mean Absolute Error)** for each of these regressions and also test for the difference between the two set of residuals.
The results show that the difference between MAEs are statistically significant for 3 out of 4 models; However, it is interesting to note that the economic magnitude of the difference is minor between financial and non-financial firms.
Consistent with prior results, across all four models, the forecasting model that includes lagged CDS, lagged IV and lagged HV, possesses the smallest level of MAE and outperforms the rest of the models.

Lagged CDS = previous observation day's CDS for the issuer firm.
Lagged IV = previous observation day's IV for the same Option ID.
Lagged HV = previous observation day's HV for the issuer firm.

	(1)	(2)	(3)
Forecasted Var: IV	Non-Financial	Financial	Test of difference
Model 1: IV_t on $\{ IV_{t-1}\}$	0.0408	0.0460	0.0052***
Model 2: IV_t on $\{ IV_{t-1}, HV_{t-1}\}$	0.0348	0.0359	0.0009
Model 3: IV_t on $\{ IV_{t-1}, CDS_{t-1}\}$	0.0386	0.0431	0.0045***
Model 4: IV_t on $\{ IV_{t-1}, HV_{t-1}, CDS_{t-1}\}$	0.0335	0.0358	0.0022***

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Chapter 5

CONTRIBUTIONS AND CONCLUSIONS

The Bear Stearns hedge funds failure in August 2007, and later collapse of Lehman Brothers in September 2008, one of the largest bankruptcies in U.S. history, played a major role in the unfolding of the recent global financial crisis and magnitude and impact of default risk in capital markets. The proliferation and popularity of credit risk transfer instruments, create additional quality data and measures in contemporary research in the field of default risk. This dissertation fits into the very same framework and sheds light on some key issues surrounding default risk in capital markets.

This dissertation covers three essays in an attempt to further investigate major issues surrounding default risk in capital markets and the relationship with the 2008 financial crisis by focusing on important segments of capital markets including investments portfolio (hedge funds), derivatives (equity options market), fixed income (credit default swaps), and equity market.

The first essay "The role of Leverage in Hedge Funds Failure" investigates the role of financial leverage, including the use of margins and derivative products, in the hedge funds failure during the 2008 financial crisis. Motivated by failure of the two Bear Stearns hedge funds at the beginning of the financial crisis in 2007, this paper examines why some hedge funds failed during and after the recent financial crisis, and why some also survived. Leverage is defined in three ways, as (a) debt/equity ratio, (b) the use of margins, and (c) the use of derivative products.

The research uses a 15-year panel dataset of 17,202 failed and survived hedge funds from the Lipper TASS Hedge Fund database. The empirical analysis, using probit regression to estimate the likelihood of failure, shows that during the financial crisis period, financial leverage is more significant in increasing the probability of failure, whereas financial leverage becomes insignificant in explaining the probability of hedge fund failure during non-crisis periods after controlling for fund structure, size, incentive fees, prior performance, and off-shore registration. The results are consistent with Ang et al. (2011) who analyze the cyclical leverage for financial institutions and hedge funds and find that financial leverage decreases during financial crisis period because the funds sell some assets to meet their margin requirements and that forces some funds into liquidation. Further analysis shows that some hedge funds which follow specific styles such as "Emerging Markets", "Equity Market", "Long/Short Equity Hedge", and "Multi-strategy", which have higher than average betas are also more likely to fail during the financial crisis.

The second essay "Does default risk impact Equity Options?" explores the impact of default risk on equity option pricing. The impact is studied in detail by empirically examining to what extent the firm-specific default risk matters in pricing individual equity options. Since credit default swaps (CDS) are similar to put options in that both offer a hedging tool and an effective protection against downside risk, we use CDS spread as credit risk proxy to investigate the effects of default risk on put pricing.

Recent financial crisis showed that for many financial firms equity options experienced high implied volatility (IV) when the underlying CDS spreads went up. By examining an exhaustive sample of US-listed firms with both CDS and put options data available over the period from 2002 to 2010, and studying the primary determinants of option IVs cross-sectionally and over time, the findings show that default risk is a significant factor in the prices of equity options. Moreover, the impact of default risk remains significant after controlling for firm-specific and macroeconomic factors. This study relates to recent literature that explains how default risk can get injected from the fixed income market to the equity options market and why default risk is important in the pricing of equity option and implied volatility.

The third essay "Forecasting Option-Implied Volatility using credit risk" investigates the topic of forecasting option implied volatility which is of interest to option market participants, who routinely formulate volatility and option price forecasts for trading and hedging purposes. As shown in essay two, credit risk matters for option pricing since options are valued on firms with significant trading liquidity, yet subject to default risk, similar to liquidity risk. If credit risk matters for option prices, this essay particularly explores whether better out-of-sample forecasts for option implied volatility (IV) can be developed using lagged credit risk measures. Various time-series forecasts of daily, weekly, and monthly for option implied volatility show that inclusion of default risk as measured by credit default swap (CDS) can significantly improve out-of-sample performance, measured through decreased mean squared error (MSE) as well as smaller root mean squared error (RMSE).

Overall, the three essays covered in this dissertation contribute to the current literature by investigating recent issues of default risk and showing the findings in each of the tested segments of capital markets.