



City Research Online

City St George's, University of London

Citation: Fei, Fei (2013). Essays on Quantitative Risk Management.
(Unpublished Doctoral thesis, City University London)

This is the unspecified version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/2949/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Essays on Quantitative Risk Management



Fei Fei

Faculty of Finance
Cass Business School
City University London

A thesis submitted for the degree of
Doctor of Philosophy

May 2013

Contents

List of Figures	iii
List of Tables	iv
Acknowledgements	v
Declaration	vii
Abstract	viii
1 Introduction	1
1.1 The Challenges of Risk Management	2
1.2 Motivation and Objective of the Thesis	4
1.2.1 Credit rating migration and the business cycle	4
1.2.2 The dynamics of default risk dependency with equity markets and the role of regimes	6
1.2.3 Downside extreme risk and its impact on portfolio optimization	8
1.3 Layout of the Thesis	10
2 Credit Rating Migration Risk and Business Cycles	13
2.1 Introduction	13
2.2 Classical Credit Migration Estimators	17
2.2.1 Rating process and transition probabilities	17
2.2.2 Cohort or discrete multinomial approach	19
2.2.3 Hazard-rate or duration approach	20
2.3 Business Cycles and Credit Migration Risk	21

2.4	Empirical Analysis	25
2.4.1	Data description	25
2.4.2	Cohort and hazard-rate migration risk	29
2.4.3	Business cycle-adjusted migration risk	32
2.4.4	Empirical distribution of default rates	39
2.4.5	Out-of-sample forecast evaluation	43
2.4.6	Economic relevance: risk capital attribution	50
2.5	Conclusions	53
2.A	Standard & Poor's Rating Definitions	55
2.B	Three-Regime MMC Estimator	56
2.C	Sectoral MMC Migration Risk During Contraction	58
3	Regime-Switching Dependence between Credit Default Swap and Equity Markets	59
3.1	Introduction	59
3.2	Copula Methodology	65
3.2.1	Sklar's theorem and marginal processes	65
3.2.2	Dynamic copulas	69
3.2.3	Regime-switching dynamic copulas	70
3.2.4	Estimation of copula parameters	72
3.3	Data description	73
3.4	Empirical Results	77
3.4.1	In-sample fit of static, dynamic and regime-switching copulas .	77
3.4.2	Out-of-sample copula forecasts for risk management	87
3.5	Conclusion	95
3.A	Merton's Structural Model and Extensions	98
3.B	Dependence Measures	100
3.C	Static Copulas	101
3.D	Log-likelihood Functions of Static Copulas	104
3.E	Estimation of Regime-Switching Copula	106
3.F	Timeline of late 2000s Crises	107
3.G	Copula VaR Simulation	111

4	An Extreme Risk Management Framework for Portfolio optimization	112
4.1	Introduction	112
4.2	GARCH-EVT-Copula Modeling of Portfolio Returns	117
4.2.1	ARMA-GARCH filtering of asset returns	118
4.2.2	Extreme Value Theory	120
4.2.3	Dynamic Copula	123
4.3	Portfolio optimization Methods	124
4.4	Data and Empirical Results	126
4.4.1	Data description and preliminary statistics	126
4.4.2	Joint probability density function	130
4.4.3	Performance of portfolio optimization strategies	138
4.4.3.1	One-step-ahead forecasts using the entire sample . .	138
4.4.3.2	Performance of realized portfolios with optimal weights	142
4.5	Conclusion	145
4.A	Coherent Risk Measures	147
4.B	Multivariate Elliptical Copula	147
4.C	Monte Carlo simulation for 1-day-ahead returns	148
5	Conclusions and Suggestions for Future Research	150
5.1	Concluding Remarks	150
5.1.1	Business-cycle adjusted credit migration for calculating capital requirement	151
5.1.2	Regime switching copula for characterizing CDS-equity dependence	152
5.1.3	Downside extreme risk for portfolio optimization	153
5.2	Further Research	154
	References	157

List of Figures

2.1	Business-Cycle Dynamics	23
2.2	Sectoral Breakdown of Bond Issues	26
2.3	Upgrades, Downgrades, Defaults and Overall Economic Activity . . .	29
2.4	Term Structure of Default Probabilities for <i>CCC</i> bonds	35
3.1	Daily Time Series Plots of CDS and Equity Indices	75
3.2	Smoothed Probability of High Dependence Regime	83
3.3	Rank Correlation Estimates	85
3.4	Tail Dependence Estimates	86
4.1	Sectoral iTraxx Europe Indices	128
4.2	QQ Plot of Normal Distribution	129
4.3	Correlograms of ARMA-GARCH filtered returns	131
4.4	QQ Plot of Generalized Error Distribution	134
4.5	Generalized Pareto Distribution Fit at Tails	136
4.6	Optimal Portfolio Efficient Frontier at 5% Confidence Level	139
4.7	Portfolio Optimal Frontier: Confidence Level of CVaR	140
4.8	Optimal Portfolio Efficient Frontier at 5% Confidence Level: Distribu- tion Comparison	141

List of Tables

2.1	Summary Statistics of US Corporate Ratings	28
2.2	One-Year Rating Transition Risk	31
2.3	Term Structure of Default Probabilities	33
2.4	One-Year Corporate Default Risk Estimates from Existing Studies . .	37
2.5	Rating Transition Matrices Conditional on the Economic State	38
2.6	Default Rates for CCC Bonds per Sector	39
2.7	Empirical Distribution of One-Year Default Risk Estimates	41
2.8	Out-of-Sample Forecast Errors	49
2.9	Economic Capital Attribution for One-Year Risk Horizon	52
3.1	Descriptive Statistics of Daily Returns	76
3.2	Estimation Results for Marginal Models	78
3.3	Goodness-of-Fit Measures for Competing Copula Models	80
3.4	Estimation Results for Static, Dynamic and RS Student's t Copula .	81
3.5	Value-at-Risk Portfolios of Stoxx and CDS Indices (Student's t copula)	93
3.6	Value-at-Risk Portfolios of Stoxx and CDS Indices (Gumbel Copula) .	94
4.1	Summary Statistics of Daily Logarithmic Return	129
4.2	Univariate Model Selection	132
4.3	Estimated Parameters for ARMA-GARCH-GED model	133
4.4	Estimates of EVT Model of Upper and Lower Tail Behavior	135
4.5	Estimation Result of Copula Models	137
4.6	Statistics of Realized Optimal Portfolio Returns	144

Acknowledgements

First and foremost, I would like to take this opportunity to express my greatest and most sincere gratitude to my supervisors, Prof. Ana-Maria Fuertes and Dr. Elena Kalotychou, for allowing me the freedom to pursue my own research interests. Without their great patience, invaluable guidance, constructive criticism and continuous encouragement throughout these years this thesis would not have come into existence. It has been my great honor to work with them.

I would like to thank seminar participants at Cass Business School, London, and participants at the 2010 *Society for Computational Economics* 16th International Conference on Computing in Economics and Finance, Limassol, the 2011 5th *CSDA International Conference* on Computational and Financial Econometrics, London. I appreciate Prof. Andrew Stark, the editor of the *Journal of Business Finance and Accounting*, accepting my paper “Credit Rating Migration Risk and Business Cycles” as published in 2012.

I would like to thank to my colleagues and friends, Dr. Ka Kei Chan and Dr. Yile Wu for helpful discussions, comments and helps. Special thanks is due to Cookie, the most adorable stray cat, who brought me endless fun during the summer in 2010.

This thesis is dedicated to my parents, Heyi Fei and Shisheng Hu, my parents-in-law, Baokuan Zhang and Zhiying Zhang, and my wife, Qian Zhang, for their undoubtable support, limitless understanding, unfailing encouragement and unconditional love.

Declaration

The author of this thesis grants the power of discretion to the City University Library to allow this thesis to be copied in whole or in part without further reference to him. This permission covers only single copies made for study or research purposes, subject to normal conditions of acknowledgments.

Abstract

The costly lessons from global crisis in the past decade reinforce the importance as well as challenges of risk management. This thesis explores several core concepts of quantitative risk management and provides further insight.

We start with rating migration risk and propose a Mixture of Markov Chains (MMC) model to account for stochastic business cycle effects in credit rating migration risk. The model shows superior in-sample estimation and out-of-sample prediction than its rivals. Compared with the naive approach the economic application suggests banks with MMC estimator will increase capital requirement in economic expansion and free up capital during recession hence it is aligned with Basel III macroprudential initiative by reducing the recession-vs-expansion gap in capital buffers.

Subsequently we move to the key concept of dependence by investigating the importance of dynamic linkages between credit and equity markets. We propose a flexible regime-switching copula model to explore the dynamics of dependence and possible structure breaks with special consideration on tail dependence. The study reveals a high-dependence regime that coincides with the recent financial crisis. The backtesting results acknowledge the new model's superiority on out-of-sample VaR forecasting over purely dynamic or static copula. It can serve to emphasise the relevance for risk management of appropriately modeling complex dependence structures.

Finally we discuss the risk measures and how they affect the portfolio optimisation. We contend that more successful portfolio management can be achieved by combining extreme value analysis to describe downside tail risk and dynamic copulas to model nonlinear dependence structures. Conditional Value-at-Risk is adopted as pertinent measure of downside tail risk for portfolio optimisation. Using both realised portfolio returns and a set of out-of-sample Monte Carlo experiments, our novel portfolio strategy is confronted with the de facto mean-variance approach. The results suggest that the MV approach produces suboptimal portfolios or a less desirable risk-return tradeoff.

Introduction

Over the past two decades, financial markets have witnessed several large scale catastrophes. In the aftermath of the Asian crisis in 1997 and the Russian sovereign debt crisis in 1998, the bursting of Dot-Com bubble during 2000-2001 caused the world economy to plunge into a recession and the global markets, especially the internet sector, to shrink with losses in excess of \$5 trillion. A few years later, the burst of the American housing bubble led to the subprime mortgage crisis during 2007 to 2008, which spread quickly and globally. In September 15, 2008, the shock collapse of Lehman Brothers - the fourth largest investment bank in the USA with assets worth \$600 billion - became the symbolic start of the most dramatic financial disaster to unfold since the Great Depression. The world economy was once again dragged into deep recession. The effects of the ensuing European sovereign debt crisis still persist and cast doubt over the prospect of global economic recovery.

Each of the aforementioned financial crises impinged heavily on the economy: stock markets collapsed, firms went bankrupt or were taken over, and the unemployment rate increased. Although there are ongoing debates on causes of crisis, people have widely attributed them to various aspects such as a loosely regulated financial sector, over-complicated financial products, poor public monetary policies, global-

ization and failure of risk controls. Against this backdrop the the need for better financial risk management practices has emerged and has created new impetus for the the quest for effective risk measurement techniques and stress testing procedures.

1.1 The Challenges of Risk Management

Whereas specific definitions of risk might vary, Charette (1990) claims that three characteristics must be satisfied: firstly, there must exist a loss associated with a certain situation; secondly, there must be some uncertainty with respect to the eventual outcome and finally, some choice or decision of how to deal with the uncertainty and potential for loss. The first two conditions are measurable, which in turn define the risk as the quantifiable likelihood of suffering a loss. The third condition implies some actions that can be deployed when a risk emerges. To accommodate such features, Kloman (1990) defines the general essence of risk management as:

Risk management is a discipline for living with the possibility that future events may cause adverse effects.

For financial institutions who make profits from taking risks, risk management is the code activity and is a central part of their day-to-day practices. It is the art of making the tactical and strategic decisions to control risks. With the help of advanced techniques and financial instruments risk management mitigates misfortune and exploits desirable future opportunities. Therefore, it manages the direction and the extent of exposure of risk and adjusts our risk-taking behavior in a sensible way. Risk measurement is indispensable to support the management of risk. It is the task of communicating and quantifying risk. It is obvious that some quantitative tools are critical aids for supporting good risk management.

Although the second Basel Accord (Basel II) issued in 2004 by the Basel Committee on Banking Supervision, a regulatory body under the Bank of International Settlements (BIS), broadly classifies financial risks into three major categories: market risk, the risk from movements in market prices; credit risk, the risk of not receiving

contractual repayments often due to obligor's default and operational risk, the risk of losses from inadequate business operation, the boundaries of the three types of risks are not always clearly distinguished and one risk is often propagated from the others. Under the Basel II framework there is a mandatory capital conservation charge for banks in order to create a safeguard buffer for their exposures sensitive to credit risk, market risk and operational risk. Banks, especially the more sophisticated institutions, are allowed to opt for their own risk assessment approaches, such as internal value-at-risk (VaR) models for market risk and internal-rating-based (IRB) approaches for credit risk, in calculating capital requirements. The looming Basel III framework aims to response to the deficiencies in financial regulation revealed by the late-2000s financial crisis by further strengthening bank capital requirements. It is therefore expected that there will be an increasing demand for more rigorous quantitative risk modeling and stress testing tools across financial institutions in order to maintain financial stability of the banking sector.

The calculation of risk capital charges requires the use of the profit-and-loss (P&L) distribution, which is the distribution describing uncertainty of the changes in value of a bank's portfolio over a specific time period. An appropriate measure of risk also has to be determined. An obvious yet important challenge in quantitative risk management is presented by the multivariate nature of risk. We are generally interested in some form of aggregate risk that depends on high-dimensional vectors of underlying risk factors such as individual asset values in market risk or credit spreads and counterparty default indicators in credit risk. A particular concern over our multivariate modeling is the phenomenon of dependence between extreme outcomes, when many risk factors experience simultaneous adverse moves. How to accurately describe these individual risk factors and, more importantly, their dependence structure is a crucial aspect of better risk assessment. Another challenge is the need to address unexpected extreme outcomes rather than expected average outcomes. New quantitative risk management techniques therefore are required to go beyond the classical normal model and attempt to capture the related phenomena of heavy tails and extreme value clustering. Furthermore, in risk management, we are mainly concerned about

the probability of large losses and hence with the upper tail of the loss distribution. However, the return distribution is quite likely to behave asymmetrically at the tails, that is, the negative extreme return joint distribution could exhibit different stylized facts than the opposite tail. Thus, the central concern in modern risk management is to have a measure of risk which can address the aforementioned properties of downside (extreme) risk, yet is easy to understand and utilize.

1.2 Motivation and Objective of the Thesis

This thesis attempts to explore a range of quantitative risk management aspects like credit rating migration, dependence structure between asset returns, extreme risk measures and to provide insights on their impact on value-at-risk analysis, capital allocation and portfolio optimization. The thesis is divided into three chapters focusing on the aforementioned risk management topics. In this section, we summarize the motivation and objectives of each chapter.

1.2.1 Credit rating migration and the business cycle

The Basel II Accord (2004) permitted banks and other financial institutions to use their own internal models for calculating economic capital. The accurate estimation of probability of defaults (PDs) and the probability of an obligor's change of creditworthiness together with the loss given default (LGD) and exposure at default (EAD) play a vital role in calculating the risk of banks' loan books as well as determining bank risk capital allocation. In Chapter 2 we delve into the issue of estimation of the credit rating migration matrix, which gives the probabilities that an obligor's creditworthiness will change (upgrade, downgrade or even default) over a specific time horizon, with primary focus the accurate estimation of PDs.

The classical discrete-time cohort method, proposed by Jarrow et al. (1997), and its continuous-time extension the hazard rate model, developed by Lando and Skodeberg (2002), are both based on two assumptions that rating migrations are time-homogeneous and follow a Markov chain process. However such strong assumptions

have been questioned recently. Altman and Kao (1992) and Carty and Fons (1994) firstly document non-Markovian behavior like ratings momentum or ratings drift. Those “upgrade then upgrade” and “downgrade then downgrade” feature of rating migrations are also reported by Kavvathas (2001), Carling et al. (2002), Couderc and Renault (2005), and Fuertes and Kalotychou (2007).

The time-homogeneity assumption has been questioned mainly because of the presence of business cycles. Credit rating migrations are likely to be correlated with macroeconomic factors and, in particular, it seems natural to expect an increase in portfolio credit risk during economic downturns. Bangia et al. (2002) find that variation of migration probabilities is higher in contractions than they are on average, supporting the existence of two distinct economic regimes. Hamilton and Cantor (2004) document that within a rating class the hazard rates for default and other migrations vary considerably over time. Duffie and Singleton (2003), Duffie et al. (2007) document stronger correlation between macroeconomic and firm characteristics with the rating migrations for speculative grade issuers when economic activity shrinks.

The literature on transition matrix estimation that incorporates business cycle information is very limited (see Trueck (2008) for a review). Nickell et al. (2000) propose an ordered probit model of credit migration probabilities conditional on exogenous variables so as to investigate the dependence of ratings transition probabilities on different borrowers characteristics (e.g. industry or country domicile) and on the current state of the economy. Their model is criticized in Wei (2003) because of requiring large samples to obtain reliable estimates. Bangia et al. (2002) account for business cycle effects in the discrete cohort-type migration estimator by dividing the ratings data into two subsets in order to match the business cycle and estimate separately expansion and contraction matrices (hereafter, this is referred to as naive business cycle cohort estimator). Lando and Skodeberg (2002) deploy the non-parametric Aalen-Johansen estimator that extends the cohort estimator to infinitely short-time intervals so as to allow for general time heterogeneity in the underlying continuous Markov process. They note that, apart from its high computational costs, this approach does not yield significantly different estimates from the time-homogeneous

hazard rate estimator for large data sets.

Against this backdrop, Chapter 2 aims to provide further insight into the credit rating migration risk. We advocate a Mixture of Markov Chains (MMC) estimator of rating migration risk which explicitly recognizes the stochastic business cycle. The MMC estimator is compared with the naive cyclical counterpart and with classical through-the-cycle estimators in three frameworks: statistical, forward-looking, and economic. The analysis has important implications for the ongoing financial regulatory reforms and is in the light of the Basel III initiatives to improve the financial sector's resilience to stress scenarios which calls for a reassessment of banks' credit risk models and, particularly, of their dependence on business cycles. .

1.2.2 The dynamics of default risk dependency with equity markets and the role of regimes

The global credit derivative market has grown dramatically over a short period of time and expanded in volume from \$300 billion in 1998 to \$25.9 trillion at the end of 2011 according to statistics from the International Swaps and Derivatives Association (ISDA). This is primarily attributed to the development of credit default swap (CDS) indices. CDS is a financial contract to transfer risk from the protraction buyer to the seller when credit events, usually default, happen. The spread of a CDS is a percentage of the notional annual amount the protection buyer must pay the protection seller over the length of the contract thus it is considered, together with the credit rating, as a good indicator of credit risk.

In the past decades a large body of literature has been devoted to investigating the factors driving the credit spreads. Among them, the so-called structural modeling approach has become popular among industry practitioners and academic researchers. The model, inspired from the work of Black and Scholes (1973) and Merton (1974), assumes that the value of the firm follows a stochastic process and default occurs when its value falls below its debts. The empirical literature argues in favor of the existence of a negative link between credit spreads and firm's equity value, see Longstaff

et al. (2003), Norden and Weber (2009), and positive relationship with firm's equity volatility, see Madan and Unal (2000), Blanco et al. (2005), and Zhang et al. (2009). However, all the aforementioned studies focus on single-name (at firm level) CDS spreads, which are susceptible to liquidity criticisms. As responses, Bystrom (2006), Alexander and Kaeck (2008) and Fung et al. (2008) study the comovement of CDS index with the underlying equity markets. However their models are based on the OLS regression framework implying constant linear correlation and lacking consideration of the dependence of extreme events. The number of studies on the linkages between CDS indices and the equity market is very limited hence Chapter 3 of the thesis is an attempt to enrich the literature in this field by providing a realistic description of such features.

There is mounting evidence that the dependence between financial returns is non-linear and time-varying (see Scheinkman and Lebaron (1989), Hsieh (1989), Longin and Solnik (1995), Brooks (1996) and Ang and Chen (2002)), which has triggered a quest for more flexible dependence structures that go beyond the linear correlation. In particular, the comovement between assets tends to be stronger when financial markets are more volatile and the dependence does not disappear even if returns take extreme values. These properties of asymmetric dependence and tail dependence invalidate the use of the conventional Pearson's linear correlation and the multivariate normal distribution for the same reason. The recent financial crisis has also provided vivid evidences of the aforementioned dependence features. Empirically Hull et al. (2004b) and Zhang et al. (2009) both document the nonlinear relation between CDS spreads and equity volatility in the Merton's framework. Increased linkages between the credit and equity markets are an important source of systemic risk and for this reason there is ongoing pressure from regulators, investors and rating agencies on financial institutions to build appropriate models that allow measurement of the different risks faced. A crucial step in this process is modeling the temporal dynamics in the dependency between the returns of different asset classes.

To this end, we resort to the copula models as a better alternative to measuring correlations. Copulas naturally have the capacity and enough flexibility to describe

the non-normality in asset return distributions, e.g. nonlinear, asymmetric relationships, and dependence in tails, which makes them a suitable vehicle for modeling multivariate correlations. There is a bulk of empirical evidence supporting the use of copula models as a more realistic characterization of the dependence between financial returns (see Embrechts et al. (2002), Mashal and Zeevi (2002)). However these models are typically static and not capable to effectively capture the dynamic aspects of dependence between financial risks.

In Chapter 3 we follow Patton (2006) who extends the copula models into a dynamic framework and we propose a regime-switching dynamic copula for modeling to reveal the sudden changes in dependence between the European credit default swap and the underlying equity market from a “low” (normal periods) to a “high” (crisis periods) dependence regime. The regime-switching behavior is also explored at the tails of the joint distribution. The regime-switching effects in the credit-equity relationship are then assessed from a risk management viewpoint by means of Value-at-Risk (VaR) backtesting. Thus this work may benefit the empirical implication for better risk management.

1.2.3 Downside extreme risk and its impact on portfolio optimization

In the past decade we have witnessed a sequence of crises, 2007-2008 credit crunch, 2008-2010 automotive industry crisis, European sovereign-debt crisis since 2009 (see Appendix 3.F for details), which have left tremendous marks on global financial markets. Such disasters have become more frequent in the recent years than ever. The question emanating is whether current risk management frameworks can provide adequate responses to such effect. If not, how can they be improved to better capture the downside risk associated with these dangerous occurrences? These concerns may challenge conventional ideas about portfolio construction.

Portfolio optimization is the art as well as science of balancing return against risk. It is one of the most important tasks for risk managers, how they efficiently

allocate capital to assets in a portfolio to achieve maximum returns given specific risks or equivalently to minimize risks for a given set of returns is always on the top of their agenda. The conventional approach in modern portfolio theory, pioneered by the path-breaking work of Markowitz (1952), measures risks as the portfolio standard deviation, a measure of how much the returns deviate from the expected or mean return, hence, it is also called *mean variance* (MV) framework.

The core assumption of this framework is that future asset returns are normally distributed and thus, a portfolio is multivariate normal, or generally speaking the multivariate elliptical. The statistical properties of normality posit that the mean vector together with covariance matrix of returns fully describe their joint behavior and variance alone determines the weight of the tails. However, there is considerable evidence that asset return distributions are leptokurtic, fat-tailed, asymmetric and far from normal. The thin tails of the normal distribution, implying a very low likelihood for situations where all asset classes are falling significantly in market crash scenarios, may underestimate the true risk. In order to address this problem the literature has turned towards introducing heavy-tailed distributions. However, one problem of these “one-piece” models is that they use the same distribution function to characterize both the tail and central behaviors which implicitly assumes the tails have the same properties of the central part of the return distribution. Another issue is that fitting a parametric distribution to data sometimes results in a model that agrees well with the data in high density regions but poorly in areas of low density due to fewer observations in the tails. Taking advantage of the latest advances in the statistical area of Extreme Value Theory (EVT), in Chapter 4, we attempt to separate the estimation of the tails of asset return distribution that characterizes the occurrence of extreme events from the central part of the distribution. To this end, in Chapter 4 we resort the dynamic copula models in order to provide a more realistic description of the dependence evolution and evaluate the economic gains of such approaches from a portfolio construction viewpoint.

Another drawback of the conventional MV framework is that it takes the portfolio standard deviation as the measure of portfolio risk. It implicitly assumes that in-

vestors treat the desirable upside returns as same as the undesirable downside losses, which is not true as investors are typically more concerned about potential losses rather than gains especially during periods of crisis. The recent literature, Artzner et al. (1999) and Rockafellar and Uryasev (2002), has proposed an alternative risk measure, Conditional Value-at-Risk (CVaR), which has the attractive feature of focusing on the extent of downside losses in situations where things do get bad and hence it is better aligned with the modern risk management outlook.

In Chapter 4 , we deploy EVT and dynamic copula models to better describe the joint density function while considering non-normality, fat-tails and the clustering of extreme events. We evaluate the properties of alternative portfolio optimization strategies based on CVaR against the traditional mean-variance model. The performance of the competing models is examined by a set of Monte Carlo out-of-sample simulation experiments in order to address that the optimal portfolios hinge on the assumptions made for the distribution of asset returns and their dependent structure, as well as the choice of risk measure to focus on.

1.3 Layout of the Thesis

The remaining part of this thesis consists of three empirical essays regarding quantitative aspects of risk management. We start from the rating migration model in Chapter 2, which describes the default probability or the probability of changes of obligors' creditworthiness and is recognized as one of the core areas of risk modeling. Then we move to another key area of portfolio risk management to study the dependence structure of asset returns. Financial investors, especially institutional investors, with large exposures on different assets have to face risks from both the devaluation of assets and that of the whole industry or economy. This raises questions about how to accurately characterize the correlation among asset classes. To this end, we resort to copula models in Chapter 3 to examine the dependence structure between the CDS and equity markets. Following up, we examine the impact of appropriately measuring downside extreme risk in a portfolio optimization framework in Chapter

4.

Chapter 2 compares the two classical models known as cohort and hazard rate approaches for estimations of credit migration probabilities and discusses their flaws on dealing with cyclicity. A Mixture of Markov Chains (MMC) model is then proposed to estimate rating migration risk that explicitly recognizes the stochastic evolution of the economy between phases of the business cycle. The merits of the model are gauged by both in-sample estimations and out-of-sample forecast evaluation. A real-time leading indicator of business cycles is employed to generate out-of-sample predictions of credit migration risk. A number of forecast error metrics is deployed to compare the forecast accuracy of MMC model against more basic approaches. Finally, we investigate the economic relevance of capturing business cycles via our MMC estimator by an application to economic capital attribution.

Chapter 3 investigates the dynamics of dependence between the credit market and the corresponding equity market and the role of regime shifts in dependence during crisis periods. To this end, we firstly introduce the concept of copula and the its properties and estimation methods. An ARMA-GARCH skew- t model is adopted to filter out stylized facts of financial index returns like leptokurtosis, asymmetry, heavy-tails, auto-correlation and heteroskedasticity. Then a regime-switching (RS) dynamic copula model is proposed to capture the dependence structure between iTraxx CDS spreads and equity returns for marketwide and sectoral pairs. In-sample statistical analysis reveals the relevance of the regimes at the center and at the tails of the joint distribution. The suggested high dependence regimes coincide with the credit crunch and the European sovereign debt crisis. Finally the superiority of the regime-switching approach over purely dynamic or static copula is also underlined via out-of-sample VaR backtesting relevant for risk management.

Chapter 4 focuses on modeling downside extreme risk and investigates the performance of portfolio optimization strategies when different measures of risk are considered. We start from ARMA-GARCH models with different error terms for modeling univariate time series. Tail distributions are characterized by Extreme Value Theory. A dynamic t copula model is then employed to provide a realistic character-

ization of the correlation evolution, especially the dependence at the tails. A set of Monte Carlo out-of-sample simulation are deployed to compare the performance of return-risk allocation under the mean-conditional value-at-risk strategy against the traditional mean-variance approach.

Chapter 5 concludes the thesis by providing an overview of our research and a summary of the results. Finally, the last section of the chapter suggests potential directions for future research.

Credit Rating Migration Risk and Business Cycles

2.1 Introduction

The Basel II Accord, issued in 2004 by the Basel Committee on Banking Supervision (BCBS), permitted banks to use internal models to calculate capital requirements. The old Basel I (1988) rules endorse a standardized risk-weighting approach to determine the capital needed for backing different assets. Under the internal ratings-based approach, encouraged by Basel II and Basel III, banks can use their “in-house” risk models to predict for each asset in their portfolio the corresponding probability of default (PD), exposure at default (EAD) and loss given default (LGD). The resulting numbers are then plugged into a formula that assigns a risk weight.¹

Historical credit ratings² are the main inputs to classical estimation of credit migration probabilities, of which default risk is a measure of special interest. The late 2000s global financial crisis (GFC), which started with the collapse of the US housing bubble, has prompted a lot of skepticism over agency ratings for not being

¹The expected likelihood that firms or sovereign borrowers will default in its contractual payments is a crucial input not only to the assignment of economic capital but also to other risk management applications such as portfolio risk analysis and the pricing of bonds or credit derivatives.

²A rating is an indication of creditworthiness.

informative enough on the credit quality of structured debt obligations and for lacking timeliness. A clear instance is the bipartisan investigation into the origins of the crisis that led to the Levin and Coburn (2011) report of the U.S. Senate stating that:

the crisis was not a natural disaster, but the result of high risk, complex financial products; undisclosed conflicts of interest; and the failure of regulators, the credit rating agencies, and the market itself to rein in the excesses of Wall Street.

Recent studies suggest though that rating actions convey new information to the marketplace and trigger capital restructuring. Hill and Faff (2010) and Faff et al. (2007) document significant causality running from credit rating events to, respectively, international equity markets and fund flows. In a model to explain credit default swap (CDS) prices, Batta (2011) finds that ratings are significant determinants of corporate credit risk and impound relevant accounting variables such as earnings and leverage. Kisgen (2009) documents that downgraded firms reduce leverage by about 1.5%-2% in the year following the rating change.

The recent GFC has also served as a stark reminder to the marketplace of the crucial role of systematic stress testing of financial institutions' portfolios, particularly, their lending books. In response to the regulatory deficiencies thus revealed, Basel III is seeking to achieve the broader macroprudential goal of protecting the banking sector from periods of excess credit growth by requesting longer horizon default probabilities, downturn loss-given-default measures and improved calibration of risk models (see Basel Committee on Banking Supervision (2011a)). The calibration of models that translate credit ratings and/or market data into default probabilities has direct implications for the determination of the risk-adjusted capital (i.e., core Tier 1 capital ratio) that banks need to hold to back their loans and safeguard solvency. In this chapter we propose an approach to estimate credit rating migration risk that controls for the business-cycle evolution during the relevant time horizon in order to ensure adequate capital buffers both in good and bad times. The approach allows the default risk associated with a given credit rating to change as the economy moves

through different points in the business cycle.

There is a body of research linking portfolio credit risk with macroeconomic factors showing, for instance, that default risk tends to increase during economic downturns. Figlewski et al. (2012) document that unemployment and real GDP growth are strongly correlated with default risk. Stefanescu et al. (2009) develop a Bayesian credit score model to capture the typical internal credit rating system of most major banks and show that macroeconomic covariates such as the S&P500 returns have good explanatory power. Thus point-in-time methodologies that account for business cycles should provide more realistic credit risk measures than through-the-cycle models that smooth out transitory fluctuations (perceived as random noise) in economic fundamentals.³

Some attempts have been made in the *classical* credit risk measurement literature to incorporate cyclicity. Nickell et al. (2000) subdivide the historical ratings into those observed in ‘normal’, ‘peak’ and ‘trough’ regimes according to real GDP growth and deploys a discrete time (cohort) estimator of migration risk separately on each subsample. This naive approach to accommodating cyclicity is also deployed in several other studies although conditioning instead upon NBER-delineated expansion and recession phases: Jafry and Schuermann (2004) and Bangia et al. (2002) in a credit-portfolio stress testing context, Hanson and Schuermann (2006) in their statistical comparison of continuous time (hazard) versus discrete migration estimators, and Frydman and Schuermann (2008) in a Markov mixture estimator that allows for firm heterogeneity. In effect, the naive estimator implicitly assumes that the current economic conditions prevail throughout the prediction time horizon of interest. In contrast, the estimator proposed in this chapter relaxes this assumption by allowing the economy to evolve randomly between states of the business cycle during the risk horizon. The importance of doing so is implicitly underlined by the Basel III requirement to use long term data horizons of at least one year to estimate probabilities of

³Rating agencies tend to adopt a through-the-cycle approach seeking to provide stable risk assessments over the life of at least one business cycle. The survey by Cantor et al. (2007) reflects that whereas bond issuers tend to favor ratings stability, plan sponsors and fund managers/trustees have a stronger preference for point-in-time accuracy.

default.

This chapter contributes to the credit risk modeling literature as follows. It demonstrates the advantages of using a Mixture of Markov Chains (MMC) model to estimate rating migration risk that explicitly recognizes the stochastic evolution of the economy between phases of the business cycle. This is the first study that comprehensively evaluates a MMC estimator against a naive counterpart that conditions deterministically on the current economic conditions by assuming that they prevail throughout the prediction time horizon, and against classical through-the-cycle estimators. It also departs from the aforementioned studies, which consider economic dynamics in credit risk modeling, by assessing the estimators in a strictly forward-looking sense. More specifically, we exploit a real-time leading indicator of business cycles based on a principal components methodology to generate out-of-sample predictions of credit migration risk. Overall the horse race of estimators is carried out in three complementary ways: a simulation analysis of their in-sample statistical properties with specific emphasis on accuracy, an out-of-sample forecasting exercise based on a range of (a)symmetric loss functions, and an economic Value-at-Risk (VaR) analysis drawing upon the CreditRisk+ risk management framework of Credit Suisse First Boston (1997).

To preview our key results, the in-sample analysis reveals efficiency gains in default risk measures derived from the MMC model and more so during economic contraction due to the paucity of ratings data. Acknowledging the risk that economic conditions randomly evolve over the risk horizon is shown to improve the accuracy of out-of-sample default probability predictions. This is clearly revealed through novel asymmetric loss functions that attach a relatively high penalty to the under(over)prediction of down(up)grade risk. Such accuracy gains of the MMC estimator vis-à-vis the naive counterpart increase with the length of the forecast horizon. Both business cycle estimators make an important difference for economic capital attribution since they imply more prudent capital buffers than through-the-cycle estimators during contraction. However, the naive cyclical approach suggests relatively high (low) contraction (expansion) risk-capital holdings. The MMC estimator of default risk implies about

17% less capital in downturns than the naive estimator, which could be channeled into lending to stimulate the real economy, while the suggested capital during expansions exceeds by 6% that from the naive estimator. The MMC estimator notably reduces the expansion-versus-recession gap in risk capital relative to the naive counterpart and can be cast as an efficient way to perform stress testing. This is an important property because exaggerated cyclicalities can fuel ‘irrational exuberance’ and deepen recessions by making lending too capital intensive, which is one of the main criticisms of Basel II (see Gordy and Howells (2006)). Thus, relative to its competitors, the MMC approach prescribes capital build-up in good times that banks can draw upon in bad times and so it is more aligned with the Basel III macroprudential initiative to dampen the procyclical transmission of risk and promote countercyclical capital buffers.

This chapter is organized as follows. Section 2.2 below reviews the classical migration estimators. Section 2.3 presents the MMC hazard rate approach. The data and empirical results are outlined in Section 2.4, and Section 2.5 concludes.

2.2 Classical Credit Migration Estimators

2.2.1 Rating process and transition probabilities

A *credit rating* is a financial indicator of an obligor’s level of creditworthiness. Most firms issuing publicly traded debt are rated at least by one of the three major rating agencies, Moody’s, Standard & Poor’s (S&P), and Fitch Ratings.⁴ Let the credit rating of a firm at time t be denoted $R(t) \in S = \{1, 2, \dots, K\}$ where S is the rating space with 1 and $K - 1$ representing, respectively, the best and worst credit quality; K represents default. For instance, the coarse S&P’s rating system (AAA, AA, A, BBB, BB, B, CCC) together with the default state D imply $K = 8$. The rating definitions

⁴Several papers examine the rating actions of the three agencies. For instance, Hill and Faff (2010) offer evidence of stronger stock market reaction to S&P’s rating changes than to those of the other two agencies. Mählmann (2009) proposes a structural self-selection model to explain the decision by firms to seek a third ‘optional’ rating from Fitch based on their borrowing costs. Using survey research methods, Baker and Mansi (2002) find that issuers and institutional investors perceive corporate bond ratings from Moody’s and S&P as more accurate than those of Fitch.

provided by the agencies are qualitative in nature which makes their mapping onto specific quantitative risk measures crucial; Appendix 2.A provides summary S&P definitions.

The goal is to estimate the transition or migration probabilities over horizon $[t, t + \Delta t]$ denoted

$$\mathbf{Q}(\Delta t) \equiv \mathbf{Q}(t, t + \Delta t) = \begin{pmatrix} q_{11}(\Delta t) & q_{12}(\Delta t) & \cdots & q_{1K}(\Delta t) \\ q_{21}(\Delta t) & q_{22}(\Delta t) & \cdots & q_{2K}(\Delta t) \\ \vdots & \vdots & \ddots & \vdots \\ q_{K1}(\Delta t) & q_{K2}(\Delta t) & \cdots & q_{KK}(\Delta t) \end{pmatrix}, \quad (2.1)$$

where $q_{ij}(\Delta t) \equiv P[R(t + \Delta t) = j | R(t) = i] \geq 0, \forall i, j \in S$, is the chance that an obligor rated i at time t is assigned rating j at $t + \Delta t$, hence $\sum_{j=1}^K q_{ij}(\Delta t) = 1$. The K^{th} column contains the probabilities of default (PDs). Since default is treated as an “absorbing” state, $q_{Ki}(\Delta t) = 0$ and $q_{KK}(\Delta t) = 1$, implying that $R(t)$ will settle to the default steady-state in the limit as $\Delta t \rightarrow \infty$.

As noted earlier, two assumptions typically underlie the rating migration process:

1. *Markovian behavior.* The probability of transition to a future state j only depends on the current state and is independent of the rating history. Formally,

$$P[R(t + \Delta t) = j | R(t), R(t - 1), R(t - 2), \dots] = P[R(t + \Delta t) = j | R(t)] \quad \forall j \in S.$$

and thus the current rating contains all relevant information to predict future ratings.⁵

2. *Time homogeneity.* For a given risk horizon of interest, Δt , the transition probabilities are time-constant meaning that they only depend on Δt and thus

⁵Rating momentum or drift, first documented in Altman and Kao (1992) and Carty and Fons (1994), is a prime counterexample to Markovian dynamics. In the same vein, Lando and Skodeberg (2002) and Fuertes and Kalotychou (2007) show that a downgraded issuer is prone to further subsequent downgrades.

there is a family of matrices

$$\mathbf{Q}(\Delta t) \equiv \mathbf{Q}(t, t + \Delta t) = \mathbf{Q}(t - k, t - k + \Delta t) \quad \forall k.$$

A stochastic process satisfying (1) is called a Markovian process. If its state space is countable, it is called a *Markov chain* process. So a time-homogeneous Markov chain satisfies

$$P[R(t + \Delta t) = j | R(t) = i] = P[R(t - k + \Delta t) = j | R(t - k) = i].$$

Hence, the n -year migration matrix is given by the n^{th} power of the annual one, $\mathbf{Q}(n) = \mathbf{Q}(1)^n$.

Next we outline the two classical migration risk estimators known as *cohort* and *hazard rate* approaches. Both build on the Markov and time-homogeneity assumptions but differ mainly in that they are formulated in a discrete- and continuous-time framework, respectively.

2.2.2 Cohort or discrete multinomial approach

Let $N_i(t)$ denote the number of firms that start year t at rating i and $N_{ij}(t, t + 1)$ the subset of them that have migrated to rating j by year-beginning $t + 1$. Let the migration frequency be denoted $\hat{q}_{ij}(t, t + 1) = \frac{N_{ij}(t, t + 1)}{N_i(t)}$ in years $t = 1, 2, \dots, T$. Assuming a time-homogeneous Markov rating process, the maximum likelihood (ML) estimator of the one-year credit migration risk is

$$\hat{q}_{ij} \equiv \hat{q}_{ij}(1) = \sum_{t=1}^T w_i(t) \hat{q}_{ij}(t, t + 1) = \frac{\sum_{t=1}^T N_{ij}(t, t + 1)}{\sum_{t=1}^T N_i(t)} = \frac{N_{ij}}{N_i}, \quad (2.2)$$

where $w_i(t) = N_i(t) / \sum_{t=1}^T N_i(t)$ are yearly weights. Thus $\hat{q}_{ij}$ can be simply computed as the total number of annual migrations from i to j divided by the total number of obligors that were in grade i at the start of any sample year. Time-homogeneity is called upon to obtain the n -year cohort migration matrix as $\hat{\mathbf{Q}}(n) = \hat{\mathbf{Q}}^n$ where $\hat{\mathbf{Q}}$ is

obtained using Eq.(2.2). The major rating agencies routinely publish these migration estimates for horizons $\Delta t = 1, 2, \dots, 10$ years.

One weakness of the cohort approach is that it neglects within-year rating transitions and rating duration information. For instance, if firm X is rated *AA* on 01/01/1992, *A* on 03/09/1992 and *AA* on 31/12/1992, the cohort estimator would consider its rating unchanged in year 1992. Thus this approach is very sensitive to data sparsity which is especially typical of transitions from top ratings to default. Another issue is that the discrete annual (or n -year) horizon of the baseline cohort estimator is too rigid as more flexibility is needed to price payoffs occurring at arbitrary points in time. These shortcomings call for a continuous-type credit migration estimator.

2.2.3 Hazard-rate or duration approach

Let the *transition intensity* or *generator* matrix of a continuous Markov chain be denoted $\mathbf{\Lambda}_t \equiv \{\lambda_{ij}(t)\}_{ij \in S}$ where $\lambda_{ij}(t)_{i \neq j}$ is the *hazard rate* function or intensity representing the instantaneous transition rate at time t ; the diagonal entries are given by $\lambda_{ii}(t) \equiv \lambda_i(t) = -\sum_{j \neq i} \lambda_{ij}(t)$. The probability of migration from rating i to j over an arbitrarily small time horizon τ is given by

$$P[R(t + \tau) = j \mid R(t) = i] = \lambda_{ij}(t)\tau \quad \text{for } i \neq j$$

Under time-homogeneity it follows that $\lambda_{ij}(t) = \lambda_{ij}$ and its ML estimator is

$$\hat{\lambda}_{ij} = \frac{N_{ij}(0, T)}{\int_0^T Y_i(s) ds} \quad \text{for } i \neq j \quad (2.3)$$

where $N_{ij}(0, T)$ is the total number of transitions from i to j observed over the sample period, $Y_i(s)$ is the number of firms rated i at time s and thus $D_i \equiv \int_0^T Y_i(s) ds$ gives the rating *duration* or total time spent in state i by all sampled obligors. The

transition risk matrix estimator is

$$\hat{Q}(\Delta t) \equiv \hat{Q}(t + \Delta t) = e^{(\Delta t)\hat{\Lambda}} \quad (2.4)$$

where the matrix exponential can be obtained via a Taylor-series expansion, $e^{(\Delta t)\hat{\Lambda}} = \sum_{k=0}^{\infty} \frac{[(\Delta t)\hat{\Lambda}]^k}{k!}$.

One appealing feature of Eq.(2.4) is its flexibility to measure credit migration risk over any arbitrary time horizon, Δt . Moreover, it exploits rating transitions that occur at any point in the sample as well as rating duration information. For illustration, with reference to the example at the end of Section 2.2.2, the hazard-rate estimator exploits the intermediate within-year migrations to/from A through the transition intensities as expressed in Eq.(2.3). Furthermore, suppose that the transition $AAA \rightarrow D$ is not observed but there are transitions $AAA \rightarrow BB$ and $BB \rightarrow D$. By contrast with the cohort measure, the hazard-rate PD estimate for AAA -rated bonds is non-zero, albeit small, in line with economy theory since no bond is default free.⁶

2.3 Business Cycles and Credit Migration Risk

In this section we present a Mixture of Markov Chains (MMC) estimator of credit migration risk that accounts for the current (time t) economic phase while acknowledging the stochastic business-cycle evolution over the migration horizon of interest $(t, t + \Delta t)$. It mixes two time-homogeneous Markov chains, one that models the ratings process and another that models the business cycle process. We focus the exposition and ensuing analysis on two phases, expansion (E) and contraction (C), but the estimator can be readily extended to more phases; Appendix 2.B presents the

⁶This is an important aspect implicitly recognized, for instance, by the S&P definitions which state "For example, a corporate bond that is rated AA is viewed by S&P as having a higher credit quality than a corporate bond with a BBB rating. But the AA rating isn't a guarantee that it will not default, only that, in our opinion, it is less likely to default than the BBB bond." (Source: www.standardandpoors.com).

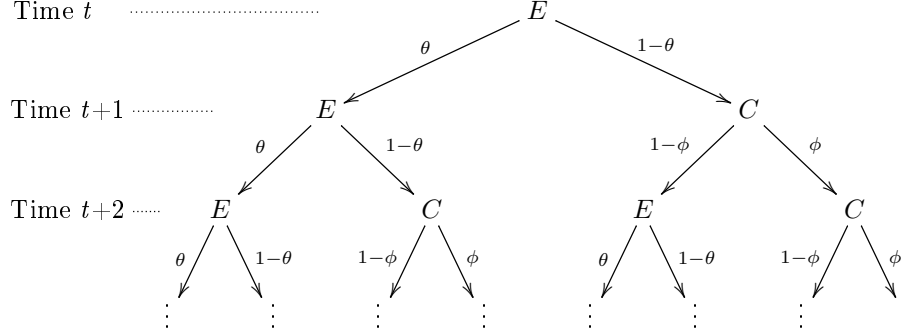
3-regime case.⁷ Let the following matrix characterize the *economic* evolution over a one-period horizon

$$\mathbb{S}(1) \equiv \mathbb{S}(t, t+1) = \begin{pmatrix} \theta & 1 - \theta \\ 1 - \phi & \phi \end{pmatrix}, \quad (2.5)$$

where θ is the probability that the next phase at $t+1$ is an expansion conditional on the time t phase being an expansion, and $(1 - \theta)$ is the probability of switching to a contraction. The parameters in $\mathbb{S}$ are treated as exogenous and obtained in the spirit of the *hazard-rate* approach via the corresponding transition intensity matrix $\mathbf{\Lambda}_{\mathbb{S}}$. For instance, if the baseline 1-period window is one quarter, the transition intensity $\hat{\lambda}_{E,C}$ can be computed as the number of $E \rightarrow C$ transitions over the entire sample divided by the total duration of expansion phases in months; likewise for $\hat{\lambda}_{C,E}$. The quarterly regime-switching matrix Eq.(2.5), simply called $\mathbb{S}$, can then be estimated as $\hat{\mathbb{S}} = e^{3\hat{\mathbf{\Lambda}}_{\mathbb{S}}}$. Figure 2.3 characterizes for current (time t) expansion the subsequent unfolding of the economy, or business-cycle dynamics, as a binomial tree. Within each economic phase, expansion or contraction, the *ratings* evolution follows another time-homogeneous Markov chain characterized, respectively, by the conditional transition matrices $\mathbf{Q}_E$ and $\mathbf{Q}_C$. These two matrices are estimated by splitting the observed ratings into two subsets according to whether they have been observed during expansion or contraction, and deploying the hazard-rate estimator Eq.(2.4) separately on each.

⁷One could define three regimes as expansion, ‘mild’ recession and ‘severe’ recession, where mild and severe are qualified in terms of the time-length or severity measured, say, as the percentage decrease in real GDP growth. Or one might identify ‘above’, ‘below’ and ‘full’ capacity phases using the Hodrick-Prescott filtered real GDP.

Figure 2.1: Business-Cycle Dynamics



Notes: This figure characterizes for current expansion at time t the subsequent unfolding of the economy, or business-cycle dynamics, as a binomial tree. E denotes expansion and C denotes contraction. The parameter $\theta(\phi)$ denotes the probability that the next phase at $t+j$ is an expansion (contraction) conditional on the time $t+j-1$ phase being an expansion (contraction).

The *mixture* process is characterized by the following one-period transition matrix

$$\mathbb{M}(t, t+1) = \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{pmatrix} \equiv \begin{pmatrix} & \begin{array}{c} \text{Expansion} \\ \text{AAA} \quad \dots \quad \text{CCC} \quad \text{D} \end{array} & \begin{array}{c} \text{Contraction} \\ \text{AAA} \quad \dots \quad \text{CCC} \quad \text{D} \end{array} \\ \begin{array}{c} \text{Expansion} \\ \text{AAA} \\ \vdots \\ \text{CCC} \\ \text{D} \end{array} & \begin{array}{c} \theta \mathbf{Q}_E \\ \\ \\ \end{array} & \begin{array}{c} (1-\theta) \mathbf{Q}_C \\ \\ \\ \end{array} \\ \begin{array}{c} \text{Contraction} \\ \text{AAA} \\ \vdots \\ \text{CCC} \\ \text{D} \end{array} & \begin{array}{c} (1-\phi) \mathbf{Q}_E \\ \\ \\ \end{array} & \begin{array}{c} \phi \mathbf{Q}_C \\ \\ \\ \end{array} \end{pmatrix} \quad (2.6)$$

all entries in Eq.(2.6), simply called $\mathbb{M} \equiv \mathbb{M}(1)$, are non-negative and each row sums to 1. The probability that an i graded obligor at current time t , pertaining to economic expansion, is downgraded to j at $t+1$, also in expansion, is given by $\theta q_{i,j,E}$. Given current expansion, $E(t)$, the one-period MMC transition matrix, $\mathbb{Q}_E(1) \equiv \mathbb{Q}_E(t, t+1)$, is calculated by adding the transition probabilities associated to the two possible

business-cycle pathways, $E(t) \rightarrow E(t+1)$ and $E(t) \rightarrow C(t+1)$, as

$$\mathbb{Q}_E(1) \equiv \mathbf{M}_1 + \mathbf{M}_2 = \mathbb{X} = \theta \mathbf{Q}_E + (1 - \theta) \mathbf{Q}_C. \quad (2.7)$$

For a two-period horizon, as Figure 2.3 illustrates, there are four possible business-cycle pathways which result in the MMC rating transition matrix for current expansion

$$\begin{aligned} \mathbb{Q}_E(2) \equiv \mathbf{M}_1 \mathbb{X} + \mathbf{M}_2 \mathbb{Y} = \mathbb{F} \mathbb{L}' = \\ \theta \mathbf{Q}_E \theta \mathbf{Q}_E + \theta \mathbf{Q}_E (1 - \theta) \mathbf{Q}_C + (1 - \theta) \mathbf{Q}_C (1 - \phi) \mathbf{Q}_E + (1 - \theta) \mathbf{Q}_C \phi \mathbf{Q}_C \end{aligned} \quad (2.8)$$

where $\mathbb{Y} \equiv \mathbb{Q}_C(1) \equiv \mathbf{M}_3 + \mathbf{M}_4 = (1 - \phi) \mathbf{Q}_E + \phi \mathbf{Q}_C$ is the one-period MMC transition matrix conditional on current contraction, $\mathbb{F} \equiv (\mathbf{M}_1 \ \mathbf{M}_2)$, and $\mathbb{L}'$ is the transpose of $(\mathbb{X} \ \mathbb{Y})$. Over an n -period horizon, $\Delta t \equiv n$, the ratings evolution will have 2^n unique pathways but the MMC migration matrix for current expansion can be fashioned in an elegant closed-form solution as

$$\mathbb{Q}_E(n) \equiv \mathbb{F} \mathbb{M}^{n-2} \mathbb{L}'. \quad (2.9)$$

The MMC migration matrix for current contraction, denoted $\mathbb{Q}_C(n)$, is derived similarly by defining instead $\mathbb{F} \equiv (\mathbf{M}_3 \ \mathbf{M}_4)$ in Eq.(2.9). The MMC approach implicitly addresses to some extent two other issues beyond time-heterogeneity. One is cross-sectional dependence in default rates across obligors due to common macroeconomic conditions or systematic risk; the other is serial dependence (e.g., ratings drift or non-Markovian behavior) induced by the cyclical behavior of latent macro factors. A simpler way to embed business cycle effects in credit migration risk models is to assume a deterministic business-cycle evolution, namely, that the economy remains in the same (as the current or time t) phase throughout the horizon of interest t to $t+n$.⁸ This naive estimator emerges as a particular case of the MMC hazard-rate es-

⁸Cyclicity is accounted for in this manner by Bangia et al. (2002) using the cohort estimator, Jafry and Schuermann (2004) using the hazard-rate estimator, Hanson and Schuermann (2006) using both estimators and Frydman and Schuermann (2008) using a Markov mixture estimator designed

timator Eq.(2.9) by conceptualizing $\mathbb{S}$ as the identity matrix, i.e. assuming $\theta = \phi = 1$ in Eq.(2.5) which gives $\mathbb{Q}_E(1) = \mathbf{Q}_E$ and $\mathbb{Q}_E(n) = \mathbf{Q}_E^n$; likewise for $\mathbb{Q}_C(n)$.

2.4 Empirical Analysis

2.4.1 Data description

Our sample contains 7,514 US corporate bond rating histories over the 26-year period 01/01/1981 to 31/12/2006 from the S&P CreditPro 7.7 database. Like Altman and Kao (1992) *inter alios*, we track the ratings of individual bond issues in order to increase the number of observed migrations.⁹ Among the debt issues sampled, mainly from large corporations, 2,218 are industrials, 1,677 utilities and 1,494 financials. Figure 2.2 illustrates that the representation of financials over our sample period initially experienced a gradual increase at the expense of industrials and utilities but the relative proportions have remained roughly steady for over half of the sample.¹⁰

The S&P rating scale comprises 22 fine categories but they are typically shrunk into a coarse rating system which excludes the +/- modifiers and has become the industry standard (8 rating categories plus default).¹¹ Each bond issue has experienced more than 3 rating transitions over the sample period, a total of 1,166 bonds finally default and there are 4,202 *Not Rated* (*NR*) assignments in total.¹² Transitions to *NR* may be due to debt expiration, calling of the debt or failure to pay the required fee to the rating agency. Following Altman and Kao (1992), Carty and Fons (1994)

to capture firm heterogeneity.

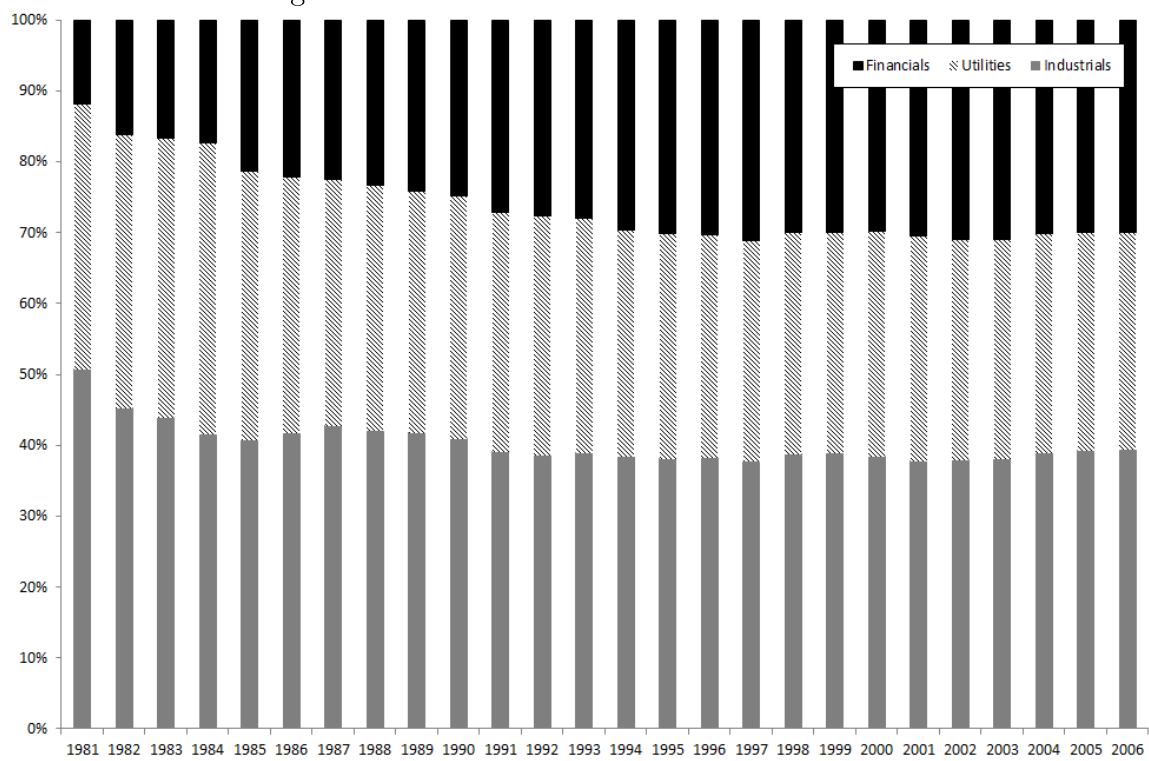
⁹S&P maps individual issue ratings into issuer ratings through the implied long-term senior unsecured rating.

¹⁰Ratings for sovereigns and municipals are not included. The *industrial* sector amalgamates aerospace, automotive, capital goods, metal/forest and building products, homebuilders, healthcare, chemicals, high technology, computers and office equipment firms. Energy and natural resources, transportation and telecommunication companies are included in the *utility* sector. The *financial* sector comprises financial institutions and insurance firms. Other sectors include consumer service (miscellaneous retailers) and leisure time/media firms.

¹¹The coarse classification implies, for instance, that CCC+, CCC, CCC-, CC and C ratings are grouped as CCC.

¹²Among the 3,933 issues whose rating is withdrawn (*NR*) at some point, 802 are re-rated, of which 285 finally default; 251 bond rating histories have at least 2 different NR episodes. Only 5 of the 1,166 issues that enter default are rated again but, since the default state is conceptualized as absorbing, we discard their post-default rating history.

Figure 2.2: Sectoral Breakdown of Bond Issues



Notes: This figure plots the relative proportion of industrial, utility and financial bond issues rated by Standard & Poor's on each year from 01/01/1981 to 31/12/2006.

and Frydman and Schuermann (2008) *inter alios*, we keep *NR* as another “rating” category in the sample. However, since rating migrations from/to *NR* do not provide any information about the obligor’s credit quality, they are not counted as upgrades or downgrades. By incorporating new issues in the sample and discarding existing ones after default, we allow the cross-section to vary over time. These considerations help us to increase the sample size for each transition. The ratings are allocated into two categories, respectively, those observed during economic expansion and contraction (or recession). Recession is conceptualized by the National Bureau of Economic Research (NBER) as a significant decline in US economic activity lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales.¹³ There are three official periods of recession in our sample as identified by the NBER dating: 1) the early 1980s recession linked to the hike in oil prices after the late 70s energy crisis, 2) the early 1990s recession characterized by decreases in industrial production and manufacturing-trade sales, 3) the early 2000s recession following the collapse of the dotcom bubble, the 9/11 terrorist attacks and accounting scandals like Enron. These three economic stress episodes represent overall merely 11% of quarters in our 26-year sample period.

Table 2.1 reports for the entire sample and for the NBER expansion and contraction subsamples: *i*) total rating assignments, *ii*) rating duration in quarters, and *iii*) proportion of up/downgrades. Most of the assignments are to the intermediate *A*, *BBB*, *BB* or *B* ratings. A much lower percentage of the speculative or non-investment grade (i.e., *BB* or below) rating assignments and durations pertain to financials than to industrials/utilities whereas the opposite is true for the top-quality ratings; this maybe because it is very difficult to keep banks operating when a low credit rating has damaged customers’ trust.¹⁴ Downgrades are more likely than upgrades particularly in contraction.

¹³The NBER Business Cycle Dating Committee has since 1978 delineated peak/trough months of economic activity. We adopt the first day of the peak/trough month as the business cycle turning point. See www.nber.com/cycles.

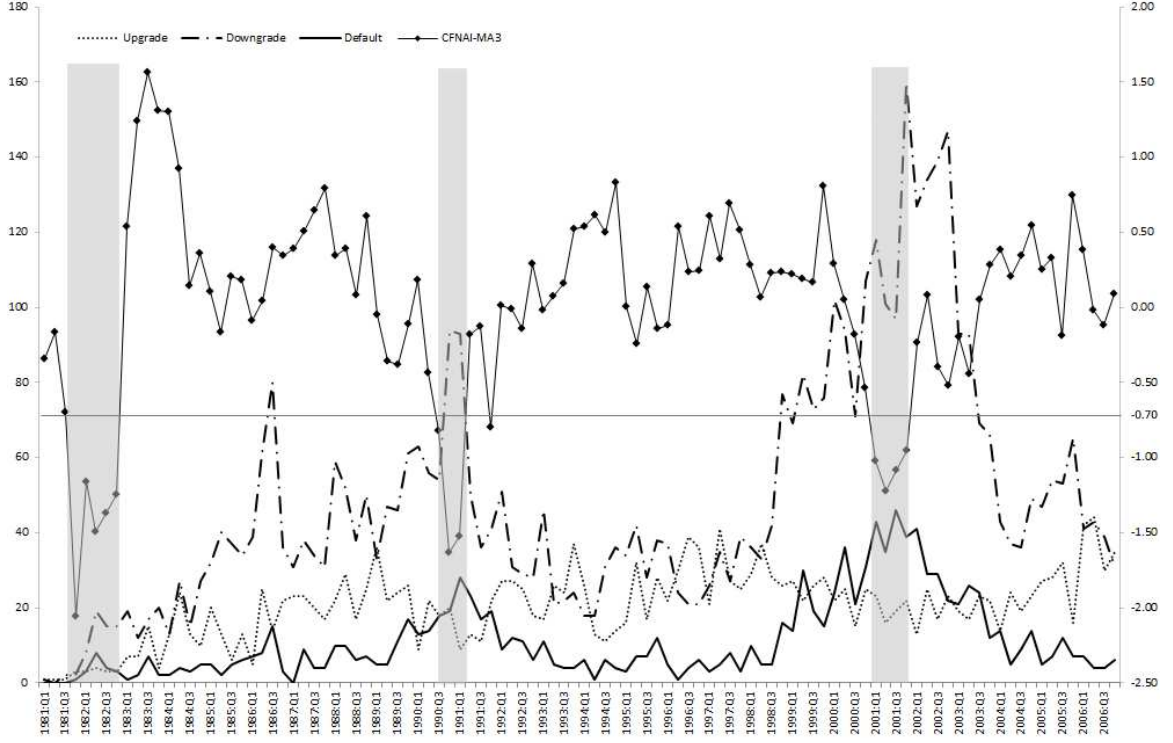
¹⁴For instance, 10.41% of the *AA* assignments correspond to industrials, 24.32% to utilities and 54.93% to financials. In contrast, 39.66% of the *B* assignments correspond to industrials, 17.83% to utilities and 5.41% to financials.

Table 2.1: Summary Statistics of US Corporate Ratings

State i	$\mathcal{E}_i$					$\mathcal{E}_i$				$\mathcal{C}_i$			
	N_i	$N_{i \rightarrow j, i \neq j}$	D_i	$N_{i,u} \%$	$N_{i,d} \%$	N_i^E	D_i^E	$N_{i,u}^E \%$	$N_{i,d}^E \%$	N_i^C	D_i^C	$N_{i,u}^C \%$	$N_{i,d}^C \%$
AAA	240	185	6,792.55	0	100.00	224	6,356.00	0	100.00	16	435.69	0	100.00
AA	1,480	825	23,592.24	6.42	93.58	1,346	21,924.52	6.80	93.20	134	1,664.96	3.28	96.72
A	3,777	1,749	52,324.63	20.12	79.88	3,409	48,794.15	20.91	79.09	368	3,523.14	11.11	88.89
BBB	4,469	2,067	49,441.55	39.89	60.11	4,076	46,035.30	41.53	58.47	393	3,397.91	21.70	78.30
BB	4,105	2,313	35,443.58	36.36	63.64	3,859	33,106.71	37.15	62.85	246	2,331.50	27.78	72.22
B	5,418	3,004	41,775.67	36.34	63.66	5,144	38,863.57	38.90	61.10	274	2,965.03	14.44	85.56
CCC	1,405	970	4,055.99	19.50	80.50	1,245	3,688.46	21.75	78.25	160	358.89	5.45	94.55
D	1,166	0	45,943.63	–	–	987	43,147.33	–	–	179	2,787.09	–	–
NR	4,202	802	143,291.98	–	–	3,941	133,925.84	–	–	261	9,287.68	–	–
Total	26,262	11,915	402,661.81	29.58	70.42	24,231	375,841.88	31.12	68.88	2,031	26,751.89	15.01	84.99

Notes: This table presents summary statistics for the sample of S&P's rating events from 01/01/1981 to 31/12/2006 employed in our subsequent analysis. N_i denotes the total assignments to state i that represents either a rating, default or NR. $N_{i \rightarrow j, i \neq j}$ is the total number of migrations from state i to state j . D_i is the sum of quarters spent in rating i across all obligors. $N_{i,u}$ and $N_{i,d}$ are, respectively, the proportion of upgrades and downgrades relative to total migrations from rating i . E and C denote, respectively, expansion and contraction phases according to NBER chronology. Full details are given in Section 2.4.1 of the chapter.

Figure 2.3: Upgrades, Downgrades, Defaults and Overall Economic Activity



Notes: This figure shows the quarterly number of S&P's corporate bond rating upgrades, downgrades and defaults alongside the real-time history of the Chicago Fed National Activity Index as a 3-month moving average (CFNAI-MA3). Shaded areas indicate official NBER-delineated contraction episodes.

Figure 2.3 reports the quarterly evolution of upgrades, downgrades, and defaults over the sample period. Shaded areas are NBER contraction quarters. The graph confirms business cycle time-heterogeneity in the rating migration process by illustrating that the number of downgrades (defaults) rises in contractions. The quarterly regime-switching matrix $\hat{S}$ estimated via transition intensities (with duration in months) has off-diagonal entries $1 - \theta = 0.0276$ and $1 - \phi = 0.241$. Hence, if the economy is currently in expansion, the probability that it enters contraction over the next quarter is 2.8% and the probability that it switches from contraction to expansion is 24.1%.

2.4.2 Cohort and hazard-rate migration risk

The finest sampling interval adopted for tracking the ratings evolution is one quarter so as to match the window length of the expansion/contraction switching probability

matrix $\mathbb{S}$. Hence, in the notation of Sections 2.2 and 2.3, the one-year horizon for the migration matrices corresponds to $\Delta t \equiv n = 4$, the two-year horizon to $\Delta t \equiv n = 8$ and so forth. Thus the cohort estimator Eq.(2.2) is deployed on a quarterly basis, giving $\hat{\mathbf{Q}}$, and the one-year migration matrix is computed as $\hat{\mathbf{Q}}(4) = \hat{\mathbf{Q}}^4$. In the hazard-rate framework, first, we deploy Eq.(2.3) to obtain the intensities or entries of $\hat{\mathbf{\Lambda}}$, defined as rating transitions per quarter, and the one-year transition matrix is given by $\hat{\mathbf{Q}}(4) = e^{4\hat{\mathbf{\Lambda}}}$. Table 2.2 reports the one-year cohort and hazard rate migration risk measures.¹⁵

As expected, both credit migration matrices are diagonally dominant implying relatively large ratings stability over a one-year horizon. However, the diagonal entries are smaller for *speculative grade* ratings than for *investment grade* ones, confirming that low ratings are more volatile. Unsurprisingly also, the default likelihood increases monotonically as credit quality deteriorates. Another commonality across the two matrices is that the immediate off-diagonal elements are generally larger for downgrades than upgrades, e.g. the cohort probability for a *BB* issuer to attain *BBB* next year is 5% whilst the chance of being downgraded to *B* is 7%. The estimates essentially confirm the stylized row monotonicity in rating migrations, i.e. the migration likelihood generally decreases the further away from the diagonal, reflecting the typical practice by S&P (and other rating agencies) of changing ratings in one-notch steps. These findings are in line with the evidence in Nickell et al. (2000), Bangia et al. (2002), Lando and Skodeberg (2002), and Fuertes and Kalotychou (2007).

The hazard-rate approach overcomes the cohort estimation pitfall of producing zero default risk measures for *AAA* bonds in the absence of *AAA* $\rightarrow$ *D* sample migrations. More specifically, the hazard-rate estimator suggests a non-zero, albeit small, PD at 1.5bp for *AAA* bonds. The estimated PD for *CCC* bonds offers also an interesting contrast: 28% (cohort) and 41% (hazard). On the other hand, the cohort approach overestimates (relative to the hazard-rate method) the default probabilities

¹⁵In order to identify an appropriate truncation lag k for the Taylor-series expansion $e^{(\Delta t)\hat{\mathbf{\Lambda}}} \equiv \sum_{k=0}^{\infty} \frac{[(\Delta t)\hat{\mathbf{\Lambda}}]^k}{k!}$ in (2.4), we follow Löffler and Posch (2007) and compute the sum of the squared elements of the k th summand. If this is smaller than 10^{-320} the series is truncated at that k , otherwise one more summand is added.

Table 2.2: One-Year Rating Transition Risk

	AAA	AA	A	BBB	BB	B	CCC	D	NR
Panel I: <i>Discrete Cohort estimator</i>									
Transition probabilities									
AAA	89.861	5.910	0.348	0.058	0.174	0.000	0.000	0.000	3.650
AA	0.551	86.748	7.604	0.635	0.050	0.167	0.033	0.017	4.195
A	0.053	1.542	87.686	5.551	0.481	0.173	0.038	0.075	4.400
BBB	0.008	0.150	3.616	85.164	4.177	0.671	0.118	0.237	5.859
BB	0.033	0.066	0.231	5.504	77.456	6.987	0.703	0.989	8.842
B	0.000	0.056	0.139	0.250	4.971	75.693	3.951	4.674	10.266
CCC	0.000	0.000	0.380	0.285	1.139	8.444	48.767	27.799	13.188
D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000
NR	0.019	0.061	0.147	0.270	0.275	0.318	0.011	0.591	98.308
Panel II: <i>Continuous Hazard-Rate estimator</i>									
Transition intensities									
AAA	-0.109	0.061	0.005	0.001	0.001	0.000	0.000	0.000	0.042
AA	0.006	-0.140	0.085	0.005	0.001	0.001	0.000	0.000	0.042
A	0.001	0.017	-0.134	0.064	0.004	0.002	0.000	0.000	0.048
BBB	0.000	0.002	0.040	-0.168	0.057	0.005	0.001	0.001	0.064
BB	0.000	0.001	0.003	0.057	-0.262	0.099	0.006	0.002	0.093
B	0.000	0.001	0.002	0.004	0.057	-0.288	0.086	0.023	0.117
CCC	0.000	0.000	0.005	0.007	0.011	0.131	-0.960	0.634	0.172
D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NR	0.000	0.001	0.002	0.003	0.004	0.005	0.000	0.007	-0.023
Transition probabilities									
AAA	89.669	5.376	0.651	0.093	0.110	0.017	0.001	0.015	4.067
AA	0.559	86.994	7.416	0.695	0.082	0.080	0.013	0.020	4.131
A	0.067	1.467	87.630	5.525	0.466	0.162	0.008	0.022	4.654
BBB	0.017	0.180	3.429	84.815	4.601	0.630	0.063	0.099	6.166
BB	0.030	0.078	0.365	4.650	77.335	7.617	0.573	0.520	8.831
B	0.002	0.054	0.174	0.446	4.369	75.565	4.725	3.874	10.789
CCC	0.002	0.011	0.327	0.477	0.871	7.226	38.563	40.980	11.541
D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000
NR	0.024	0.078	0.199	0.318	0.366	0.438	0.024	0.710	97.842

Notes: This table presents the rating transition probabilities in percentage points obtained from the discrete-time cohort estimator (Panel I) and continuous hazard-rate estimator (Panel II, second exhibit). The first exhibit of Panel II reports the transition intensity or generator matrix that represents the instantaneous rate of transition from one rating i to another rating j ($i \neq j$) expressed as number of rating transitions per quarter. Full details of the computations are given in Sections 2.2.2 and 2.2.3 of the chapter.

for middle ratings which is in line with the results in Jafry and Schuermann (2004), Hanson and Schuermann (2006) and Fuertes and Kalotychou (2007). For instance, the cohort PD estimate for BBB bonds is 23.7bp whereas the hazard-rate PD is 9.9bp. This may be because longer durations in the middle ‘stepping stone’ ratings (of which downgrade drift is largely responsible) reduce the transition intensities and, in turn, the hazard-rate transition probability estimates.

2.4.3 Business cycle-adjusted migration risk

We now compare the term structure of default probabilities implied by the naive and MMC business-cycle estimators. The results are set out in Table 2.3. To ease the comparison, Panel I reports also the estimates from the classical hazard-rate approach. Panels II and III pertain, respectively, to the naive expansion and contraction matrices $\hat{\mathbf{Q}}_E(n)$ and $\hat{\mathbf{Q}}_C(n)$ obtained as follows. The two intensity matrices $\hat{\Lambda}_E$ and $\hat{\Lambda}_C$ with rating duration in quarters are computed, respectively, by deploying (2.3) on the expansion and contraction rating subsamples. The naive business-cycle transition risk for 1-year horizon is then given by $\hat{\mathbf{Q}}_E(4) = e^{4\hat{\Lambda}_E}$ and $\hat{\mathbf{Q}}_C(4) = e^{4\hat{\Lambda}_C}$. Panels IV and V show, respectively, the MMC migration matrices $\hat{\mathbf{Q}}_E(n)$ and $\hat{\mathbf{Q}}_C(n)$ obtained using as inputs the quarterly matrices $\hat{\mathbf{S}}$, $\hat{\mathbf{Q}}_E \equiv \hat{\mathbf{Q}}_E(1) = e^{\hat{\Lambda}_E}$ and $\hat{\mathbf{Q}}_C \equiv \hat{\mathbf{Q}}_C(1) = e^{\hat{\Lambda}_C}$. Thus the 1-year PDs for current expansion in Panel IV are obtained from $\hat{\mathbf{Q}}_E(4)$, i.e. Eq.(2.9) is deployed for $n = 4$ quarters.

Both the naive and MMC business-cycle estimators yield much higher default risk in current contraction than in expansion. Intuitively, this means that a firm operating in a contracted economy at time t is more likely to default over the horizon $(t, t + \Delta t)$ than another similarly rated firm currently in expansion *ceteris paribus*. Moreover, accounting for business cycles increases the default risk estimates relative to the baseline hazard-rate approach, particularly, in contraction. For example, the 1-year PD for a *CCC* issuer is 40.913% according to the classical hazard-rate estimator, and increases to 54.988% (65.135%) with the MMC (naive) contraction estimator.

It is worth emphasizing that the term “naive” for the estimates shown in Panels II-III refers to the implicit assumption that the economy remains in the same time t

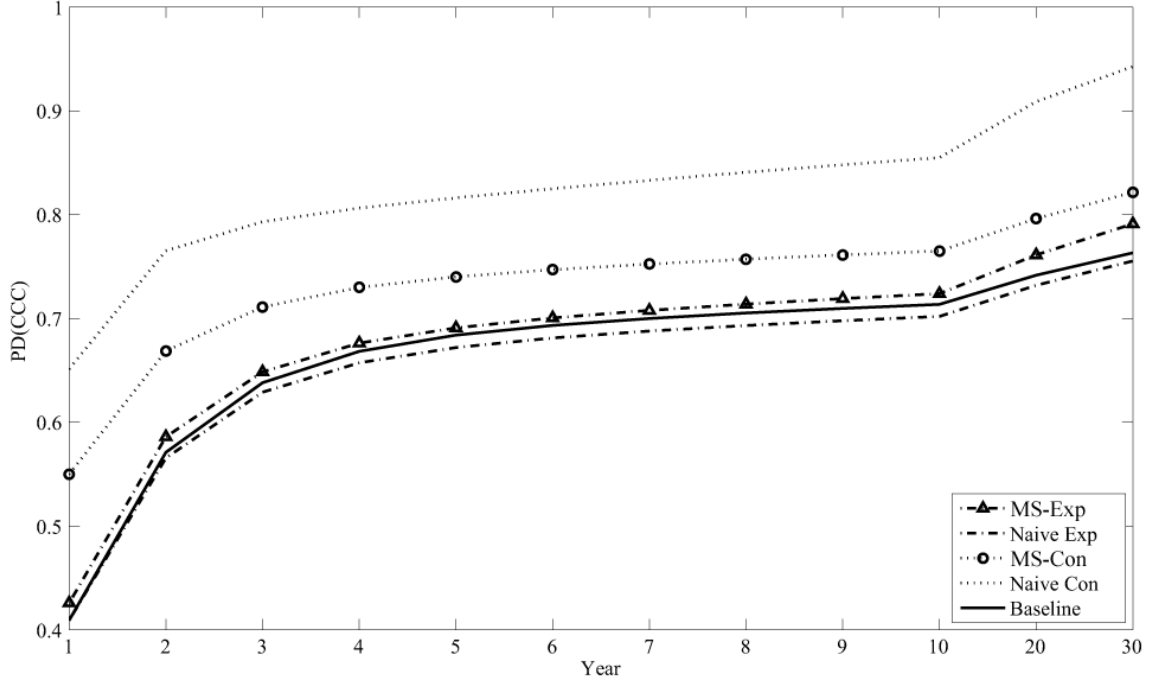
Table 2.3: Term Structure of Default Probabilities					
	1 Year	2 Year	3 Year	4 Year	5 Year
Panel I: <i>Baseline Hazard-Rate estimator</i>					
AAA	0.015	0.060	0.136	0.242	0.379
AA	0.020	0.079	0.174	0.304	0.470
A	0.022	0.091	0.211	0.383	0.606
BBB	0.099	0.300	0.604	1.006	1.494
BB	0.520	1.513	2.825	4.311	5.863
B	3.874	8.811	13.441	17.396	20.678
CCC	40.980	57.096	63.804	66.829	68.386
Panel II: <i>Naive Expansion estimator</i>					
AAA	0.017	0.069	0.155	0.276	0.432
AA	0.023	0.091	0.200	0.351	0.542
A	0.034	0.123	0.272	0.480	0.747
BBB	0.132	0.385	0.755	1.229	1.793
BB	0.666	1.828	3.300	4.922	6.586
B	4.314	9.545	14.301	18.287	21.547
CCC	40.903	56.554	62.886	65.719	67.196
Panel III: <i>Naive Contraction estimator</i>					
AAA	0.626	2.292	4.709	7.635	10.875
AA	1.506	3.898	6.883	10.224	13.741
A	1.938	4.736	8.044	11.611	15.268
BBB	2.617	6.191	10.138	14.139	18.044
BB	8.700	15.982	21.805	26.539	30.540
B	18.015	30.343	37.819	42.685	46.243
CCC	65.135	76.518	79.317	80.630	81.618
Panel IV: <i>MMC Expansion estimator</i>					
AAA	0.046	0.196	0.438	0.758	1.150
AA	0.108	0.358	0.710	1.143	1.651
A	0.145	0.463	0.906	1.450	2.081
BBB	0.282	0.843	1.598	2.494	3.495
BB	1.185	3.128	5.356	7.639	9.859
B	5.217	11.559	17.126	21.646	25.255
CCC	42.581	58.587	64.844	67.613	69.089
Panel V: <i>MMC Contraction estimator</i>					
AAA	0.297	0.713	1.133	1.585	2.089
AA	0.775	1.387	1.942	2.518	3.143
A	1.013	1.750	2.408	3.090	3.827
BBB	1.404	2.437	3.381	4.362	5.406
BB	5.018	7.668	9.844	11.898	13.856
B	11.770	18.389	23.129	26.821	29.764
CCC	54.988	66.863	71.105	72.984	74.010

Notes: This table reports estimates of the probability that a firm currently rated as indicated in the first column (from AAA to CCC) enters default over a time horizon $(t, t + \Delta_t)$ from 1 to 5 years. Panel I corresponds to the hazard-rate estimator. Panels II and IV correspond, respectively, to the naive and Mixture of Markov Chains (MMC) estimator conditional on current expansion. Panels III and V correspond, respectively, to the naive and MMC estimator conditional on current contraction. Full details of the computations are given in Section 2.2.3 and 2.3 of the chapter. All figures are in percentage points.

conditions, expansion or contraction, throughout the horizon $(t, t + \Delta t)$. In reality, the economy randomly alternates between business-cycle phases over time. By contrast, the MMC estimator not only takes into account the time t conditions but also acknowledges the stochastic evolution of the business cycle. A comparison of Panels II and IV reveals that the expansion PDs are too low (i.e., over-optimistic) for the naive estimator relative to the MMC estimator, and vice versa (i.e. over-conservative) in contraction. The 1-year naive PD estimates for *AAA* and *AA* when the economy is currently in expansion are, respectively, 1.7bp and 2.3bp and increase to 4.6bp and 10.8bp with the MMC estimator; the latter, but not the former, are in line with the 3bp floor imposed by the Basel II Accord on any PD estimate. Similarly for contraction, the 1-year PDs for *AAA* and *AA* are 29.7bp and 77.5bp according to the MMC approach, halving those implied by the naive one. These contrasting results from the two cyclical estimators, MMC and naive, are driven by the deterministic nature of the latter which rules out the chance of economic regime-switching over the horizon of interest; thus the naive PDs are, by construction, too high (low) for current contraction (expansion). The gap between the naive and MMC estimates is visibly larger in contraction than in expansion. Historically the US economy has stayed longer in expansion which, in turn, implies a smaller probability of switching from expansion to contraction than vice versa (respectively, $1 - \hat{\theta} = 2.8\%$ and $1 - \hat{\phi} = 24.1\%$, over the entire sample period). Since the MMC estimator collapses to the naive one for $1 - \theta = 1 - \phi = 0$ in Eq.(2.5), effectively, the gap between the two should be narrower when conditioning on current expansion.

Figure 2.4 plots default risk estimates for *CCC*-rated issues over a time horizon of up to 30 years. The discrepancy between the expansion and contraction PDs from the MMC approach gradually starts to narrow down for large time horizons. This is intuitively plausible since, as time passes, the effect of the current economic regime starts to dilute and the permanent component of default risk outweighs the temporary variations. This matches the evidence in Galbraith and Tkacz (2007) and Shcherbakova (2008) that the additional information content of models conditional on macroeconomic variables tends to decline as the forecast horizon increases. By contrast, the

Figure 2.4: Term Structure of Default Probabilities for *CCC* bonds



Notes: This figure shows the default probability for CCC bonds over a time horizon from 1 to 30 years. The five lines plotted correspond, as indicated in the legend, to the classical hazard-rate estimator that ignores business cyclicity (Baseline), the naive estimator conditional on current contraction (Naive Con) and expansion (Naive Exp), and the MMC estimator conditional on current contraction (MMC Con) and expansion (MMC Exp).

gap between the expansion and contraction PDs implied by the naive estimator does not dampen over time, in line with the fact that this simple approach implicitly assumes that the economy stays put (i.e., the current state prevails) throughout the estimation horizon. The upshot is that the expansion-versus-contraction PD differential implied by the naive estimator is inflated relative to that of the MMC estimator and more so as the time horizon lengthens.

Table 2.4 reports the 1-year default risk estimates for *B* and *CCC* corporate bonds from existing studies in the literature together with ours. Although the data sources and time spans differ, two common findings across studies that deploy the *classical* estimators (top panel) are: i) the default risk for *CCC*-rated issues suggested by cohort estimates is typically lower than that implied by hazard-rate ones, ii) for *B*-rated issues, however, the cohort estimator yields relatively higher PDs than

the hazard-rate estimator because of relatively longer durations (time spent) in this rating. Studies that incorporate *business cycle* information into the migration risk estimation (bottom panel), consistently suggest lower annual default risk in expansion than contraction.

Table 2.5 shows the entire 1-year migration matrix from the two business-cycle approaches, naive and MMC. The upper off-diagonal entries suggest that the chance of a downgrade (to either neighbor or extreme ratings) is higher if the current phase is contraction. The diagonal entries are larger for expansion than contraction in line with a rise in ratings volatility when economic conditions deteriorate. Overall, the contraction-versus-expansion gap in migration risk suggested by the naive estimator is magnified relative to that implied by the MMC estimator.

Finally, we examine the robustness of the results per sector. Table 2.6 presents the term structure of PDs for *CCC* bond issues from the classical hazard-rate model and the MMC model separately for industrials, utilities and financials. The default risk in the utility and financial sectors is lower than that in the overall economy and vice versa for the industrial sector. Sector by sector, the PDs appear again underestimated if the business cycle is ignored, particularly, in contraction. The long term structure of PDs for *CCC*-rated issues in each of the three sectors (unreported, to preserve space) is qualitatively similar to that shown in Figure 2.4 confirming that the differences between the MMC default risk in contraction and expansion trail off in the limit as $\Delta t \rightarrow \infty$. A feature of the diagonal entries in the sectoral MMC matrices is that at low credit quality levels (*BBB* and below), ratings volatility is higher for financials than industrials/utilities whereas the opposite holds at top credit quality levels (*A* and above). This pattern is common across expansion and contraction phases although somewhat stronger for the latter, e.g. the probability of staying in rating *B* is 11.2 (expansion) and 12.5 (contraction) percentage points higher for industrials than financials; Appendix 2.C reports the sectoral 1-year MMC contraction matrices. The fact that lower-rated financials appear relatively more volatile than industrials/utilities, mirrored in more frequent upgrades, may be linked to bank bail-outs or ‘gambling for resurrection’ strategies characterized by excessive

Table 2.4: One-Year Corporate Default Risk Estimates from Existing Studies

Panel A: *Classical estimators: time-homogeneity*

Estimator	Nickell et al. (2000)	Bangia et al. (2002)	Lando & Skodeberg (2002)	Jafray & Schuermann (2004)	Hanson & Schuermann (2006)	Christensen et al. (2004)	Ours
Cohort	6.90; 20.6	5.45; 23.69	5.00; 20.48	14.61; 30.92	6.52; 28.54	—	4.67; 27.80
Hazard rate	—	—	4.40; 37.60	13.56; 44.02	4.70; 42.49	4.82; 43.54	3.87; 40.98
Sample period	12/70-12/97	01/81-12/98	01/88-12/98	01/81-12/02	01/81-12/02	01/87-12/91	01/81-12/06
Data source	Moody's	S&P	S&P	S&P	S&P	Moody's	S&P

Panel B: *Estimators controlling for economic cyclicity*

	Nickell et al. (2000)	Bangia et al. (2002)	Present study
			Naive
			MMC
Expansion	4.5; 22.4	3.90; 27.16	4.31; 40.90
Contraction	9.4; 23.5	8.16; 42.58	18.02; 65.14
Sample period	12/70-12/97	01/81-12/98	01/81-12/06
Data source	Moody's	S&P	S&P

Notes: This table summarizes existing default risk estimates for B and CCC corporate bonds over a 1-year horizon. The two default probabilities shown in each case are for B -rated bonds and CCC -rated bonds, $(P_{B,D}; P_{CCC,D})$, in percentage points. The reported figures in Panel B from Nickell et al. (2000) and Bangia et al. (2002) are for the naive cyclical approach that deploys the cohort estimator separately on 'expansion' and 'contraction' ratings. Nickell et al. (2000) categorize the business cycle phases as 'peak', 'normal' and 'trough'; the reported figures are for the former and the latter phases, respectively. Bangia et al. (2002) report quarterly transition probabilities that we have annualized.

Table 2.5: Rating Transition Matrices Conditional on the Economic State

	AAA	AA	A	BBB	BB	B	CCC	D	NR
Panel I: <i>Naive Expansion estimator</i>									
AAA	89.030	5.683	0.704	0.102	0.117	0.019	0.001	0.017	4.327
AA	0.595	86.196	7.827	0.767	0.076	0.095	0.014	0.023	4.407
A	0.071	1.555	86.912	5.786	0.517	0.188	0.009	0.034	4.927
BBB	0.019	0.193	3.618	83.970	4.695	0.683	0.077	0.132	6.614
BB	0.032	0.083	0.384	4.894	76.165	7.828	0.589	0.666	9.359
B	0.003	0.050	0.202	0.467	4.606	74.220	4.673	4.314	11.465
CCC	0.002	0.013	0.354	0.516	0.949	7.650	37.210	40.903	12.403
D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000
NR	0.026	0.083	0.207	0.338	0.386	0.463	0.028	0.751	97.719
Panel II: <i>Naive Contraction estimator</i>									
AAA	56.194	14.169	6.214	2.091	0.301	0.106	0.012	0.626	20.286
AA	0.785	49.686	18.436	4.841	0.995	0.222	0.027	1.506	23.503
A	0.155	3.242	56.628	12.161	1.950	0.621	0.081	1.938	23.224
BBB	0.019	1.048	6.874	58.507	6.993	1.365	0.113	2.617	22.463
BB	0.009	0.228	1.970	6.205	46.317	6.926	0.865	8.700	28.780
B	0.007	0.137	0.643	1.803	3.136	44.912	4.201	18.015	27.147
CCC	0.004	0.036	0.110	0.557	0.811	2.275	15.696	65.135	15.376
D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000
NR	0.034	0.260	0.540	1.014	0.677	0.560	0.068	4.299	92.549
Panel III: <i>MMC Expansion estimator</i>									
AAA	86.809	6.340	1.071	0.226	0.124	0.023	0.002	0.046	5.359
AA	0.611	83.679	8.647	1.037	0.135	0.101	0.015	0.108	5.668
A	0.077	1.684	84.858	6.259	0.613	0.217	0.014	0.145	6.133
BBB	0.018	0.254	3.854	82.264	4.872	0.730	0.078	0.282	7.649
BB	0.030	0.092	0.497	4.995	74.129	7.782	0.614	1.185	10.676
B	0.003	0.055	0.232	0.563	4.504	72.223	4.661	5.217	12.541
CCC	0.002	0.014	0.337	0.523	0.943	7.276	35.661	42.581	12.661
D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000
NR	0.026	0.095	0.228	0.383	0.405	0.469	0.030	0.973	97.390
Panel IV: <i>MMC Contraction estimator</i>									
AAA	70.961	10.697	3.716	1.156	0.203	0.067	0.007	0.297	12.896
AA	0.711	65.902	14.057	2.997	0.577	0.163	0.019	0.775	14.798
A	0.120	2.548	70.218	9.425	1.313	0.435	0.048	1.013	14.880
BBB	0.019	0.677	5.499	70.022	6.051	1.082	0.095	1.404	15.152
BB	0.019	0.157	1.285	5.652	59.671	7.397	0.743	5.018	20.058
B	0.005	0.096	0.448	1.221	3.795	58.059	4.461	11.770	20.144
CCC	0.003	0.023	0.193	0.531	0.872	4.410	24.946	54.988	14.032
D	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000
NR	0.030	0.181	0.387	0.705	0.549	0.524	0.049	2.627	94.948

Notes: This table presents the rating transition probabilities obtained using the naive business-cycle estimator when the current economic phase is expansion (Panel I) or contraction (Panel II), and the MMC business-cycle estimator when the current economic phase is expansion (Panel III) or contraction (Panel IV). Full details of the computations are given in Section 2.3. All figures are in percentage points.

Table 2.6: Default Rates for CCC Bonds per Sector

	1 Year	2 Year	3 Year	4 Year	5 Year
<i>Baseline Hazard-Rate estimator</i>					
Industrial	42.257	57.759	63.789	66.413	67.769
Utility	39.403	55.910	63.158	66.593	68.406
Financial	37.699	53.852	61.111	64.612	66.468
Overall	40.913	57.096	63.804	66.829	68.386
<i>MMC Expansion estimator</i>					
Industrial	43.513	58.817	64.515	67.008	68.382
Utility	40.974	57.582	64.544	67.752	69.444
Financial	41.020	56.982	63.405	66.264	67.722
Overall	42.581	58.587	64.844	67.613	69.089
<i>MMC Contraction estimator</i>					
Industrial	56.869	68.708	72.678	74.380	75.322
Utility	52.485	66.363	71.708	74.154	75.456
Financial	55.375	66.695	70.738	72.480	73.355
Overall	54.988	66.863	71.105	72.984	74.010

Notes: This table provides estimates of the probability that a firm currently rated *CCC* enters default over a time horizon $(t, t + \Delta_t)$ from 1 to 5 years. The top panel corresponds to the hazard-rate estimator that ignores business cycles. The mid and bottom panels pertain, respectively, to the MMC estimator conditional on current expansion and contraction. The labels in the first column indicate that the estimates are based on the rating histories of firms either in the industrial sector, utility sector, financial sector or the entire sample.

risk-taking influenced by moral hazard; see Goodhart (2006). Through explicit or implicit deposit insurance, banks on the road to insolvency can disguise the problem by aggressively raising money through unsustainable high interest rates.

2.4.4 Empirical distribution of default rates

We now analyze the small-sample properties of the default risk estimators. Relative accuracy is gauged by comparing confidence intervals around the PD estimates as a way of quantifying estimation error or noise. In order to assess the statistical significance of differences in the default risk measures, the PDs are re-estimated M times using bootstrap (artificial) ratings samples.¹⁶ Each bootstrap sample has the

¹⁶ A simple approach to calculate analytical confidence intervals for the discrete cohort PDs, which many studies such as Nickell et al. (2000) and Christensen et al. (2004) adopt, is through the binomial distribution under the simplifying assumption that the ratings are independent over time and across

same number of obligors (cross-section size N) as the original dataset and is obtained as follows: a bond's entire rating history is randomly drawn with replacement so as to preserve the serial (e.g. business cycle) dependence in ratings; N random draws are thus made. This process is repeated $M = 1,000$ times.¹⁷ This non-parametric bootstrap approach where the unit of resampling is a realized bond-history is advocated by Hanson and Schuermann (2006) and Löffler and Posch (2007) to circumvent having to choose a data generating process for the ratings.

The bootstrap simulation results for the classical cohort and hazard-rate PDs over a 1-year horizon are set out in the top left panel of Table 2.7. The mean and standard deviation of the PD estimates over replications are given, first, followed by the 95% confidence interval and the interval length. The bootstrapped intervals for hazard-rate PDs are much tighter than those for cohort PDs, especially with top ratings (e.g., about 16 times tighter for rating *A*). Our findings are in line with the studies by Christensen et al. (2004) from Moody's 1987-1991 firm ratings, Jafry and Schuermann (2004) from S&P's 1981-2002 firm ratings and Fuertes and Kalotychou (2007) from Moody's 1981-2004 sovereign ratings, in suggesting that the hazard-rate estimator is more efficient (accurate) than the cohort one. Those studies opt instead for a parametric bootstrap that uses a fitted Markov process as the basis for generating artificial rating histories.

We further investigate the impact of controlling for cyclicalities on the accuracy of the PD estimates. We first focus on the naive estimator. The top middle and right panels of Table 2.7 show that the mean naive PD over bootstrap replications is markedly higher for current recession than for contraction. The confidence intervals of PDs in economic recession are around 16 to 80 times wider (investment grades) and 2 to 14 times wider (junk grades) than in expansion. This accuracy loss stems

obligors. However, as discussed earlier, this is not a realistic assumption. To sidestep it and in order to make the comparison across estimators more informative we employ bootstrapped confidence sets throughout.

¹⁷This M choice follows from the fact that there is no evidence of non-normality in the empirical distribution of PDs (e.g., the baseline hazard-rate PD density for BBB-rated bonds exhibits small skewness and kurtosis at 0.290 and 3.017, ranging between 0.071-0.910 and 2.698-3.458, respectively, for all ratings). Efron and Tibshirani (1997) show that $M = 1,000$ bootstrap replications are sufficient to obtain a good approximation in this context.

Table 2.7: Empirical Distribution of One-Year Default Risk Estimates

	Cohort estimator			Length	Naive Expansion estimator			Length	Naive Contraction estimator			Length
	Mean	St. Dev.	(Lower, Upper)		Mean	St. Dev.	(Lower, Upper)		Mean	St. Dev.	(Lower, Upper)	
AAA	0	0	(0.000, 0.000)	0.000	0.017	0.002	(0.013, 0.022)	0.008	0.633	0.103	(0.447, 0.847)	0.400
AA	0.016	0.018	(0.000, 0.068)	0.068	0.023	0.004	(0.017, 0.034)	0.017	1.529	0.351	(0.917, 2.259)	1.342
A	0.075	0.023	(0.032, 0.128)	0.096	0.034	0.008	(0.024, 0.051)	0.028	1.920	0.300	(1.374, 2.559)	1.184
BBB	0.239	0.046	(0.174, 0.335)	0.161	0.132	0.023	(0.090, 0.177)	0.087	2.629	0.371	(1.956, 3.398)	1.443
BB	0.983	0.102	(0.824, 1.203)	0.379	0.647	0.062	(0.531, 0.770)	0.239	8.764	0.851	(7.139, 10.467)	3.328
B	4.692	0.201	(4.216, 5.122)	0.906	4.019	0.177	(3.686, 4.375)	0.689	18.047	1.025	(16.198, 20.052)	3.854
CCC	28.054	1.615	(25.066, 31.477)	6.411	41.002	1.606	(37.974, 44.340)	6.366	65.215	3.306	(58.978, 71.613)	12.635
MMC Expansion estimator												
AAA	0.015	0.002	(0.011, 0.019)	0.007	0.046	0.006	(0.040, 0.054)	0.014	0.300	0.048	(0.212, 0.400)	0.188
AA	0.020	0.004	(0.014, 0.030)	0.015	0.110	0.024	(0.081, 0.143)	0.062	0.788	0.196	(0.446, 1.196)	0.750
A	0.022	0.001	(0.019, 0.025)	0.006	0.144	0.021	(0.120, 0.175)	0.055	1.002	0.165	(0.707, 1.349)	0.641
BBB	0.098	0.018	(0.065, 0.138)	0.072	0.283	0.032	(0.225, 0.350)	0.125	1.410	0.204	(1.039, 1.843)	0.804
BB	0.519	0.052	(0.425, 0.631)	0.206	1.191	0.088	(1.016, 1.350)	0.334	5.055	0.487	(4.127, 6.024)	1.896
B	3.865	0.160	(3.559, 4.171)	0.612	5.225	0.203	(4.839, 5.615)	0.776	11.790	0.611	(10.667, 12.994)	2.327
CCC	41.015	1.562	(38.086, 44.366)	6.281	42.682	1.589	(39.733, 46.105)	6.372	55.091	2.367	(50.613, 59.402)	8.789
MMC Contraction estimator												
Hazard-Rate vs Naive Expansion												
	Mean	St. Dev.	(Lower, Upper)	H0: $\Delta = 0$	Mean	St. Dev.	(Lower, Upper)	H0: $\Delta = 0$	Mean	St. Dev.	(Lower, Upper)	H0: $\Delta = 0$
AAA	-0.002	0.001	(-0.004, 0.000)	Not Reject	-0.618	0.102	(-0.817, -0.418)	Reject	-0.031	0.005	(-0.041, -0.022)	Reject
AA	-0.003	0.002	(-0.006, 0.000)	Not Reject	-1.509	0.352	(-2.198, -0.820)	Reject	-0.089	0.024	(-0.135, -0.043)	Reject
A	-0.012	0.007	(-0.027, 0.003)	Not Reject	-1.898	0.300	(-2.485, -1.310)	Reject	-0.122	0.021	(-0.163, -0.081)	Reject
BBB	-0.033	0.020	(-0.072, 0.006)	Not Reject	-2.531	0.370	(-3.256, -1.805)	Reject	-0.184	0.027	(-0.237, -0.132)	Reject
BB	-0.128	0.046	(-0.237, 0.002)	Not Reject	-8.245	0.841	(-9.893, -6.598)	Reject	-0.672	0.062	(-0.794, -0.550)	Reject
B	-0.154	0.081	(-0.312, 0.008)	Not Reject	-14.182	0.970	(-16.084, -12.281)	Reject	-1.360	0.079	(-1.514, -1.205)	Reject
CCC	0.014	0.347	(-0.667, 0.695)	Not Reject	-24.199	2.927	(-29.936, -18.463)	Reject	-1.667	0.273	(-2.202, -1.132)	Reject
Hazard-Rate vs MMC Contraction												
AAA	-0.285	0.047	(-0.378, -0.192)	Reject	-0.029	0.005	(-0.039, -0.020)	Reject	0.333	0.054	(0.226, 0.439)	Reject
AA	-0.768	0.196	(-1.152, -0.384)	Reject	-0.086	0.024	(-0.132, -0.040)	Reject	0.741	0.156	(0.435, 1.047)	Reject
A	-0.980	0.165	(-1.304, -0.657)	Reject	-0.110	0.020	(-0.149, -0.071)	Reject	0.918	0.135	(0.653, 1.182)	Reject
BBB	-1.312	0.203	(-1.709, -0.915)	Reject	-0.151	0.024	(-0.199, -0.104)	Reject	1.219	0.168	(0.889, 1.548)	Reject
BB	-4.536	0.475	(-5.467, -3.605)	Reject	-0.524	0.056	(-0.634, -0.414)	Reject	3.709	0.366	(2.992, 4.427)	Reject
B	-7.925	0.548	(-8.999, -6.851)	Reject	-0.906	0.064	(-1.032, -0.780)	Reject	6.257	0.424	(5.426, 7.088)	Reject
CCC	-14.075	1.762	(-17.530, -10.621)	Reject	-1.681	0.226	(-2.124, -1.237)	Reject	10.124	1.174	(8.24, 12.424)	Reject

Notes: This table reports the mean, standard deviation, and 95% confidence interval (upper limit, lower limits and length) of the 1-year default risk estimates (top panels) and their differential (bottom panels) obtained over $M = 1,000$ bootstrap replications. $H_0 : \Delta = 0$ denotes the null hypothesis that the difference in PDs obtained from estimators i and j ($\Delta \equiv PD^i - PD^j$) is negligible. All numbers are in percentage points.

from the relatively sparse set of rating migrations observed in contraction periods since the sample entails 976 contraction days versus 8,520 expansion days.

Turning now attention to the MMC estimator, the mean PD over bootstrap replications is higher if the current phase is recession than if it is expansion, as with the naive cyclical estimator. The confidence bands for the MMC default rates are wider in contraction than in expansion but notably less so than with the naive estimator; around 5.9-7.4 times wider for investment grades and 1.6-4.5 times wider for junk grades. The PDs from the MMC estimator for current contraction are notably more accurate than those from the naive one; the confidence band length is halved.¹⁸ Such efficiency gain is not observed in expansion which is not surprising given that the expansion ratings subsample is much richer than the contraction one. The clear improvement in accuracy of the MMC estimator relative to the naive one in contraction rationalizes the large gap between the MMC and naive estimates shown previously in contraction also (c.f., Table 2.3 and Figure 2.4).

The bottom half of Table 2.7 summarizes to the distribution of the PD differential. For current contraction, the hazard-rate default risk estimates are significantly understated relative to the MMC ones. The question of whether the two business cycle-adjusted models yield statistically different PDs is addressed in the bottom (middle and right) panels of Table 2.7. In line with the findings in Section 2.4.4, the 95% confidence bands suggest that the naive approach conditional on current expansion (contraction) significantly under(over)estimates the PD relative to the MMC approach.

One important message from this simulation analysis is that risk management practices for economic capital attribution that build upon through-the-cycle (i.e.,

¹⁸The simulation results suggest that conditional on current expansion there is a gain in accuracy, albeit overall very modest, in the naive versus the MMC estimator. This maybe because the latter estimator is less parsimonious (more parameters) since it additionally controls for the fact that the business cycle is stochastically evolving over time through the switching matrix $\mathbb{S}$. The naive estimator instead assumes that the economy remains in the same economic phase throughout the migration horizon of interest. When the current phase is expansion, this assumption is relatively mild since historically expansion has been more pervasive than contraction. Thus over the entire sample the estimated probability of being in expansion over the next quarter given current expansion is 97.2%.

classical cohort or hazard) default risk estimates or upon those obtained by simply splitting the sample into contraction and expansion ratings (i.e., naive business-cycle approach) can suffer from various, bias and inefficiency, distortions especially in economic stress scenarios. This issue is further investigated in Section 2.4.6.

2.4.5 Out-of-sample forecast evaluation

In this section we conduct an out-of-sample prediction exercise to shed further light on the relative merits of the MMC estimator. We choose as evaluation or holdout period the last eight sample years (1999-2006) which amounts to one third of the data. This is a sensible choice since, as illustrated in Figure 2.3, it comprises an entire business cycle with both expansion and contraction phases. Akin to Frydman and Schuermann (2008), Koopman et al. (2008) and Stefanescu et al. (2009), we consider recursive estimation windows such that one year of ratings data is added at each iteration, i.e. 1981-1998, 1981-1999 and so forth. Using the ratings information in each window, 1- to 3-year migration matrix predictions are obtained according to the risk models entertained in the chapter. For the sake of simplicity, most of the methodological discussion focuses on the 1-year horizon.

In order to deploy the naive and MMC business-cycle estimators for prediction purposes, the forecaster needs to acknowledge the prevailing economic conditions at the current time point, namely, at the beginning (end) of the forecast horizon (estimation window) referred to as time t . In the first iteration, the economic conditions on year-end 1998 (i.e., last quarter of estimation window) are taken as the time t regime and, accordingly, a forecast is generated conditional either on current expansion or contraction for the migration risk over the subsequent 1-year horizon ending at $t+1$ (i.e., first out-of-sample year 1999) and so on.

With the naive approach, the ratings in a given estimation window ending at t are classified as ‘expansion’ or ‘contraction’, and two distinct forecasts are generated for the migration risk over the horizon $(t, t+1)$ which apply, respectively, when the time t economic phase is either expansion or contraction. Thus the prevailing economic conditions on year-end 1998 determine how the future migration risk over 1999 is

forecasted, and so on. Likewise with the MMC approach but, in contrast with the naive one, it does not assume that the current economic phase remains over the entire forecast horizon. Instead the MMC estimator uses as input the time-varying (i.e., recursively estimated) regime-switching matrix $\mathbb{S}$, Eq.(2.5), that governs the stochastic business-cycle evolution. To illustrate, since expansion prevails at the end of the first estimation window (1998:Q4), the corresponding 1-year-ahead transition risk matrix forecast incorporates the prediction that the economy will remain in expansion at the end of 1999:Q1 with probability $\hat{\theta}_t$ and will switch to contraction with probability $(1 - \hat{\theta}_t)$; and so forth over the remaining quarters of the first out-of-sample year 1999 according to the Markov chain portrayed in Figure 2.3.

Our forecasts are out-of-sample in the sense that, say, the prediction of credit migration risk over 1999 is based on data up to year-end 1998. But in order for the predictions from the naive and MMC business cycle estimators to be strictly forward-looking, we need real-time identification of the prevailing economic conditions at the point the forecasts are made (time t). For this purpose, we utilize the Chicago Fed National Activity Index (CFNAI) or, more specifically, its three-month moving average release denoted CFNAI-MA3. A practical problem with the NBER-dating employed in our in-sample analysis, and in several related studies,¹⁹ is that the announcement of a peak or trough (turning point) usually occurs many months after the event. Therefore, the NBER-dating cannot be relied upon to identify the current economic phase in a real-time framework. By contrast, the CFNAI-MA3, which is released (toward the end of) each calendar month, has been designed as an objective timely indicator of economic conditions.²⁰

¹⁹For instance, see Bangia et al. (2002), Jafry and Schuermann (2004), Hanson and Schuermann (2006) and Frydman and Schuermann (2008). The real GDP growth criterion employed in Nickell et al. (2000) to classify the ratings sample into ‘normal’, ‘peak’ and ‘trough’ suffers from the same time delay problem.

²⁰The CFNAI was developed by Stock and Watson (1999) for inflation forecasting purposes. It is the first principal component (i.e., a weighted average) of 85 inflation-adjusted economic indicators drawn from four broad categories: 1) production and income (23 series), 2) (un)employment and hours (24 series), 3) sales, orders, and inventories (23 series), 4) personal consumption and housing (15 series). It is based on projections for about 1/3 of the 85 series and therefore its real-time release is subject to subsequent revisions; however, due to its weighted-average nature the revision changes are far smaller than those of the individual series. The CFNAI has been successfully adopted as macroeconomic covariate in the credit rating model of Stefanescu et al. (2009) *inter alios*.

We employ the real-time history of the CFNAI-MA3 to label the end (i.e., final month) of each estimation window or current time t as expansion or contraction according to the ‘official’ threshold rule: *i)* a CFNAI-MA3 value below -0.7 after a period of economic expansion signals that a recession has begun, *ii)* conversely, a CFNAI-MA3 value above -0.7 after a period of economic recession is taken as suggestive that a recession has ended.²¹ Several studies have shown that this index matches remarkably well the NBER-designated business cycles and can be used to obtain good forecasts of inflation and of overall economic activity; see Brave and Butters (2010), Evans et al. (2002) and Stock and Watson (1999). Figure 2.3 plots the real-time history of the CFNAI-MA3 and illustrates that the -0.7 threshold rule yields a timely classification of expansions and recessions that is virtually identical to the lagged official NBER chronology over our sample period.

Two distinct forecast *evaluation* approaches are adopted. First, the migration risk predictions are compared with the ‘true’ migration risk. A practical difficulty, also common to the volatility forecasting literature, is that the variable being forecasted is unobserved (latent) and a proxy is needed. Stefanescu et al. (2009) proxy the true default risk by the observed yearly default frequencies (i.e., obtained by deploying the discrete cohort estimator over each of the holdout sample years) but they acknowledge a deficiency of this proxy, namely, since top-rated bonds have experienced no direct default over the sample period, their true default risk is unrealistically set to zero. Koopman et al. (2008) adopt instead a hazard-rate type proxy for the true default risk by deploying the non-parametric Aalen-Johansen estimator. In this same spirit, we deploy the continuous hazard-rate estimator (Eq.(2.4)) over each out-of-sample year (biennium or triennium) sequentially and the resulting measures are taken as true 1-year (2-year or 3-year) migration risk denoted generically $\mathbf{Q}_{t+1}$.²²

²¹See www.chicagofed.org/cfnai. Berge and Jordà (2009) develop a routine using a receiver operating characteristics (ROC) curve that yields -0.8 as alternative threshold rule which places equal weight on avoiding misclassifying a recession month as a non-recession month and a non-recession month as a recession month. The -0.7 threshold put forward by Chicago Fed researchers places marginally more weight on the second type of error.

²²Although the Aalen-Johansen (AJ) estimator allows for time-heterogeneity, several studies have shown that for large (cross-section) samples it does not produce significantly different estimates nor efficiency gains relative to the classical hazard-rate approach over short horizons, say, one to

Several criteria are adopted to compare the $K \times K$ migration risk matrix predictions ($\hat{\mathbf{Q}}_{t+1}$) and the ‘true’ migration risk ($\mathbf{Q}_{t+1}$): the $\mathbb{L}^1$ and $\mathbb{L}^2$ Euclidean distances and asymmetric extensions thereof, and a singular value decomposition (SVD) measure. For a given out-of-sample year (biennium or triennium) denoted $t+1$ the element-by-element forecast error is given by $\hat{e}_{i,j,t+1} = \hat{q}_{i,j,t+1} - q_{i,j,t+1}$ and the Euclidean distance metrics are computed as

$$MAE_{\mathbb{L}^1} \equiv \frac{1}{K^2} \sum_{i=1}^K \sum_{j=1}^K |\hat{e}_{i,j,t+1}|, \quad (2.10)$$

and its counterpart $MSE_{\mathbb{L}^2}$ that replaces the absolute errors by squared errors. Two novel asymmetric criteria are considered in the present context to allow for asymmetry in up/downgrades regarding the losses associated with over/underpredictions. A prudential view on capital requirements may imply that, from the point of view of regulators, underpredicting the probability of a downgrade is more worrisome than overpredicting it; likewise, overpredictions regarding the probability of upgrades (or of ratings stability) are less desirable than underpredictions. Accordingly, we segment the transition matrix as: *i*) upper diagonal elements (i.e., downgrades), and *ii*) lower (upgrades) and diagonal (stability) elements. Since all absolute forecast errors are less than unity, by taking their square root a heavier penalty is placed on them, i.e. $\sqrt{|\hat{e}_{t+1}|} > |\hat{e}_{t+1}|$. We extend the Mean Mixed Error (MME) loss function in Brailsford and Faff (1996) so that underprediction (U) and overprediction (O) errors corresponding to downgrades enter the loss function, respectively, in square root and absolute form; and vice versa for upgrades/no rating changes, as follows

$$MME = \frac{1}{K^2} \left[\sum_{t,i < j} \sqrt{|\hat{e}_{i,j,t+1}^U|} + \sum_{t,i < j} |\hat{e}_{i,j,t+1}^O| + \sum_{t,i \geq j} \sqrt{|\hat{e}_{i,j,t+1}^O|} + \sum_{t,i \geq j} |\hat{e}_{i,j,t+1}^U| \right]. \quad (2.11)$$

three years; the difference is much less significant than the one between the cohort and hazard-rate estimators (see Lando and Skodeberg (2002), Jafry and Schuermann (2004) and Fuertes and Kalotychou (2007)). Jafry and Schuermann (2004) illustrate that whether the hazard-rate estimator or the AJ approach is utilized makes little difference from the point of view of 1-year risk capital attribution. Since our cross-section is very large and we deploy the hazard-rate estimator over each out-of-sample year or biennium/triennium (with rating durations in quarters) the resulting ‘true’ migration risk matrices should be at worst as trustworthy as those obtained from the computationally rather expensive AJ estimator.

Moreover, the $MSE_{\mathbb{L}^2}$ criterion lends itself to an asymmetric extension (with a specific focus on rating mobility) that we put forward where all the errors enter squared but underpredictions are weighted more heavily than overpredictions for downgrades; and vice versa for upgrades. Formally,

$$MSE_{\mathbb{L}^2}^{asy} = \left[w_I^U \sum_{t,i < j} (\hat{e}_{i,j,t+1}^U)^2 + w_I^O \sum_{t,i < j} (\hat{e}_{i,j,t+1}^O)^2 \right] + \left[w_{II}^U \sum_{t,i > j} (\hat{e}_{i,j,t+1}^U)^2 + w_{II}^O \sum_{t,i > j} (\hat{e}_{i,j,t+1}^O)^2 \right] \quad (2.12)$$

with $w_I^U > w_I^O$, $w_{II}^U < w_{II}^O$ and $w_I^U + w_I^O + w_{II}^U + w_{II}^O = 1$. We consider three weight combinations: $(\frac{3}{10}, \frac{2}{10}, \frac{2}{10}, \frac{3}{10})$, $(\frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{2}{6})$ and $(\frac{4}{10}, \frac{1}{10}, \frac{2}{10}, \frac{3}{10})$. The latter case amounts to assuming a larger loss differential between underpredictions and overpredictions for downgrades. Finally, we deploy the average of singular values metric proposed by Jafry and Schuermann (2004) defined as

$$SVD \equiv \frac{1}{K} \sum_{i=1}^K \sqrt{\lambda_{i,t+1} \left(\tilde{\mathbf{Q}}'_{t+1} \tilde{\mathbf{Q}}_{t+1} \right)}, \quad (2.13)$$

where $\lambda_{i,t+1}(\mathbf{z})$ is the i th eigenvalue of $\mathbf{z}$, and $\tilde{\mathbf{Q}}_{t+1} \equiv \mathbf{Q}_{t+1} - \mathbf{I}$, with $\mathbf{I}$ denoting the identity matrix.²³ In this context, the prediction error $|\hat{e}_{t+1}|$ is defined as the absolute difference between the above SVD formula deployed on the forecasted migration risk matrix and on the ‘true’ migration matrix. Each of the forecast error metrics is calculated as outlined above over every out-of-sample year $t+1$ and then averaged out over $t = 1, 2, \dots, 8$; likewise, over the 2- and 3-year periods.

Our second forecast evaluation approach, following Frydman and Schuermann (2008), circumvents the difficulty of having to proxy the true migration risk. Forecast ability is gauged by subtracting the forecasted probability of each rating transition realized by the end of each out-of-sample year (biennium or triennium) from 1. For instance, take a corporate bond which is rated BBB at the end of the first estimation window (i.e. year-end 1998) and remains rated BBB at the end of the first out-of-sample year (i.e. year-end 1999) and a second BBB bond that was instead rated AA

²³By subtracting the identity matrix, the resulting migration matrix reflects just mobility, that is, the focus of the SVD metric (like $MSE_{\mathbb{L}^1}^{asy}$ which is computed from off-diagonal elements) is the dynamic part of $\mathbf{Q}_{t+1}$.

at year-end 1999. Suppose that the 1-year migration risk matrix forecast for 1999 gives $\hat{p}_{BBB,t+1}=0.93$ as diagonal entry corresponding to BBB and $\hat{p}_{BBB,AA,t+1} = 0.42$. Hence, the forecast error for the first bond is small, namely, $1 - 0.93$, but relatively large for the second bond, $1 - 0.42$. We deploy this approach and summarize the resulting error over all corporate bonds using mean absolute and mean square error metrics subsequently denoted, respectively, MAE_{1-p} and MSE_{1-p} .

The average forecast errors are summarized in Table 2.8 as percentage reduction relative to a benchmark. A general message that comes across is that the naive business-cycle estimator provides (from very little to) no forecast gains vis-à-vis the classical hazard-rate estimator. By contrast, the MMC business-cycle estimator entails improvements in out-of-sample forecasting performance relative to both benchmarks, hazard-rate and naive cyclical estimator, and across all forecast horizons. For instance, in terms of Mean Mixed Error (MME), the 1-year MMC migration risk estimator affords a forecast error reduction of 4% vis-à-vis the naive cyclical estimator and 5.48% vis-à-vis the through-the-cycle hazard rate estimator. Overall, with all criteria and horizons, the MMC estimator provides forecast improvements relative to the naive counterpart ranging from 13.82% (SVD) to 3.97% (MAE_{1-p}) for the 1-year horizon, and from 59.35% ($MSE_{\mathbb{L}^2}$) to 2.72% (MAE_{1-p}) for the 3-year. The out-of-sample forecast error reduction of the MMC cyclical estimator relative to the hazard-rate benchmark falls between 0.58% (MAE_{1-p}) and 12.34% (SVD) over the 1- to 3-year horizons. The improvements in forecast accuracy afforded by the MMC estimator relative to the through-the-cycle hazard rate benchmark are generally more sizable on the basis of asymmetric loss functions than with the symmetric ones (e.g. $MSE_{\mathbb{L}^2}^{asy}$ versus $MSE_{\mathbb{L}^2}$). Thus conditioning on the economic state becomes even more relevant according to novel “regulatory” oriented asymmetric loss functions that attach a heavier penalty to underpredictions of rating downgrade risk than to overpredictions and vice versa for upgrades.

The percentage forecast error reduction of the MMC estimator relative to the naive counterpart is more noticeable as the horizon of interest increases, for instance, it more

Table 2.8: Out-of-Sample Forecast Errors

Estimator	Bench- mark	$MAE_{\mathbb{L}^1}$	$MSE_{\mathbb{L}^2}$	MME	$MSE_{\mathbb{L}^2}^{asy}$	SVD	MAE_{1-p}	MSE_{1-p}
1-year horizon								
N.C.	H.R.	-8.50	-55.98	1.54	2.96	-1.71	-3.37	-9.41
MMC-C.	H.R.	3.63	8.77	5.48	8.59	12.34	0.74	1.56
	N.C.	11.17	41.51	4.00	5.80	13.82	3.97	10.03
2-year horizon								
N.C.	H.R.	-40.75	-106.72	-3.64	-3.75	-64.65	-2.75	-20.91
MMC-C.	H.R.	2.60	2.53	3.55	8.13	1.11	0.63	0.93
	N.C.	30.80	52.85	6.93	11.45	39.94	3.29	18.06
3-year horizon								
N.C.	H.R.	-51.01	-144.53	-8.09	-11.56	-84.20	-2.19	-24.66
MMC-C.	H.R.	1.55	0.59	3.85	8.92	0.90	0.58	0.80
	N.C.	34.81	59.35	11.05	18.36	46.20	2.72	20.43

Notes: This table shows the percentage reduction in the average out-of-sample forecast error of each estimator (column 1) vis-à-vis a benchmark (col. 2); positive numbers denote a decrease in average forecast error. $MAE_{\mathbb{L}^1}$, MME , $MSE_{\mathbb{L}^2}^{asy}$ and SVD , as formalized in (10), (11), (12) and (13), respectively, are forecast error metrics that require ‘true’ rating migration risk. $MSE_{\mathbb{L}^2}$ is the squared error version of $MAE_{\mathbb{L}^1}$. H.R. refers to Hazard Rate. N.C. refers to Naive cyclical. MMC-C. denotes MMC cyclical. The 1-year, 2-year and 3-year migration risk matrices obtained by applying the continuous time hazard-rate estimator over each out-of-sample year (biennium or triennium) are taken as ‘true’ migration risk. MAE_{1-p} and MSE_{1-p} are Frydmann and Schuermann (2008) evaluation metrics that are based on the forecasted probabilities of each actual transaction over the out-of-sample period and do not require a ‘true’ migration risk proxy. Naive cyclical and MMC cyclical are the two estimators that account for the current economic conditions. Hazard rate is the continuous through-the-cycle estimator. $MSE_{\mathbb{L}^2}^{asy}$ is based on weights $(\frac{4}{10}, \frac{1}{10}, \frac{2}{10}, \frac{3}{10})$. The out-of-sample period are the last 8 sample years.

than trebles from 11.17% (1-year) to 34.81% (3-year) with the $MAE_{\mathbb{L}_1}$ loss function.²⁴ The intuition behind this pattern is that the naive estimator's implicit assumption that the current economic conditions prevail over the entire forecast horizon becomes less innocuous as the latter lengthens. This is important in the light of the new Basel III Accord (under preparation) which states as one of its goals to increase the mandatory time horizon for the estimation of default risk.²⁵

2.4.6 Economic relevance: risk capital attribution

The Basel Committee requires banks to hold sufficient Tier 1 and Tier 2 capital to cover unexpected credit losses over a 1-year horizon at the 99.9% confidence level. Accordingly, a bank failure should be observed only once in a thousand years. The mean of the credit loss distribution is the expected loss associated with all possible changes in credit quality over a target horizon which is covered by credit reserves, while Value-at-Risk (VaR) is the quantile of the credit loss distribution which will not be exceeded at a given probability level α . The unexpected loss, or discrepancy between the α -th VaR measure and the expected credit loss, defines economic capital at confidence level α .²⁶ We now compare the various rating migration measures through the lens of risk capital attribution.

To this end, we utilize the popular CreditRisk⁺ model to derive the portfolio default loss distribution over a 1-year horizon.²⁷ In this model, obligors are classified

²⁴The reported $MSE_{\mathbb{L}_2}^{asy}$ are for weights $(\frac{4}{10}, \frac{1}{10}, \frac{2}{10}, \frac{3}{10})$ but qualitatively similar results are obtained for the other two sets of weights, e.g. the forecast error reduction of the MMC versus naive estimator is 5.08% (1-year), 10.36% (2-year) and 17.92% (3-year) for $(\frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{2}{6})$. Moreover, when the focus is exclusively on default risk the same pattern is observed, e.g. the $MSE_{\mathbb{L}_2}$ criteria illustrates forecast gains of MMC versus naive of 5.54% (1-year), 14.24% (2-year) and 21.08% (3-year).

²⁵For a detailed exposition of the Basel III regulatory framework see www.bis.org/bcbs/basel3.htm.

²⁶The bank's risk manager faces the task of justifying that the estimated economic capital (based on default risk estimates) reflects the actual level of credit risk the institution is taking and to present evidence to the regulators. VaR backtesting has become standard in this regard. Regulatory capital acts as a constraint for banks in the sense that target capital ratios usually exceed regulatory capital, the so called "headroom", for strategic reasons (e.g. to be able to take advantage of growth opportunities), operational reasons (e.g. to avoid the direct and indirect costs of having to raise capital at short notice) and to mitigate regulatory intervention; see Francis and Osborne (2012).

²⁷For details see <http://www.csfb.com/institutional/research/assets/creditrisk.zip>

into n independent sectors; a ‘sector’ is a group of obligors under the influence of a common systematic factor which induces default correlations. Each factor $k = 1, \dots, n$ is assumed to be Gamma distributed with mean default intensity $\lambda_k = \sum_j \theta_{jk} \lambda_j$ where $\lambda_j = -\log(1 - \overline{PD}_j)$ and volatility of default intensity $\sigma_k = \sum_j \sigma_{PD_j}$ where $\overline{PD}_j$ and σ_{PD_j} are, respectively, the mean and standard deviation of the j th obligor default probability; each of the sector weights θ_{jk} represents the extent to which sector k influences obligor j , so that $\sum_k \theta_{jk} = 1$.²⁸ Hence, the model inputs needed to derive the closed-form distribution of the portfolio losses are, for each obligor: the sector weights (θ_{jk}), the loss given default (LGD), the exposure at default (EAD), the mean probability of default ($\overline{PD}$) and its volatility (σ_{PD}).

We build a fictitious credit portfolio of 100 bonds and assign to each a random initial rating j uniformly drawn from the space $S = \{AAA, \dots, CCC\}$. $\overline{PD}_j$ and σ_{PD_j} are taken from the bootstrap distributions set out in Table 2.7. The EAD for each bond is randomly drawn from a uniform distribution with range \$1 to \$1m summing to a total of \$50,030,818 for the portfolio. All obligors pertain to the same sector (thus $n = 1$ and $\theta_{jk} = \theta_j = 1$ for all j) and have full LGD. Table 2.9 reports the economic risk capital estimates at confidence levels $\alpha = \{99.0\%, 99.9\%\}$. The risk capital suggested by the baseline hazard rate estimator at \$6.64m (99.0% level) and \$8.25m (99.9% level) is about 15% larger than that suggested by the less efficient cohort estimator; this confirms the earlier evidence in Jafry and Schuermann (2004) that the choice between a discrete-time or continuous-time estimator can matter substantially for economic capital assessment.

At the 99.9% level, the naive business-cycle estimator suggests a risk capital of \$8.19m in expansion, a modest 0.7% decrease versus the classical hazard-rate estimator; by contrast, there is a dramatic 70% increase for the required risk capital in contraction. This asymmetry arises because the naive estimator implicitly assumes that the same current economic conditions stay during the 1-year horizon. Histori-

²⁸CreditRisk⁺ classifies the obligors in each sector into $i = 1, \dots, m(k)$ sub-portfolios or bands of similar exposure at default. The distribution of the number of default events in each exposure band is treated as Poisson with mean equal to the expected number of defaults in each sub-portfolio over one year. The default loss distribution for each sector is thus obtained by aggregating with weights θ_{ik} the individual sub-portfolio loss distributions.

Table 2.9: Economic Capital Attribution for One-Year Risk Horizon

	<i>Cohort</i>	<i>Hazard</i>	<i>Naïve(Exp)</i>	<i>Naïve(Con)</i>	<i>MMC(Exp)</i>	<i>MMC(Con)</i>
99.0% level	\$5,467,943	\$6,644,600	\$6,600,314	\$11,589,362	\$7,038,285	\$9,469,067
99.9% level	\$6,918,574	\$8,250,077	\$8,193,990	\$13,774,556	\$8,693,018	\$11,416,857
		<u><i>Cohort</i></u>	<u><i>Naïve(Exp)</i></u>	<u><i>Naïve(Con)</i></u>	<u><i>MMC(Exp)</i></u>	<u><i>MMC(Con)</i></u>
		<i>Hazard</i>	<i>Hazard</i>	<i>Hazard</i>	<i>Hazard</i>	<i>Hazard</i>
99.0% level	–	82.29%	99.33%	174.42%	105.92%	142.51%
99.9% level	–	83.86%	99.32%	166.96%	105.37%	138.38%
			<u><i>MMC(Exp)</i></u>	<u><i>MMC(Con)</i></u>	<u><i>Naïve(Con)</i></u>	<u><i>MMC(Con)</i></u>
			<i>Naïve(Exp)</i>	<i>Naïve(Con)</i>	<i>Naïve(Exp)</i>	<i>MMC(Exp)</i>
99.0% level	–	–	106.64%	81.70%	175.59%	134.54%
99.9% level	–	–	106.09%	82.88%	168.11%	131.33%

Notes: This table shows the capital requirements implied by the rating migration risk measures obtained from the two classical through-the-cycle estimators (cohort, hazard-rate) and the two cyclical estimators (naïve and MMC) which are inputs to the CreditRisk+ portfolio model of Credit Suisse First Boston (CSFB, 1997). The hypothetical credit portfolio is made up of 100 bonds with random exposure at default (EAD) ranging from \$1 to \$1m, and random initial rating. Cohort and Hazard refers to the two classical through-the-cycle estimators. *Exp* denotes current economic expansion and *Con* denotes current economic contraction. Rows 1 and 2 report the implied capital allocation levels in US\$. Rows 3 to 6 report relative capital allocation levels.

cally, expansions have been more pervasive than contractions and so the gap between the hazard-rate and naïve estimator is plausibly very modest in expansion. Although the naïve estimator we deploy is of continuous-time (hazard rate) type whereas that in Bangia et al. (2002) builds on the discrete-time (cohort) framework, our analyzes concur in suggesting that classical through-the-cycle approaches can greatly underestimate economic capital in contractions.

The 99.9% economic capital suggested by the MMC estimator in expansion (contraction) at \$8.69m (\$11.42m) represents an increase (decrease) of 6% (17%) vis-à-vis the naïve counterpart measure. The contraction risk capital is 1.7 times that in expansion according to the naïve estimator but only 1.3 times larger according to the MMC estimator. Thus the naïve estimator underestimates risk capital in expansion and substantially overestimates it in recession. In times of economic stress banks could free up to 17% of capital by opting for the MMC business-cycle approach instead of the naïve counterpart. This may have important macroeconomic implications since holding a large capital buffer is costly for banks and impairs their ability to grant credit. Excessive cyclicalities in risk capital – a grievance of the naïve cyclical estima-

tor – may materialize, unfortunately, in less lending during a downturn or a “credit crunch” period which could further aggravate the economic conditions. The more contained cyclical nature of the capital requirements associated with the MMC estimator vis-à-vis the naive counterpart is attractive in the context of another of the new Basel III reforms which seeks to dampen the procyclical amplification of financial shocks.

2.5 Conclusions

The Basel Committee on Banking Supervision published in 2004 the Basel II Accord that allows banks to use internal ratings-based models in deriving loan loss distributions and credit risk-weights for their assets. Given that regulatory measures of financial strength such as the Tier 1 capital ratio are expressed as core capital to total risk-weighted assets, the choice of approach for estimating credit migration risk is a key determinant of the capital that banks hold against unexpected losses. This chapter enriches the literature by rigorously assessing the merit of accounting for economic conditions in credit risk measurement. We advocate a Mixture of Markov Chains (MMC) estimator of rating migration risk which explicitly recognizes the stochastic business cycle. A particular case of the MMC estimator is the de facto naive cyclical approach that conditions deterministically on economic phases by assuming that the same conditions prevail throughout the prediction horizon. We compare the MMC estimator with the naive cyclical counterpart and with classical through-the-cycle estimators in three different frameworks. One is purely statistical and uses simulations to assess the estimators’ in-sample properties with emphasis on accuracy. The second is a forward-looking framework that evaluates credit risk forecasts using conventional and novel (a)symmetric loss functions. Third, we confront the capital requirements implied by the different estimators. The analysis is based on a 26-year sample of Standard & Poor’s US corporate bond ratings.

Ignoring business cycles significantly understates default risk during economic contraction. The MMC approach yields more reliable default risk measures than the naive cyclical estimator, especially in contraction. The same conclusions are

reached when the analysis is conducted at the sectoral level. In terms of out-of-sample prediction, the performance of the MMC estimator is superior to that of the naive counterpart and this is clearly revealed through novel asymmetric loss functions which attach a relatively heavy penalty to under(over)predictions of down(up)grade risk. These forecast accuracy gains become more prominent as the time horizon lengthens.

An application to economic capital attribution via the CreditRisk+ model suggests that the buffers prescribed by the MMC and naive cyclical approaches are higher than those from classical through-the-cycle estimators, particularly in economic contraction. However, default risk during contraction (expansion) is statistically and economically overestimated (underestimated) by the naive cyclical approach relative to the MMC approach and more so for longer prediction horizons. The MMC estimator here proposed, which can be seen as a way to perform stress testing, prescribes about 17% less capital holdings during downturns and 6% more capital in expansions than the naive counterpart. The excess cyclicity in capital requirements associated to the naive model would make lending very costly for banks in troubled times imposing too great a cost on economic growth and potentially aggravating a contraction. Our analysis has important implications for the ongoing financial regulatory reforms. The properties of the MMC estimator here documented become quite relevant in the light of the Basel III initiatives to lengthen the time horizon over which to measure credit risk, and to promote countercyclical capital buffers in order to dampen procyclicality. Thus by adopting more sophisticated models that account for the stochastic business cycle, the banking system can serve better as shock absorber instead of transmitter of risk to the broader economy.

Appendix

2.A Standard & Poor's Rating Definitions

Long-Term Issue Credit Ratings. Long-term ratings assigned to obligations with an original maturity above 365 days which are based, in varying degrees, on S&P's analysis of the following considerations: *i)* Likelihood of payment—capacity and willingness of the obligor to meet its financial commitment on an obligation in accordance with the terms of the obligation; *ii)* Nature and provisions of the obligation; *iii)* Protection afforded by, and relative position of, the obligation in the event of bankruptcy, reorganization, or other arrangement under the laws of bankruptcy and other laws affecting creditors' rights. The general meaning of the main rating categories is:

AAA. The obligor has a extremely strong capacity to meet its financial commitment on the obligation. Highest rating.

AA . Very strong capacity to meet financial commitments.

A. Strong capacity to meet financial commitments, but somewhat susceptible to adverse economic conditions and changes in circumstances.

BBB. Adequate capacity to meet financial commitments, but more subject to adverse economic conditions.

BB. Less vulnerable in the near-term but faces major ongoing uncertainties to adverse business, financial and economic conditions.

B. More vulnerable to adverse business, financial and economic conditions but currently has the capacity to meet financial commitments.

CCC. Currently vulnerable and dependent on favorable business, financial and economic conditions to meet financial commitments.

CC. Obligation currently highly vulnerable to non-payment.

C. Obligations that have payment arrears allowed by the terms of the documents, or obligations of an issuer that is the subject of a bankruptcy petition or similar action which have not experienced a payment default.

D. Payments of an obligation are not made on the date due even if the applicable

grace period has not expired, unless S&P's believes that such payments will be made during such grace period.

(Source: Standard & Poor's.)

2.B Three-Regime MMC Estimator

The ratings sample is divided into 3 subsamples referred to as expansion, contraction and “normal” (intermediate) state. For each subsample, a one-year rating migration matrix denoted, respectively, $\mathbf{Q}_E$, $\mathbf{Q}_C$ and $\mathbf{Q}_N$, is calculated using the continuous hazard rate estimator (Eq.(2.4)).

- The regime-switching matrix is $\mathbb{S} = \begin{pmatrix} \alpha & \beta & (1 - \alpha - \beta) \\ \delta & \mu & (1 - \delta - \mu) \\ \eta & (1 - \eta - \varepsilon) & \varepsilon \end{pmatrix}$

- Assuming that the initial state is economic expansion, the mixture matrix is

$$\mathbb{M} = \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 & \mathbf{M}_3 \\ \mathbf{M}_4 & \mathbf{M}_5 & \mathbf{M}_6 \\ \mathbf{M}_7 & \mathbf{M}_8 & \mathbf{M}_9 \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{Q}_E & \beta\mathbf{Q}_N & (1 - \alpha - \beta)\mathbf{Q}_C \\ \delta\mathbf{Q}_E & \mu\mathbf{Q}_N & (1 - \delta - \mu)\mathbf{Q}_C \\ \eta\mathbf{Q}_E & (1 - \eta - \varepsilon)\mathbf{Q}_N & \varepsilon\mathbf{Q}_C \end{pmatrix}$$

- Let $\mathbb{X}$, $\mathbb{Y}$, $\mathbb{Z}$, $\mathbb{F}$, and $\mathbb{L}'$ be defined as

$$\mathbb{X} \equiv \mathbf{M}_1 + \mathbf{M}_2 + \mathbf{M}_3 = \alpha\mathbf{Q}_E + \beta\mathbf{Q}_N + (1 - \alpha - \beta)\mathbf{Q}_C$$

$$\mathbb{Y} \equiv \mathbf{M}_4 + \mathbf{M}_5 + \mathbf{M}_6 = \delta\mathbf{Q}_E + \mu\mathbf{Q}_N + (1 - \delta - \mu)\mathbf{Q}_C$$

$$\mathbb{Z} \equiv \mathbf{M}_7 + \mathbf{M}_8 + \mathbf{M}_9 = \eta\mathbf{Q}_E + (1 - \eta - \varepsilon)\mathbf{Q}_N + \varepsilon\mathbf{Q}_C$$

$$\mathbb{F} \equiv (\mathbf{M}_1 \ \mathbf{M}_2 \ \mathbf{M}_3) \text{ and } \mathbb{L}' = (\mathbb{X} \ \mathbb{Y} \ \mathbb{Z})'$$

- The Markov-switching credit migration matrix over different time horizons is

$$\text{1-year horizon: } \mathbb{Q}_E(1) \equiv \alpha\mathbf{Q}_E + \beta\mathbf{Q}_N + (1 - \alpha - \beta)\mathbf{Q}_C = \mathbb{X}$$

$$\text{2-year horizon: } \mathbb{Q}_E(2) \equiv \mathbf{M}_1\mathbb{X} + \mathbf{M}_2\mathbb{Y} + \mathbf{M}_3\mathbb{Z} = \mathbb{F}\mathbb{L}'$$

$$\text{3-year horizon: } \mathbb{Q}_E(3) \equiv \mathbb{F}\mathbb{M}\mathbb{L}'$$

$$\text{4-year horizon: } \mathbb{Q}_E(4) \equiv \mathbb{F}\mathbb{M}^2\mathbb{L}'$$

⋮

n-year horizon: $\mathbb{Q}_E(n) \equiv \mathbb{F}\mathbb{M}^{n-2}\mathbb{L}'$

If the initial state is “normal” then $\mathbb{F} \equiv (\boldsymbol{M}_4 \ \boldsymbol{M}_5 \ \boldsymbol{M}_6)$ whereas for initial economic contraction $\mathbb{F} \equiv (\boldsymbol{M}_7 \ \boldsymbol{M}_8 \ \boldsymbol{M}_9)$.

2.C Sectoral MMC Migration Risk During Contraction

	AAA	AA	A	BBB	BB	B	CCC	D	NR
Industrials									
AAA	68.252	13.577	6.576	0.939	0.611	0.103	0.008	0.253	9.682
AA	0.683	64.800	15.044	3.513	0.645	0.198	0.015	0.520	14.581
A	0.048	1.183	65.677	10.093	2.771	0.709	0.041	1.718	17.759
BBB	0.011	0.558	4.429	66.024	9.331	1.977	0.125	2.232	15.312
BB	0.009	0.181	0.816	4.638	61.760	8.825	0.530	4.046	19.192
B	0.009	0.030	0.232	0.963	3.689	59.774	3.561	10.091	21.652
CCC	0.004	0.011	0.080	1.008	1.724	4.764	24.978	56.869	11.316
D	0	0	0	0	0	0	0	100	0
NR	0.058	0.022	0.298	0.637	0.851	0.918	0.097	3.040	94.078
Utilities									
AAA	64.985	6.120	5.801	3.068	0.144	0.027	0.002	0.197	19.655
AA	0.213	59.639	15.799	4.024	0.696	0.053	0.004	0.292	19.281
A	0.016	2.888	70.400	11.524	0.734	0.118	0.009	0.719	13.592
BBB	0.022	0.528	7.658	72.683	3.385	0.577	0.066	0.617	14.284
BB	0.045	0.110	2.022	10.859	59.339	4.690	0.468	5.068	17.398
B	0.003	0.043	0.604	1.750	4.133	59.519	5.777	12.754	15.418
CCC	0.001	0.008	0.573	0.269	0.364	6.424	30.077	52.485	9.799
D	0	0	0	0	0	0	0	100	0
NR	0.006	0.035	0.586	1.007	0.133	0.104	0.009	1.242	96.878
Financials									
AAA	74.818	12.738	3.019	0.396	0.054	0.010	0.002	0.149	8.814
AA	1.192	72.759	11.638	2.181	0.159	0.024	0.014	0.985	11.047
A	0.049	3.215	73.056	6.807	0.625	0.092	0.016	0.509	15.631
BBB	0.018	1.410	5.207	65.253	6.465	0.856	0.200	1.733	18.857
BB	0.008	0.394	2.114	6.102	46.146	8.773	2.145	8.027	26.292
B	0.007	0.307	2.556	3.922	5.510	47.284	7.094	13.374	19.947
CCC	0.005	0.097	0.077	0.110	0.980	2.526	24.224	53.375	15.852
D	0	0	0	0	0	0	0	100	0
NR	0.049	0.844	0.569	0.289	0.041	0.110	0.016	0.728	97.356

Notes: This table reports the probability that an industrial, utility or financial firm rated at time t , as indicated in the first column, is rated at time $t + \Delta t$ where $\Delta t = 1$ year, as indicated in the first row. *NR* denotes Not Rated status. *D* indicates default. The MMC estimator assumes that the economy is in contraction at time t and evolves stochastically over $(t, t + \Delta t)$. The counterpart expansion matrices are qualitatively similar across the three sectors.

Regime-Switching Dependence between Credit Default Swap and Equity Markets

3.1 Introduction

Appropriately modeling the dependence structure of credit portfolios and systematic risk factors is important for risk managers in order to set trading limits, for traders in order to hedge the market risk of their credit positions and for pricing credit derivatives. In particular, the use of models that acknowledge shifts in the relationship between financial institutions' credit exposures and the underlying equity market can be beneficial towards the design of more adequate regulatory frameworks and reduce systemic risks during stressed market conditions. Merton (1974)'s theory indirectly suggests a link between credit derivative prices and equity prices. Firm-value structural models originating from Merton's theoretical framework rest on the fundamental asset value process, namely, the default of a firm is triggered when its value falls below a certain threshold, which is commonly characterized as an increasing function of firm leverage and the volatility of its asset values. As asset value and volatility are latent, the implementation of structural credit risk models for publicly-traded firms relies on the observable equity return and a volatility proxy (*e.g.*, historical or implied), while

the credit default swap (CDS) spread can be taken as a measure of firm default risk.¹

CDS spreads can be argued to provide more reliable signals on the default riskiness of corporate borrowers than bond spreads as bond prices are often distorted by tax and liquidity issues. CDS contracts are highly standardized and thus less likely to be influenced by aspects of the contractual agreement such as seniority, coupon rates, embedded options and guarantees. Moreover, the CDS spread does not hinge on the choice of risk-free benchmark. Longstaff et al. (2005) found that liquidity factors are a very important driver of bond yield spreads. Blanco et al. (2005) showed that the CDS market leads the bond market in terms of short-run price discovery and attribute it to the higher liquidity and trading volume of the CDS market which makes it informationally more efficient.² The perception of the CDS premium as a rather “direct” measure of default risk together with the rapid development of the CDS market have spurred an enthusiastic debate over the determinants of CDS spreads and, in particular, their sensitivity to structural factors such as equity volatility, macro-variables, firm-specific balance sheet information and credit ratings.

There are empirical researches investigating variables influencing CDS spreads. Norden and Weber (2009) discuss the link between changes in CDS spreads and stock returns, while Madan and Unal (2000), Blanco et al. (2005), and Zhang et al. (2009) also consider stock return volatility. It is intuitive that the drops of a firm’s market value (proxied by its equity value) increase the probability of default. Similarly hitting the default barrier becomes more likely if the firm value fluctuates widely. Ericsson et al. (2004) find that volatility and leverage alone explain a substantial proportion of the variation in CDS premia. Yu (2006) is the first to document shifts

¹Similar to traditional insurance policies, the seller of a CDS contract must compensate the buyer if the underlying loan defaults. In return for this protection, the buyer is required to make fixed periodic payments with predefined premium (or spread) to the seller. The spread of a CDS is the annual amount the protection buyer must pay the protection seller over the length of the contract, expressed as a percentage of the notional amount. In the event of default, the CDS buyer receives compensation and the seller takes possession of the loan.

²The global CDS market grew dramatically over a short period of time with a volume expansion from \$300 billion in 1998 to \$25.9 trillion at the end of 2011 according to the International Swaps and Derivatives Association (ISDA). This significant growth can be primarily attributed to the development of CDS indices. The market has slowed down in recent years; between June 2011 and the end of the year volumes in the CDS market declined by 12.5 percent, partly due to an increase in central clearing, the effectiveness of netting and collateral, and portfolio compression.

between “turbulent” and “calm” regimes in the dynamics of CDS spreads. A common denominator to the above studies is that they focus on the determinants of single-name CDS spreads which are notably less liquid than CDS indices.³ The launch of broad-based CDS indices in 2001 by JP Morgan and Morgan Stanley marks a new era in credit derivatives trading by offering more liquidity, tradability and transparency. However, research into the dependence structure dynamics between CDS index spreads and equity market indicators is still sparse. Bystrom (2008) finds that stock returns and stock market volatility are able to explain most of the variation in iTraxx CDS spreads. Using Markov-switching regressions, Alexander and Kaeck (2008) show that the determinants of CDS index spreads are regime-specific; implied volatility is strongly related to CDS spreads in the high volatility regime while stock returns play a bigger role in the tranquil regime.

While all of the aforementioned empirical studies implicitly rely on the conventional linear Pearson correlation as dependence measure, firm structural models inspired from Merton (1974) suggest that the marginal effect of a fall in equity value is non-constant (as linear approaches would predict) but instead driven by firm fundamentals such as leverage.⁴ Using an extension of Merton’s model with realized volatility and jumps, Zhang et al. (2009) provide evidence that the strength of the relation between credit risk and equity value depends on the firm’s credit rating. They document a nonlinear convex relation between CDS spreads and equity volatility. Cao et al. (2010) find that the link between the CDS market and implied volatility is stronger when CDS spreads are more volatile and credit ratings are lower. Empirical studies have consistently suggested that credit spread predictions obtained from Merton-type structural credit risk models underestimate historical credit spreads;

³CDS indices are pools of basic single-name CDSs providing protection against the basket of entities in the index. The first two families of indices (iBoxx and Trac-x) merged in 2004 to form the CDX in North America and iTraxx in Europe and Asia which comprise the most liquid single-name CDSs. They were both acquired by Markit in 2007 which since then administrates both CDS indices. Unlike single-name contracts, CDS index type contracts do not terminate after a credit event but instead they continue with the defaulted entity removed and the contract value reduced. For a comprehensive discussion of CDS index composition and performance, see Markit (2010).

⁴Zero correlation does not imply independence, it only rules out *linear* dependence.

e.g., Jones et al. (1984), and Eom et al. (2004)⁵. This may partly stem from the fact that the actual dependence structure of debt with equity has complex features that linear correlation models fail to capture. Recent work supports this conjecture. Hull et al. (2004a) show that theoretical CDS spreads implied from Merton’s model using equity value and volatility as inputs are nonlinearly related to historical CDS spreads. Using adaptive nonparametric regressions, Giammarino and Barrieu (2009) provide evidence that the relationship between iTraxx Europe CDS index returns and two systematic factors, Euro Stoxx 50 returns and changes in the VStoxx 50 volatility index, suffered several structural changes between November 2004 and January 2008.

We extend recent research on the nonlinear relation between credit spreads and tradable systematic risk factors by adopting copulas which represent a very versatile framework to estimate multivariate distributions. Although copulas have been employed in credit risk modeling before⁶, this is the first application of copula to model nonlinearities and asymmetries in CDS-equity dependence. The main appeal of the copula framework is that it facilitates separate modeling of the marginal distributions and the dependence and thus, a variety of dependence structures can be captured with more flexibility and parsimony than in competing frameworks (*e.g.*, multivariate GARCH). Patton (2006) introduces conditional or *dynamic* copulas to portray time-varying dependence structures which represent an important improvement upon the initial *static* copula models. The original dynamic copula framework is extended by Christoffersen et al. (2012) in order to accommodate asymmetries and trends in time-varying cross-market dependence. Far less attention has been paid to the possibility of regime-switching (RS) behavior in dependence structures. To the best of our knowledge, the only few exceptions are Garcia and Tsafack (2011), Chollete et al. (2009), Okimoto (2008) and Rodriguez (2007). One can argue though that existing RS copula models have the limitation of assuming constant state-specific dependence, *i.e.* a distinct *static* copula governs each regime, despite the fact that a

⁵An introduction of Merton’s structural model and its extensions can be found in Appendix (3.A).

⁶Crook (2011) use copula to analyze the dependence of default rates in consumer loans, Das and Geng (2006) focus on corporate debt default dependence while Hull and White (2006) confront the task of credit-default option pricing using copula.

given regime or state could linger on for years.

In this Chapter, we do not intend to extend the Merton's structural model to determine the theoretical price of credit spreads. Instead, motivated by Merton's theory and the predictions from subsequent firm value structural models, we investigate the empirical comovements between credit market and the corresponding equity market indices. We provide both methodological and empirical contributions to the literature. On the former, we propose flexible *Markov-switching dynamic* (autoregressive) copulas which capture asymmetry in the form of "high" or crisis dependence and "low" or normal dependence. Our models generalize existing Markov-switching static copula by allowing for distinct mean reversion in dependence within each regime. Empirically, we seek to provide a better understanding of the dynamic evolution of dependence and tail dependence for the European credit market, proxied by the iTraxx Europe CDS index, and two underlying systematic factors proxied by the Stoxx equity index return and VStoxx implied volatility index, respectively. We carry out a comprehensive in-sample statistical comparison of various copula models and draw overall inferences on cross-market (*i.e.*, CDS and equity) dependence at the center and tails of the bivariate distributions. Given that CDS indices have become a very important instrument for risk hedging and arbitrage trading and therefore, a key component of institutional investors' portfolios, we assess the relevance of the proposed Markov-switching dynamic copulas in the context of CDS-equity portfolios from a risk management perspective.⁷ More specifically, the economic significance of our proposition is assessed through a Value at Risk (VaR) simulation to set 1-day-ahead trading limits for CDS-equity portfolios.

We document various sudden changes in the dependence structure of CDS and equity markets over the period from September 2005 to March 2011. The identified

⁷CDS indices facilitate the transfer of marketwide or sectoral credit risk by institutional investors like hedge funds and insurance companies, and by capital structure arbitrageurs who can now use derivatives (CDS index options and futures) for managing the risk related to their CDS index positions. Yu (2006) provides evidence of capital structure arbitrage opportunities in the CDS market for industrials, *i.e.* it is possible to make profits out of a trading strategy that exploits the CDS mispricing error. The latter is defined as the difference between observed CDS market spreads and predicted CDS spread predictions from a Merton-type structural model with inputs the observed equity prices and information about the obligor's capital structure.

transitions to the *high* dependence regime largely reflect the onset of the automotive industry and energy crises in 2005, the credit crunch in 2007 and the most recent Greek and European sovereign debt crises in late 2009. The Markov-switching dynamic copula model reveals that, in crisis periods, shocks to dependence of CDS spreads with market volatility have longer-lasting effects than shocks to dependence of CDS spreads with market returns. Both in crisis and normal periods, changes in CDS premia are more strongly linked with the evolution of equity returns than with market volatility. The two distinct regimes of dependence are more clearly identified at sectoral than marketwide level. The proposed *Markov-switching dynamic* copula models are supported over simpler nested copulas not only by conventional in-sample statistical criteria but also by out-of-sample VaR forecast accuracy measures. Using regulatory loss functions that take into account both the frequency and magnitude of exceptions, the VaR simulation highlights the economic relevance of our copula models by showing that they lead to more cautious 1-day-ahead trading limits. A mismatch is documented between in-sample statistical fit and economic value of predictability regarding the choice of specific copula function; log-likelihood values and Akaike Information Criteria support the Student's t copula but lower average regulatory losses are associated to the VaR forecasts from the asymmetrically-tailed Gumbel copula.

Our findings have important implications. The proposed copula framework can be useful towards the Basel III macroprudential goal of making the banking sector more resilient to stress conditions through enhanced risk coverage. One of the reforms put forward by the Basel Committee on Banking Supervision (2011b) is precisely about strengthening capital requirements for credit exposures arising from banks credit derivatives such as CDS positions, and introducing stressed-VaR capital requirements for the trading book. Our study suggests that copula models that explicitly parametrize sudden shifts in the dependence structure between credit exposures and the equity market facilitate more conservative downside-risk measures. Hence, our results point into a clear direction for improvement of stress testing platforms and reduction of systemic risk. The copula framework proposed can be useful too

for capital structure arbitrageurs that seek to exploit temporary deviations between model-based CDS spread predictions and observed CDS market spreads.

The remainder of this chapter is organized as follows. Section 3.2 outlines the methodology and Section 3.3 describes the data. Section 3.4 provides an in-sample statistical comparison of copulas and inferences on CDS-equity dependence, followed by an evaluation of the economic significance of the Markov-switching dynamic copula formulation proposed. Section 3.5 concludes. Technical details are confined to an Appendix.

3.2 Copula Methodology

We begin by outlining the baseline theory of *static* and *dynamic* (or conditional) copulas, before proposing *regime-switching* extensions. Without loss of generality, the exposition is confined to a bivariate setting. We begin by presenting Sklar's theorem and the marginal distribution model employed to characterize individual asset returns. Next we turn attention to the copulas and their estimation approach. Bold font denotes vectors and matrices, lowercase is used for probability density function (*pdf*) and uppercase for cumulative distribution function (*cdf*). Finally, we describes the data set.

3.2.1 Sklar's theorem and marginal processes

Let x_1 and x_2 denote the whitened and standardized returns of two financial assets which are realizations of the random variables X_1 and X_2 , respectively. A *copula* is defined as follows.

Definition (Copula). A function $C : [0, 1]^2 \rightarrow [0, 1]$ is a copula if it satisfies (i) $C(u_1, u_2) = 0$ for $u_1 = 0$ or $u_2 = 0$; (ii) $\sum_{i=1}^2 \sum_{j=1}^2 (-1)^{i+j} C(u_1(i), u_2(j)) \geq 0$ for all $(u_1(i), u_2(j))$ in $[0, 1]^2$ with $u_1(1) < u_1(2)$ and $u_2(1) < u_2(2)$; and (iii) $C(u_1, 1) = u_1, C(1, u_2) = u_2$ for all u_1, u_2 in $[0, 1]$.

Sklar (1959)'s theorem expressed formally below provides the theoretical foundation for copulas.

Theorem (Sklar). *A joint cdf $H(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$ of the random variables $\mathbf{X} = (X_1, X_2)$ with respective marginal cdf $F_1(x_1) = P(X_1 \leq x_1)$ and $F_2(x_2) = P(X_2 \leq x_2)$ can be written as*

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \quad (3.1)$$

where C is a copula. Given H , if F_1 and F_2 are continuous, then there exists a unique C satisfying (3.1). Conversely, given C and the margins $F_1(x_1), F_2(x_2)$ then the resulting $C(F_1(x_1), F_2(x_2))$ is a joint cdf.

Copula is in essence a dependence function that maps two univariate *pdf* (or *margins*) into a joint *pdf*.⁸ For continuous variables, the density c corresponding to the copula C is given by

$$c(F_1(x_1), F_2(x_2)) = \frac{\partial^2 C(F_1(x_1), F_2(x_2))}{\partial F_1(x_1) \partial F_2(x_2)}. \quad (3.2)$$

The joint *pdf* denoted $f_{\mathbf{X}}$ can be obtained as a function of the copula density as

$$f_{\mathbf{X}}(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \prod_{n=1}^2 f_n(x_n), \quad (3.3)$$

where $f_n(x_n)$ is the margin or univariate *pdf* corresponding to $F_n(x_n) = u_n$, $n \in \{1, 2\}$, which is distributed as *Uniform*(0, 1). The log-likelihood of the joint distribution can be conveniently expressed as

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi}) &= \sum_{t=1}^T \log c(u_{1t}, u_{2t}; \boldsymbol{\theta}) + \sum_{n=1}^2 \sum_{t=1}^T \log f_n(x_n; \boldsymbol{\phi}_n) \\ &= \mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi}) + \sum_{n=1}^2 \mathcal{L}_n(\boldsymbol{\phi}_n), \end{aligned} \quad (3.4)$$

⁸A thorough exposition of copula theory can be found in Nelsen (2006) and financial applications in Cherubini et al. (2004).

where $\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi})$ is the copula log-likelihood with $\boldsymbol{\phi} = (\boldsymbol{\phi}_1, \boldsymbol{\phi}_2)'$, and $\mathcal{L}_n(\boldsymbol{\phi}_n)$, $n = 1, 2$, are the log-likelihoods of the margins. The vector $\boldsymbol{\theta}$ gathers the copula parameters that govern the dependence structure.

Let the random process r_t denote the daily returns of a financial asset which can be characterized by an autoregressive–moving-average (ARMA) model as follows

$$r_t = a_0 + \sum_{i=1}^p a_i r_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t \quad (3.5)$$

where a_0 is a constant; p and q are the order of autoregressive and moving average processes respectively for the conditional mean. The error term ε_t can be splitted into a stochastic part x_t and a time-dependent standard deviation σ_t so that $\varepsilon_t = \sigma_t x_t$. The series σ_t^2 is characterized by Bollerslev (1986)'s generalized autoregressive conditional heteroskedasticity (GARCH) model

$$\sigma_t^2 = c_0 + \sum_{i=1}^r c_i \sigma_{t-i}^2 + \sum_{i=1}^s d_i \varepsilon_{t-i}^2 \quad (3.6)$$

where c_0 is a constant; r and s are the order of the GARCH and ARCH terms of the conditional volatility process. The filtered returns $x_t = \varepsilon_t / \sigma_t$, $t = 1, \dots, T$, follow a strong white noise process with a zero mean and unit variance. In our empirical work, we adopt Hansen (1994)'s skewed Student's t distribution $x_t \stackrel{i.i.d.}{\sim} skT(0, 1; \nu, \zeta)$, with $\nu > 2$ and ζ denoting the degrees of freedom (dof) and asymmetry parameters, respectively. It has the *pdf*⁹

$$f(x; \nu, \zeta) = \begin{cases} bc \left(1 + \frac{1}{\nu-2} \left(\frac{bz+a}{1-\zeta} \right)^2 \right)^{-\frac{\nu+1}{2}} & , \text{ if } z < -\frac{a}{b} \\ bc \left(1 + \frac{1}{\nu-2} \left(\frac{bz+a}{1+\zeta} \right)^2 \right)^{-\frac{\nu+1}{2}} & , \text{ if } z \geq -\frac{a}{b} \end{cases} \quad (3.7)$$

where $a = 4\zeta c^{\frac{\nu-2}{\nu-1}}$, $b^2 = 1 + 3\zeta^2 - a^2$, $c = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)\Gamma(\frac{\nu}{2})}}$. The skewed Student's t distribution is quite general as it nests the Student's t distribution ($\zeta = 0$) and the Gaussian density ($\zeta = 0$, $\nu \rightarrow \infty$). Previous studies advocate this parametrization for the mar-

⁹There are other skewed Student's t distributions that the skewness is introduced in different ways, see Fernandez and Steel (1998) and Aas and Haff (2006).

gins as able to capture the autocorrelation, volatility clustering, skewness and heavy tails exhibited typically by financial asset returns; see e.g. Jondeau and Rockinger (2006) and Kuester et al. (2006). The latter study highlights, in particular, the excellent forecasting results from GARCH-type models based on the skewed Student's t for Value-at-Risk applications. Equations (3.5)–(3.7) represent an ARMA-GARCH-skT model. In our empirical work, we fix $r = s = 1$, and select the best p and q among $1, 2, \dots, 10$ by minimizing the Akaike Information Criterion (AIC). The model parameters ϕ_n are estimated by quasi-maximum likelihood (QML). *Uniform*(0, 1) margins denoted $u_n = F_n(x_n)$, $n = 1, 2$, can be obtained from each filtered return series via the probability integral transform. Once the vector $\mathbf{u} = (u_1, u_2)'$ is formed, the copula parameter vector can be estimated by maximizing the corresponding copula log-likelihood function $\mathcal{L}_C(\boldsymbol{\theta}, \boldsymbol{\phi})$. Further discussion on estimation can be found below.

Copulas can be broadly grouped as elliptical (*e.g.*, Gaussian and Student's t) and Archimedean (*e.g.*, Gumbel, Clayton and symmetrized Joy-Clayton denoted SJC). Unlike the Gaussian copula which is solely parametrized by the linear Pearson's correlation ρ , the Student's t copula can capture extreme return comovements via the so-called *tail dependence* parameter which is determined by the dof parameter ν alongside ρ ; the smaller ν , the more prominent the tail dependence or clustering of extreme returns. The key advantage of elliptical copulas is tractability since they can be easily extended from bivariate to high-dimensional settings, but their main shortcoming is that they impose symmetry. Archimedean copulas can additionally capture asymmetric tail dependence. Gumbel (Clayton) copula describes upper (lower) tail dependence but, by rotation, the opposite tail can be modeled.¹⁰ The SJC copula can model asymmetrically the dependence structure at both tails and hence, enables tests of symmetry. See Appendix 3.C for further details on the copula functions. The log-likelihood functions of the copulas can be found in Appendix (3.D). In early financial applications, the above copulas were mainly deployed in *static* settings.

¹⁰Rotation is by 180° degrees, namely, the copulas are formulated on minus the random variables; see Cherubini et al. (2004).

3.2.2 Dynamic copulas

In a dynamic context the copula parameters are estimated conditionally (*i.e.*, allowed to time-vary) and so are the *rank correlation* and *tail dependence* measures implied from them; see Appendix 3.C. Patton (2006) sets the foundations for time-varying copulas by proving Sklar's theorem for conditional distributions, and suggests to allow the generic copula dependence parameter θ evolves in ARMA fashion as follows

$$\theta_t = \Lambda(\omega + \varphi\theta_{t-1} + \psi\Gamma_t) \quad (3.8)$$

which permits mean-reversion in dependence. The forcing variable Γ_t is defined as

$$\Gamma_t = \begin{cases} \frac{1}{m} \sum_{j=1}^m F_1^{-1}(u_{1,t-j}) F_2^{-1}(u_{2,t-j}) & \text{elliptical} \\ \frac{1}{m} \sum_{j=1}^m |u_{1,t-j} - u_{2,t-j}| & \text{Archimedean} \end{cases}$$

where $F_n^{-1}(u_{n,t})$, $n = 1, 2$ is the inverse *cdf* of the margins and $\Lambda(\cdot)$ is the (modified) logistic or exponential function.¹¹ We use $m = 10$ as in Patton (2006).

Engle (2002)'s dynamic conditional correlation (DCC) model inspired the copula formulation¹²

$$\begin{aligned} \mathbf{Q}_t &= (1 - \varphi - \psi) \bar{\mathbf{Q}} + \varphi \mathbf{Q}_{t-1} + \psi \epsilon_{t-1} \cdot \epsilon'_{t-1}, \quad \varphi + \psi < 1; \varphi, \psi \in (0, 1) \\ \mathbf{R}_t &= \tilde{\mathbf{Q}}_t^{-1} \mathbf{Q}_t \tilde{\mathbf{Q}}_t^{-1} \end{aligned} \quad (3.9)$$

where all matrices are 2×2 in our bivariate setting; $\bar{\mathbf{Q}}$ is the unconditional covariance of $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t})'$ estimated as $\bar{\mathbf{Q}} = T^{-1} \sum_{t=1}^T \epsilon_t \epsilon'_t$ with $\epsilon_{1,t} \equiv F_1^{-1}(u_{1,t})$ and $\epsilon_{2,t} \equiv$

¹¹In the context of elliptical copulas, the dynamic parameter is the conventional correlation measure, $\theta_t = \rho_t$, and $\Lambda(y) = (1 - e^{-y})(1 + e^{-y})^{-1}$ is the modified logistic transformation to ensure $\rho_t \in (-1, 1)$. In the Gumbel copula $\theta_t = \eta_t$ and $\Lambda(y) = e^y$ to ensure $\eta_t \in (0, \infty)$ as defined in Appendix 3.C. Once these dynamic parameters are estimated, they can be mapped into time-varying rank-correlation and tail dependence measures, $\hat{\tau}_t$ and $\hat{\lambda}_t$, using the formula tabulated in Appendix 3.C. In the SJC copula, the parameter modeled in (3.8) is directly the upper tail dependence, $\theta_t = \lambda_t^U$ (or lower tail dependence $\theta_t = \lambda_t^L$), and $\Lambda(y) = (1 + e^{-y})^{-1}$ is the logistic transformation to ensure $\lambda_t^U, \lambda_t^L \in (0, 1)$.

¹²The popular DCC model put forward by Engle (2002) can be cast as a Gaussian DCC copula. Jondeau and Rockinger (2006) model the dependence between international equity indices using elliptical DCC copulas.

$F_2^{-1}(u_{2,t})$; $\mathbf{Q}_t$ is the conditional covariance matrix; $\tilde{\mathbf{Q}}_t$ is a diagonal matrix with elements the square root of $\text{diag}(\mathbf{Q}_t)$; and $\mathbf{R}_t$ is a correlation matrix with off-diagonal element ρ_t which (for elliptical copula) relates to Kendall's τ_t as shown in Appendix 3.C.

Both ARMA and DCC formulations have as common aspects: i) characterizing the dependence dynamics as ‘autoregressive’ type, and ii) nesting *static* copulas under the restriction $\varphi = \psi = 0$. But they have different merits. The DCC copula formulation can be easily extended to multivariate contexts which is rather challenging with the ARMA formulation. On the other hand, the DCC formulation is not straightforward to apply to non-elliptical copulas; see Manner and Reznikova (2010), for further comparative discussion.

3.2.3 Regime-switching dynamic copulas

We propose flexible Markov-switching (RS) copula models which accommodate dynamic dependence within each regime and hence, can capture regime-specific mean reversion. This feature represents a distinction from conventional RS copulas where a static copula function is assumed to govern each regime regardless of how long the given state prevails. Extant studies typically associate the *low* dependence regime with Gaussian copula which assumes zero tail dependence and the *high* dependence regime with non-Gaussian copula that permits tail dependence. For instance, in the RS copula formulated by Rodriguez (2007) and Okimoto (2008) the means, variances and correlations switch together and each regime is dictated by a distinct static copula. In a similar vein, Chollete et al. (2009) and Garcia and Tsafack (2011) deploy RS dependence models which are parametrized by static Gaussian copula in the normal regime regime and a mixture of static elliptical/Archimedean copulas in another regime.

In order to outline our regime-switching (RS) copula framework, let S_t be a state variable that dictates the prevailing regime. The joint distribution of X_{1t} and X_{2t}

conditional on being in regime s is defined as

$$(X_{1t}, X_{2t} \mid X_{1,t-1}, X_{2,t-1}; S_t = s) \sim C_t^{S_t} (u_{1t}, u_{2t} \mid u_{1,t-1}, u_{2,t-1}; \boldsymbol{\theta}_t^{S_t})$$

with $s \in \{H, L\}$ where H denotes the *high* dependence regime and L the *low* dependence regime. The random variable S_t follows a Markov chain of order one characterized by the transition probability matrix

$$\boldsymbol{\pi} = \begin{pmatrix} \pi_{HH} & 1 - \pi_{HH} \\ 1 - \pi_{LL} & \pi_{LL} \end{pmatrix} \quad (3.10)$$

where π_{HH} and π_{LL} are the so-called staying probabilities, namely, π_{HH} (π_{LL}) is the probability of being in the *high* (*low*) dependence regime at time t conditional on being in the same regime at $t - 1$.

First, we propose a regime-switching ARMA copula where the dependence structure evolves as follows

$$\theta_t^{S_t} = \Lambda \left(\omega^{S_t} + \varphi \theta_{t-1}^{S_{t-1}} + \psi \Gamma_t \right) \quad (3.11)$$

in each regime, with Γ_t and $\Lambda(\cdot)$ defined as in Section 3.2.2. We call this novel formulation RS-ARMA to distinguish it from conventional RS formulations where a static copula governs each regime.

Second, we propose a regime-switching DCC (RS-DCC) dependence model where the time-varying copula function that governs each regime is of DCC type, formalized as¹³

$$\mathbf{Q}_t^{S_t} = (1 - \varphi^{S_t} - \psi^{S_t}) \bar{\mathbf{Q}} + \varphi^{S_t} \mathbf{Q}_{t-1}^{S_{t-1}} + \psi^{S_t} \epsilon_{t-1} \cdot \epsilon'_{t-1}, \quad \varphi^{S_t} + \psi^{S_t} < 1; \quad \varphi^{S_t}, \psi^{S_t} \in (0, 1) \quad (3.12)$$

$$\mathbf{R}_t^{S_t} = \left(\tilde{\mathbf{Q}}_t^{S_t} \right)^{-1} \mathbf{Q}_t \left(\tilde{\mathbf{Q}}_t^{S_t} \right)^{-1}$$

with $\mathbf{Q}_t^{S_t}$ the auxiliary matrix driving the rank correlation dynamics.

¹³The RS-DCC copula can be seen as a generalization of Billio and Caporin (2005)'s regime-switching DCC model.

In our empirical analysis below, the RS-ARMA and RS-DCC models employ the same copula function (*e.g.*, Gumbel) for all regimes but allow for time-variation (mean reversion) in dependence and tail dependence within each regime. Put differently, the RS-ARMA and RS-DCC copulas are flexible enough to capture abrupt increases (decreases) in dependence as financial markets enter crisis (tranquil) regimes without imposing the restriction of static within-regime dependence. If there is only one regime (*i.e.*, $\pi_{HH} = \pi_{LL} = 1$) the RS-ARMA and RS-DCC copula collapse, respectively, to *dynamic* ARMA and DCC copulas formalized as Eq. (3.8) and Eq. (3.9), respectively. Conventional RS copulas collapse instead to *static* copulas.

3.2.4 Estimation of copula parameters

We employ *canonical maximum likelihood* (CML) estimation to obtain the copula parameters. CML is similar in spirit to the *inference functions for margins* (IFM) method where the parameters of the marginal distributions are separated from each other and from those of the copula, and then multi-step ML estimation is applied. In the first step, the parameters of the margins are estimated via univariate ML. In the second step, we estimate by ML the parameters of the copula conditional on the step-one margins.¹⁴ One main advantage of CML (versus IFM) is that, by exploiting the observed empirical distributions, it avoids having to specify *a priori* the margins. More specifically, CML relies on the concept of empirical marginal transformation which approximates an unknown parametric margin with the (uniform) empirical distribution function $\hat{u}_{1t} = \hat{F}_1(x_{1t}) = \frac{1}{T} \sum_{t=1}^T 1_{\{X_{1t} \leq x_{1t}\}}$, and likewise for $\hat{u}_{2t} = \hat{F}_2(x_{2t})$, where (x_{1t}, x_{2t}) , $t = 1, \dots, T$, are the filtered returns. The CML estimator of the

¹⁴Two-step estimation yields asymptotically efficient and multivariate normal parameters $\sqrt{T}(\hat{\theta}_{IFM} - \theta_0) \rightarrow N(0, \mathcal{V}^{-1}(\theta_0))$ where $\mathcal{V}(\theta_0) = \mathbf{H}^{-1} \mathbf{M} (\mathbf{H}^{-1})'$ is the Godambe Information Matrix defined as the covariance matrix of a parameter vector's estimation error when the distribution is non-Gaussian. Given the score function $g(\theta) = \left(\frac{\partial \mathcal{L}_{x_1}}{\partial \theta_1}, \frac{\partial \mathcal{L}_{x_2}}{\partial \theta_2}, \frac{\partial \mathcal{L}_C}{\partial \theta_C} \right)$, then the expectation of the log-likelihood Hessian, $\mathbf{H}^{-1} = \mathbf{E} \left(\frac{\partial g(\theta)}{\partial \theta} \right)$, and the covariance of the log likelihood scores, $\mathbf{M} = \mathbf{E} (g(\theta) g(\theta)')$, can be evaluated numerically at the optimum.

copula parameters is defined as

$$\hat{\boldsymbol{\theta}} \equiv \arg \max_{\boldsymbol{\theta}} \sum_{t=1}^T \mathcal{L}_C \left(c \left(\hat{u}_{1t} = \hat{F}_1(x_{1t}), \hat{u}_{2t} = \hat{F}_2(x_{2t}) \right), \boldsymbol{\theta} \right)$$

which amounts to a ML estimator conditional on the empirical margins.¹⁵

Estimation of the RS copula parameters requires inferences on the probabilistic evolution of the state variable S_t . Probability estimates based on information up to time t are called “filtered probabilities” and those based on full-sample information are “smoothed probabilities”. Our estimation approach builds on Hamilton (1989)’s *filtering* algorithm and Kim (1994)’s *smoothing* algorithm; see Appendix 3.E.

3.3 Data description

Our analysis is based on daily midpoint closing CDS spread quotes at 5-year maturity from *Bloomberg* on three indices: Markit iTraxx Europe, Markit iTraxx Europe Subordinated Financials (SubFin) and Markit iTraxx Europe Autos (Auto). As tradeable systematic equity factors we employ, respectively, the Dow Jones Stoxx Europe 600 index, which comprises the 600 largest market capitalized companies in Europe, the Stoxx Europe 600 Financials index and the Stoxx Europe 600 Automobiles & Parts index. We focus on the cost of insuring against default on automotive companies’ debt as this sector has been severely hit by the recent financial crisis; see 2000s crisis timeline in Appendix 3.F. Finally, we employ the option-implied Dow Jones Euro VStoxx 50 index as proxy for the unobservable asset volatility. Implied volatility is based on market prices of options on the Dow Jones Stoxx Europe 50 and is a forward-looking volatility indicator as it reflects traders’ expectations about future movements of the underlying equity index. The time period is September 9, 2005 to

¹⁵Since CML makes no *a priori* assumption on the parametric form of the margins, it provides superior fit to IFM when the margins are misspecified. Another estimation approach is *Exact Maximum Likelihood* (EML) which provides the parameters of the copula and the margins simultaneously by maximizing the full likelihood. This is computationally rather burdensome. For a detailed comparison of EML, IFM and CML, see Cherubini et al. (2004).

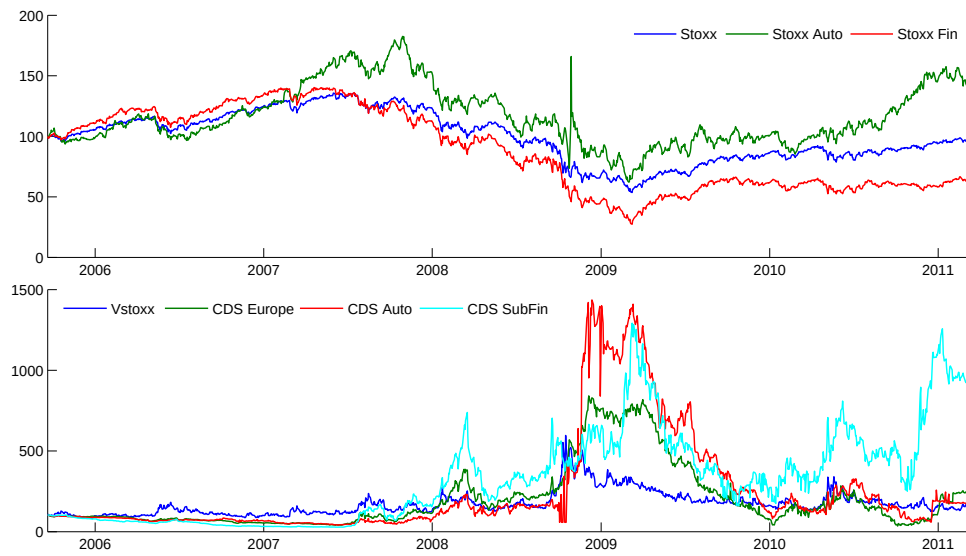
March 11, 2011 making a total of $T = 1,380$ days.¹⁶

Figure 3.1 plots for each index the daily prices (Panel A) and daily logarithmic returns (Panel B). CDS and VStoxx indices move in tandem while Stoxx indices move in the opposite direction. September 2007 marks the start of a steady downward trend in equity prices, attaining the lowest level in 2009, coupled with increased volatility and a steady rise in default risk premiums. Table 3.1 provides summary statistics. CDS SubFin has the highest mean return of 0.17%. Both Figure 1 and Table 1 reveal that CDS indices are notably more volatile than equity indices, and CDS Auto is by and large the most volatile. The Jarque-Bera test confirms that daily returns are non-Gaussian. The Ljung-Box Q test and Engle's ARCH LM test provide strong evidence, respectively, of serial dependence and heteroskedasticity in daily returns. Both the correlation parameter ρ and Kendall's rank correlation parameter τ suggest that CDS returns are negatively (positively) associated with equity returns (volatility), in line with Merton (1974)'s model predictions: i) growth in firm value reduces the probability of default, and ii) higher equity volatility implies a larger probability that the value of assets drops below the level of liabilities, triggering default.

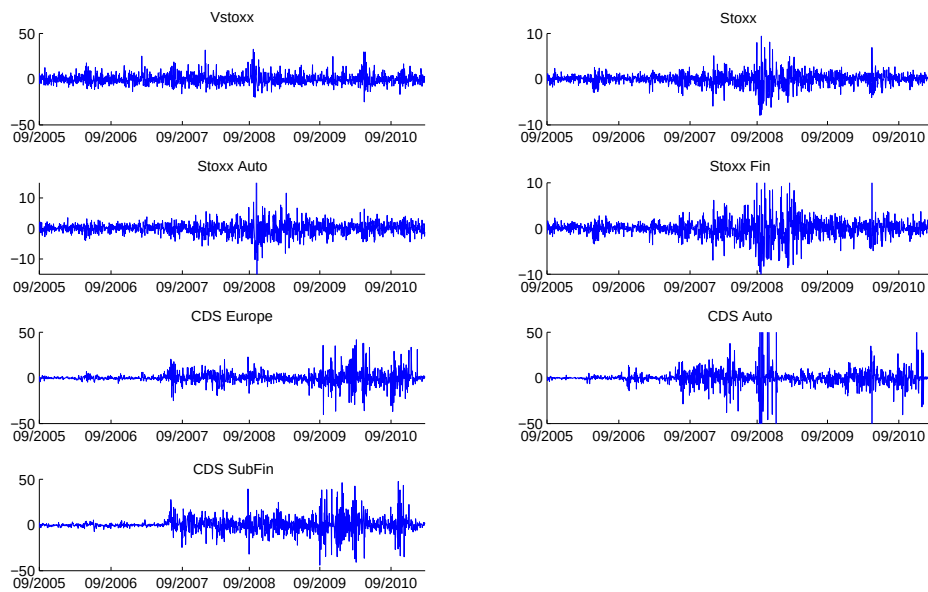
Table 3.2 reports parameter estimates and diagnostics of the (margins) ARMA-GARCH-skT models. The dof parameter ν of the skT density fitted to the standardized residuals reveals substantial leptokurtosis; CDS index returns show fatter tails than the underlying Stoxx and VStoxx index returns. The asymmetry coefficient ζ of the skT residual distribution is significant for VStoxx, Stoxx, Stoxx Fin and CDS SubFin index returns. The Kolmogorov-Smirnov (KS) test, with p -values reported in

¹⁶Markit iTraxx Europe comprises the 125 equally-weighted most liquid names in the European market. iTraxx SubFin is composed of the 25 most liquid CDS names on subordinated debt issued by companies, *e.g.* ABN, Aegon, Allianz, AVIVA, AXA, Barclays, Deutsche Bank, Swiss Re, RBS, Zurich, and iTraxx Auto comprises the 10 most liquid CDS names in the automotive industry, BMW, Volvo, Michelin, Continental, DaimlerChrysler, GKN, Peugeot, Renault, Valeo, Volkswagen. Markit iTraxx indices trade at 3, 5, 7 and 10-year maturities and are reviewed every 6 months in March and September to form a new series (for each maturity) that reflects changes in liquidity as determined by polls of the leading CDS dealers while the old series continues trading; we use Markit Series 4 which was introduced in September 2005. Index trading for 5-year maturity is the most liquid. Markit calculates the official mid-day (11am GMT) and closing (4pm GMT) levels for iTraxx indices on a daily basis. Stoxx Europe 600 Financials contains 23 out of the 25 entities that conform the iTraxx SubFin, while Stoxx Europe 600 Automobiles & Parts includes 9 out of the 10 entities in iTraxx Autos. Stoxx Europe 600 includes all of the 125 firms in the iTraxx Europe index. Stoxx and VStoxx closing prices are downloaded from *www.stoxx.com*.

Figure 3.1: Daily Time Series Plots of CDS and Equity Indices
 Panel A: Levels



Panel B: Logarithmic Returns (%)



The top graphs plot the daily levels of European equity market indices (Stoxx, Stoxx Auto, Stoxx Fin and Vstoxx) and CDS indices with all series normalized to start at 100. The bottom graphs plot the daily logarithmic returns.

Table 3.1: Descriptive Statistics of Daily Returns

	Vstoxx	Stoxx	Stoxx Auto	Stoxx Fin	CDS Europe	CDS Auto	CDS SubFin
Mean	0.038	-0.005	0.024	-0.036	0.068	0.043	0.166
Median	-0.577	0.067	0.057	0.019	-0.193	-0.021	-0.057
Maximum	32.767	9.410	40.817	14.666	41.745	199.212	47.500
Minimum	-24.919	-7.930	-35.427	-10.179	-40.297	-177.407	-43.987
Std. Dev.	5.878	1.419	2.699	2.042	7.074	13.327	8.266
Skewness	0.899	-0.053	3.192	0.303	0.394	2.610	0.343
Kurtosis	6.432	9.741	97.591	10.081	11.520	106.320	10.418
Number of obs.	1381	1381	1381	1381	1381	1381	1381
Jarque-Bera test	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Ljung-Box(10) test	0.035	0.000	0.000	0.008	0.000	0.000	0.000
ARCH(10) test	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pearson linear correlation ρ :							
Vstoxx	-	-	-	-	0.362	0.148	0.302
Stoxx	-	-	-	-	-0.366	-0.157	-0.345
Kendall's rank correlation τ							
Vstoxx	-	-	-	-	0.304	0.214	0.226
Stoxx	-	-	-	-	-0.352	-0.256	-0.274

Notes: The table presents summary statistics of the logarithmic daily returns expressed in percentage for the volatility (VStoxx) index, equity (Stoxx) indices, and CDS (iTraxx) indices. p -values are reported for the tests. The reported correlation between CDS and Stoxx are for matched Stoxx Europe 600 marketwide, Auto or Financial Services indices.

Panel B of Table 3.2, cannot refute the null hypothesis that the residuals conform to a skT distribution. The standardized ARMA-GARCH-skT residuals, x_t , $t = 1, \dots, T$, are mapped into $Uniform(0,1)$ observations, via the probability integral transform; the resulting residual series $\hat{u}_t = \hat{F}(x_t)$, $t = 1, \dots, T$, are then inputs for copula estimation. In order to establish further the goodness-of-fit of the margins, following Diebold et al. (1998) and Patton (2006) we apply the Ljung-Box Q test to various moments $(\hat{u}_t - \bar{u})^m$, $m \in \{1, 2, 3, 4\}$. The results reported in Panel B suggest that there is no remaining serial dependence.

3.4 Empirical Results

3.4.1 In-sample fit of static, dynamic and regime-switching copulas

We begin this section with a preliminary discussion of the ability of Student's t , Gumbel and SJC copulas to predict in-sample the dependence between CDS returns and tradeable systematic risk factors. We employ the different formulations discussed in Section 3.2: static, dynamic (ARMA and DCC), conventional regime-switching (RS) which nests a static copula in each regime, and finally the RS-ARMA and RS-DCC that we propose. Table 3.3 shows the AIC and log-likelihood (LL) values of the competing models. Student's t copulas, which account for tail dependence in a symmetric way, attain higher LL and lower AIC than the competing Gumbel and SJC copulas, irrespective of the formulation employed (static, dynamic or regime-switching).¹⁷ Albeit in purely *static* formulations, Student's t copula models have been shown to succeed in previous competitions such as that conducted by Breyman et al. (2003) to characterize dependence among FX spot returns. This can be partially attributed

¹⁷Table 3.3 presents the AIC of the best-fitted Gumbel copula among the various Gumbel copulas considered (0° , 90° , 180° and 270° rotated). The CDS-Stoxx dependence is best captured by the 90° Gumbel copula which focuses on lower tail dependence. The CDS-Vstoxx dependence is best described by the 0° Gumbel (non-rotated) copula which models upper tail dependence. These results mirror the evidence of asymmetry reported in Longin and Solnik (2001) suggesting that tails that describe adverse movements are strongly dependent whereas the opposite tails are essentially independent.

Table 3.2: Estimation Results for Marginal Models

	Vstoxx	Stoxx	Stoxx Auto	Stoxx Fin	CDS Europe	CDS Auto	CDS SubFin
Panel A: ARMA-GARCH-skT model estimates							
Conditional mean							
Intercept	0.015 (0.135)	0.062** (0.025)	0.107** (0.040)	0.052* (0.028)	-0.166** (0.049)	-0.143** (0.049)	-0.300** (0.054)
AR1	- -	- -	- -	- -	0.079** (0.032)	- -	- -
Conditional variance							
Intercept	1.685** (0.430)	0.017** (0.007)	0.058** (0.022)	0.019** (0.009)	0.075** (0.029)	0.080* (0.047)	0.070** (0.034)
ARCH1	0.092** (0.017)	0.112** (0.021)	0.115** (0.021)	0.113** (0.021)	0.171** (0.023)	0.216** (0.028)	0.167** (0.021)
GARCH1	0.858** (0.023)	0.881** (0.019)	0.876** (0.018)	0.887** (0.019)	0.829** (0.026)	0.784** (0.036)	0.833** (0.023)
ν	6.885 (1.206)	8.959 (2.090)	8.024 (1.949)	7.685 (1.528)	5.005 (0.485)	3.524 (0.213)	4.806 (0.436)
ζ	0.291** (0.039)	-0.112** (0.034)	-0.005 (0.008)	-0.047* (0.028)	0.030 (0.026)	-0.029 (0.020)	-0.051** (0.023)
Panel B: Goodness-of-fit tests							
1 st Moment	0.242	0.873	0.611	0.551	0.606	0.934	0.194
2 nd Moment	0.948	0.859	0.175	0.147	0.941	1.000	0.738
3 rd Moment	1.000	0.998	0.613	0.980	0.999	1.000	0.621
4 th Moment	1.000	1.000	0.999	0.860	1.000	1.000	0.996
K-S test	0.950	0.071	0.306	0.137	0.275	0.109	0.101

Notes: Panel A reports the parameter estimates for the conditional mean Eq. (3.5) and conditional variance Eq. 3.6; ν and ζ are the degrees-of-freedom and asymmetry parameter of the skewed Student's t (skT) distribution for the innovations. Standard errors are in parentheses. ** and * denote statistical significance at the 5% and 10% levels, respectively. Panel B reports the p -value of the Ljung-Box Q-test on the first four moments of the probability integral transformed series, $(u_t - \bar{u})^m$, $m = \{1, 2, 3, 4\}$, obtained from the standardized filtered returns x_t to assess the null that the transformed uniform, $u_t = F(x_t)$, has no autocorrelation up to 10 lags. The bottom row reports the p -value of the Kolmogorov-Smirnov test to assess the null hypothesis that x_t is skewed Student's t distributed.

to the ability of the Student's t copula to fit well the central part of the joint distribution which more heavily contributes to the log-likelihood. Hence, for simplicity of exposition a large part of the subsequent discussion in this section focuses on inferences from the Student's t copula function. The *dynamic* formulation (ARMA or DCC) clearly provides better in-sample fit than the *static* formulation, irrespective of the underlying copula function employed. The *regime-switching* formulation further enhances the copula's ability to describe the dependence structure of CDS-equity markets. Moreover, allowing the dependence structure to display short-memory within each regime (RS-ARMA or RS-DCC) leads to the lowest AIC and largest LL among the competing formulations.

Table 3.4 reports parameter estimates of static, dynamic and regime-switching Student's t copulas. The correlation parameter ρ of the static copula suggests significantly negative dependence for all CDS and Stoxx pairs, and significantly positive dependence for all CDS and VStoxx pairs in line with Merton (1974)'s theory; a firm's likelihood of default is a decreasing function of asset value proxied by the market value of its equity, and an increasing function of asset volatility proxied by the volatility of its equity returns.¹⁸ The correlation parameter ρ in the static copula formulation and $(\rho^U, \rho^L)'$ in the conventional RS copula reveal that, generally, changes in CDS spreads are more strongly associated with changes in equity returns than with changes in volatility. Furthermore, the CDS and equity return association is more prominent in the SubFin sector than the Auto sector, in line with extant evidence that structural credit risk models are more successful in explaining non-investment grade spreads; see Eom et al. (2004) and Zhang et al. (2009).

The significance of parameters φ and ψ in the dynamic ARMA and DCC formulations bears out that rank correlations are time-varying. The persistence measure $\varphi + \psi$ inferred from the DCC copula suggests that the rank correlation of CDS and equity markets is more persistent at sector level than marketwide; the largest persis-

¹⁸We tested the hypothesis of zero tail dependence by means of a likelihood ratio (LR) test for $H_0 : 1/\nu = 0$ which is essentially a test for the restriction that the dof parameter ν is large enough so that the Student's t copula effectively becomes the Gaussian. The hypothesis is rejected for all pairs except CDS Auto and VStoxx. The AIC and LL criteria also favor the Student's t copula over the Gaussian copula. Details are available from the authors upon request.

Table 3.3: Goodness-of-Fit Measures for Competing Copula Models

		Stoxx CDS Europe	Vstoxx CDS Europe	Stoxx Auto CDS Auto	Vstoxx CDS Auto	Stoxx Fin CDS SubFin	Vstoxx CDS SubFin
Static copulas							
Student's t (Static)	<i>AIC</i>	-379.413	-255.125	-167.542	-121.169	-214.191	-140.361
	<i>LL</i>	191.707	129.563	85.771	62.584	109.095	72.180
Gumbel (Static)	<i>AIC</i>	-362.281	-252.012	-166.509	-122.508	-218.837	-142.973
	<i>LL</i>	182.140	127.006	84.255	62.254	110.418	72.487
SJC (Static)	<i>AIC</i>	-360.039	-244.581	-159.234	-120.913	-216.928	-141.240
	<i>LL</i>	182.020	124.291	81.617	62.456	110.464	72.620
Regime-switching Static Copulas							
Student's t (RS)	<i>AIC</i>	-404.514	-280.493	-193.126	-140.579	-223.487	-146.328
	<i>LL</i>	207.257	145.247	101.563	75.289	116.744	78.164
Gumbel (RS)	<i>AIC</i>	-384.918	-268.826	-187.505	-131.069	-223.458	-143.875
	<i>LL</i>	196.459	138.413	97.752	69.534	115.729	75.937
SJC (RS)	<i>AIC</i>	-381.162	-272.480	-184.326	-124.190	-222.882	-144.416
	<i>LL</i>	196.581	142.240	98.163	68.095	117.441	78.208
Dynamic Copulas							
Student's t (DCC)	<i>AIC</i>	-408.035	-287.043	-188.93	-145.561	-228.643	-149.081
	<i>LL</i>	207.018	146.521	97.467	75.780	117.321	77.540
Student's t (ARMA)	<i>AIC</i>	-405.568	-287.440	-185.540	-141.770	-228.668	-147.691
	<i>LL</i>	206.784	147.720	96.770	74.885	118.334	77.845
Gumbel (ARMA)	<i>AIC</i>	-382.284	-274.081	-180.170	-128.327	-223.250	-140.876
	<i>LL</i>	194.142	140.040	93.083	67.163	114.625	73.438
SJC (ARMA)	<i>AIC</i>	-379.754	-276.367	-179.85	-121.389	-223.074	-145.683
	<i>LL</i>	195.877	144.183	95.925	66.695	117.537	78.842
Regime-switching Dynamic Copulas							
Student's t (RS-DCC)	<i>AIC</i>	-408.117	-287.094	-191.648	-141.984	-229.332	-148.110
	<i>LL</i>	211.058	150.547	102.824	77.992	121.656	81.055
Student's t (RS-ARMA)	<i>AIC</i>	-414.242	-289.265	-192.315	-142.914	-230.319	-149.380
	<i>LL</i>	214.121	151.623	103.158	78.457	122.160	81.690
Gumbel (RS-ARMA)	<i>AIC</i>	-392.916	-278.836	-188.819	-135.819	-226.447	-148.821
	<i>LL</i>	202.458	145.418	100.410	73.909	119.224	80.411
SJC (RS-ARMA)	<i>AIC</i>	-385.553	-281.532	-176.162	-124.429	-226.349	-146.612
	<i>LL</i>	202.776	150.766	98.081	72.214	123.175	83.306

Notes: This table reports the goodness-of-fit measures of Student's t , Gumbel and SJC copulas in a static, regime-switching, dynamic and regime-switching dynamic formulation. AIC denotes Akaike information criterion and LL denotes the optimized log-likelihood. The bold font is used to signify the best case (largest LL or lowest AIC) within each panel.

Table 3.4: Estimation Results for Static, Dynamic and RS Student's t Copula

		Stoxx CDS Europe	Vstoxx CDS Europe	Stoxx Auto CDS Auto	Vstoxx CDS Auto	Stoxx Fin CDS SubFin	Vstoxx CDS SubFin
Static	ρ	-0.495** (0.021)	0.419** (0.022)	-0.347** (0.024)	0.296** (0.025)	-0.381** (0.023)	0.316** (0.025)
	ν	8.943 (1.953)	20.443 (8.970)	22.025 (10.737)	25.052 (17.066)	11.513 (3.747)	15.921 (6.430)
RS	ρ_H	-0.679** (0.033)	0.642** (0.040)	-0.610** (0.050)	0.383** (0.031)	-0.551** (0.048)	0.475** (0.041)
	ρ_L	-0.273** (0.098)	0.194** (0.057)	-0.214** (0.045)	0.173** (0.141)	-0.267** (0.054)	0.196** (0.043)
	ν	27.804 (26.691)	32.015 (18.568)	29.246 (16.734)	36.517 (10.658)	18.553 (9.558)	25.758 (32.067)
	π_{HH}	0.969** (0.020)	0.947** (0.024)	0.981** (0.096)	0.995** (0.034)	0.991** (0.065)	0.993** (0.042)
	π_{LL}	0.966** (0.032)	0.949** (0.020)	0.992** (0.045)	0.980** (0.018)	0.995** (0.031)	0.995** (0.026)
DCC	φ	0.880** (0.038)	0.829** (0.030)	0.936** (0.027)	0.940** (0.026)	0.960** (0.012)	0.972** (0.011)
	ψ	0.057** (0.016)	0.078** (0.015)	0.039** (0.015)	0.034** (0.013)	0.022** (0.008)	0.015** (0.006)
	ν	10.299 (2.535)	29.609 (23.391)	22.854 (11.297)	39.324 (21.525)	13.465 (5.067)	18.122 (8.562)
ARMA	ω	-0.087 (0.063)	0.297** (0.141)	-0.017 (0.015)	0.014 (0.010)	-0.002 (0.022)	0.010 (0.011)
	φ	1.809** (0.192)	0.956** (0.529)	1.941** (0.073)	1.941** (0.059)	2.030** (0.090)	2.609** (0.056)
	ψ	0.211** (0.067)	0.491** (0.189)	0.105** (0.043)	0.105** (0.034)	0.082** (0.033)	0.045** (0.021)
	ν	10.550 (2.680)	31.157 (27.071)	22.010 (10.203)	41.749 (72.621)	14.316 (5.600)	18.221 (8.520)
RS-DCC	φ_H	0.346 (0.245)	0.646** (0.204)	0.371 (0.301)	0.422 (1.804)	0.383 (2.164)	0.520 (0.527)
	ψ_H	0.011 (0.031)	0.350** (0.151)	0.015 (0.096)	0.019 (0.045)	0.011 (0.014)	0.109 (0.223)
	φ_L	0.872** (0.038)	0.841** (0.040)	0.933** (0.039)	0.950** (0.022)	0.958** (0.013)	0.971** (0.013)
	ψ_L	0.065** (0.017)	0.069** (0.022)	0.054* (0.030)	0.040** (0.013)	0.027** (0.009)	0.017** (0.007)
	ν	10.682 (2.831)	35.068 (17.996)	32.453 (16.179)	46.957 (7.457)	14.115 (5.911)	19.162 (9.233)
	π_{HH}	0.927** (0.107)	0.933** (0.066)	0.958** (0.050)	0.978** (0.014)	0.926** (0.060)	0.941** (0.065)
	π_{LL}	0.990** (0.018)	0.997** (0.006)	0.978** (0.019)	0.956** (0.034)	0.983** (0.027)	0.994** (0.007)
RS-ARMA	ω_H	-1.839** (0.616)	1.471** (0.563)	-0.194** (0.055)	0.192* (0.106)	-1.548** (0.361)	0.470 (0.408)
	ω_L	-0.076* (0.043)	0.199* (0.110)	-0.002 (0.008)	0.050** (0.021)	-0.004 (0.026)	0.185 (0.165)
	φ	1.831** (0.194)	1.224** (0.579)	2.026** (0.053)	1.488** (0.585)	2.021** (0.103)	1.056* (0.605)
	ψ	0.190** (0.066)	0.399** (0.198)	0.021 (0.017)	0.118* (0.068)	0.078** (0.035)	0.098 (0.136)
	ν	12.858 (3.958)	33.443 (19.219)	31.794 (16.295)	47.260 (18.302)	14.695 (7.019)	18.590 (10.030)
	π_{HH}	0.900** (0.064)	0.900** (0.076)	0.979** (0.014)	0.996** (0.004)	0.901** (0.088)	0.993** (0.004)
	π_{LL}	0.996** (0.004)	0.994** (0.009)	0.993** (0.004)	0.984** (0.017)	0.998** (0.003)	0.996** (0.002)

Notes: This table gives the estimates of Student's t copula models in various formulations: static, regime-switching static, two dynamic copulas (DCC and ARMA) and two regime-switching extensions (RS-DCC and RS-ARMA). Superscript H (L) indicates the *high* (*low*) dependence regime. π_{HH} (π_{LL}) is the probability of staying in the *high* (*low*) dependence regime. Numbers in parentheses are standard errors. ** and * denote significance at the 5% and 10% levels, respectively.

tence at 0.982 (Stoxx) and 0.987 (VStoxx) corresponds to the subordinated financial sector.

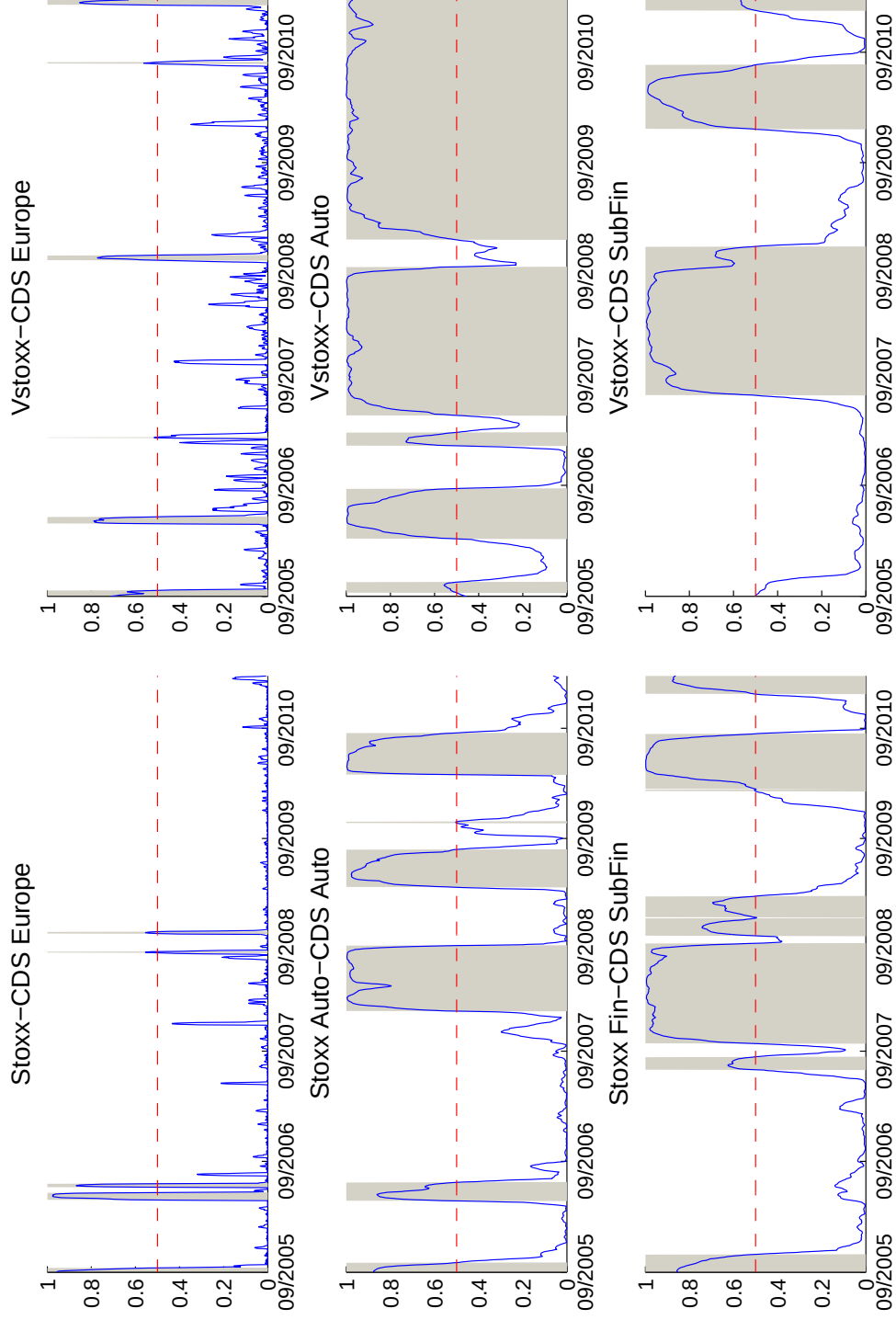
Turning attention to the regime-switching models, the probabilities π_{HH} and π_{LL} consistently suggest slightly longer duration of the low dependence regime during our sample period. The statistical significance of two dependence regimes can be tested by means of a LR test for the null hypothesis $H_0 : \omega_H = \omega_L$ which states that the RS-ARMA copula has one regime, then becoming the ARMA copula. An analogous LR test is formulated as $H_0 : \rho_H = \rho_L$ with the conventional RS copula, then becoming the static copula. The traditional asymptotic theory for these LR test statistics does not apply because of the nuisance parameter problem (*i.e.*, unidentified parameters under the null such as the residual variance in each regime). However, the conventional p -values are very small, all below 0.008, providing *prima facie* evidence of regime-switching effects. Parameters ρ_H and ρ_L of the RS copula confirm that both in “crisis” and “normal” regimes the level of dependence is stronger for CDS returns with equity returns than for CDS returns with equity volatility. Regarding the degree of dependence persistence, we learn from the RS-DCC copula (parameters $\varphi + \psi$) that in “crisis” regimes shocks to dependence between CDS returns and equity volatility die more slowly than shocks to dependence between CDS returns and equity returns.

Figure 3.2 plots the smoothed probabilities of the *high* dependence regime in the RS-ARMA copula.¹⁹ In both sectors, Auto and SubFin, the dependence between CDS and equity markets enters a *high* or “crash” regime by late 2007 reflecting the onset of the credit crunch, and lingers on for about a year.²⁰ Although the global credit crisis originated from the huge losses of subprime CDS investment in the financial sector, the automotive industry was badly hit by various illiquidity shocks, a sharp fall in consumer confidence and soaring oil prices; Appendix 3.F provides a snapshot of the 2000s crisis timeline. After a short pause, both sectors enter again the *high* dependence regime by late 2009 possibly reflecting the breakout of the European sovereign

¹⁹Inferences from the RS-DCC Student’s t copula are qualitatively similar and therefore not reported to save space.

²⁰According to the Eurocoin indicator, the eurozone was in recession from March 2007 to February 2009. According to the NBER business cycle indicator, the US was in recession from December 2007 to June 2009.

Figure 3.2: Smoothed Probability of High Dependence Regime



The graphs show the smoothed probability of *high* dependence regimes for pairs of CDS indices with equity market Stoxx and Vstoxx indices inferred from the RS-ARMA-Student's t copula.

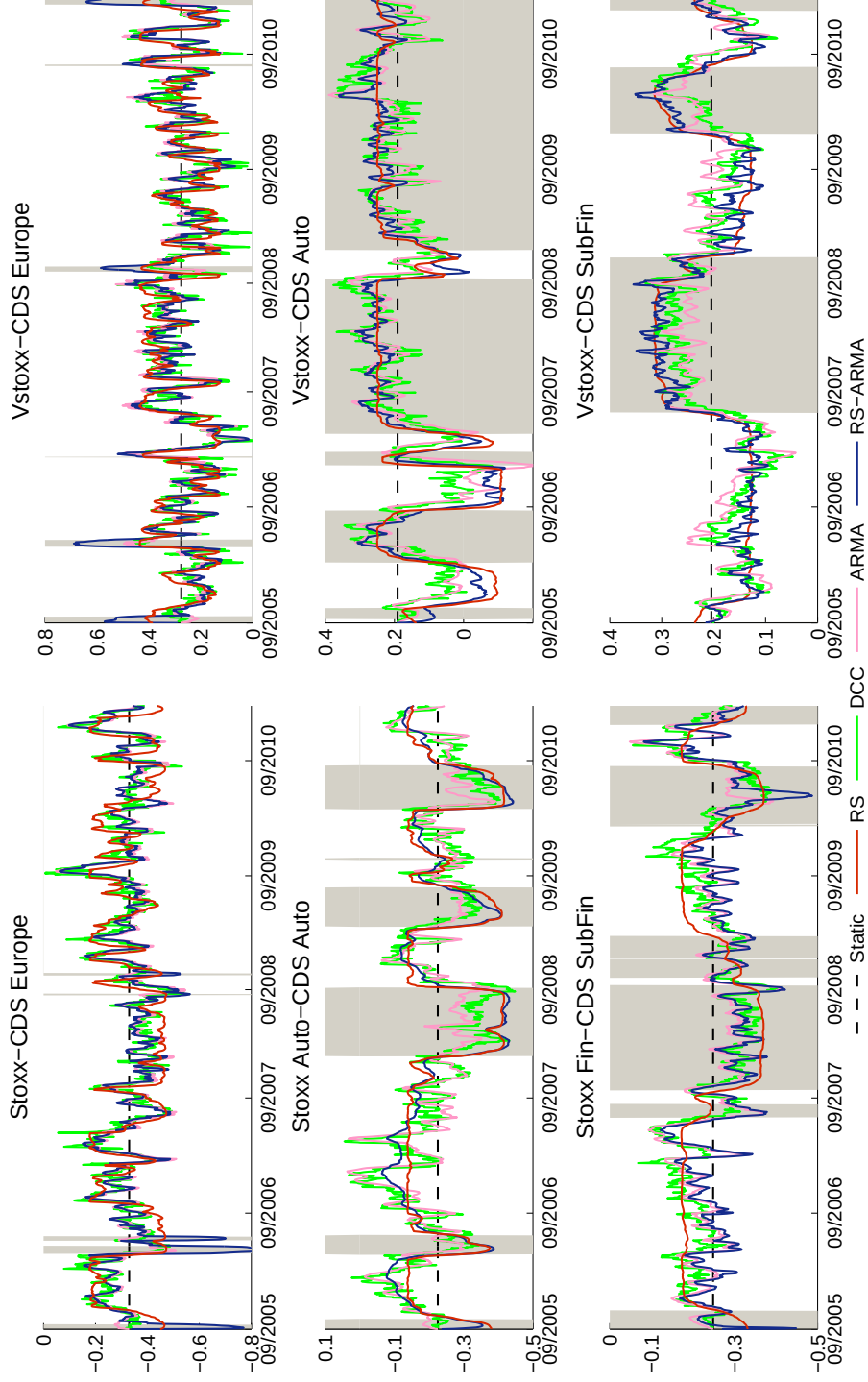
debt crisis. Fung et al. (2008) also document a significant increase in dependence between the US equity and CDS markets during the 2007 credit crunch on the basis of linear vector autoregressions and common Pearson correlation coefficients. Our findings from flexible copula models extend their evidence to the European context.²¹ Altogether for the Auto and SubFin sectors, four transitions into the *high* dependence regime are identified in Figure 3.2. One in 2005 roughly coinciding with the downgrade of two big players in the auto industry (Ford and GM), another in 2006 reflecting the deterioration of the US housing market, a third entry in 2007 reflecting the credit crunch, and a fourth entry in 2009 concurrent with the European debt crisis. The unsuccessful regime identification in the marketwide CDS-equity copulas possibly reflects the fact that diversification, *i.e.* pooling of entities from diverse sectors with different timings of transition between *high* and *low* dependence states, reduces the overall signal-to-noise ratio.

The Kendall's *rank correlation* τ inferred from various copula formulations is plotted in Figure 3.3. Several observations can be made. First, the degree of dependence clearly varies over time. Second, the RS and RS-ARMA copula suggest upward shifts in dependence between CDS and equity markets at economically plausible time points. For instance, the sudden downgrade by S&P's of two important car manufacturers, GM and Ford, from BBB to BB in May 2005 and to B in December 2005 led to a dramatic increase in CDS and equity dependence for the Auto and marketwide indices which the ARMA copulas tend to smooth out. Crude oil prices reached historically high levels of over \$77 per barrel in July 2006 which pushed the CDS-Stoxx (Auto) dependence to a *high* regime; again this pattern is better captured by the RS and RS-ARMA copulas. For the SubFin CDS-equity pairs, the most dramatic increase in dependence roughly occurs in late 2009 when a credit rating downgrade from A- to BBB+ is announced by S&P's for Greece.

The *tail dependence* parameter λ inferred from the different copula formulations is plotted in Figure 3.4. We can see evidence of *high* and *low* tail-dependence regimes

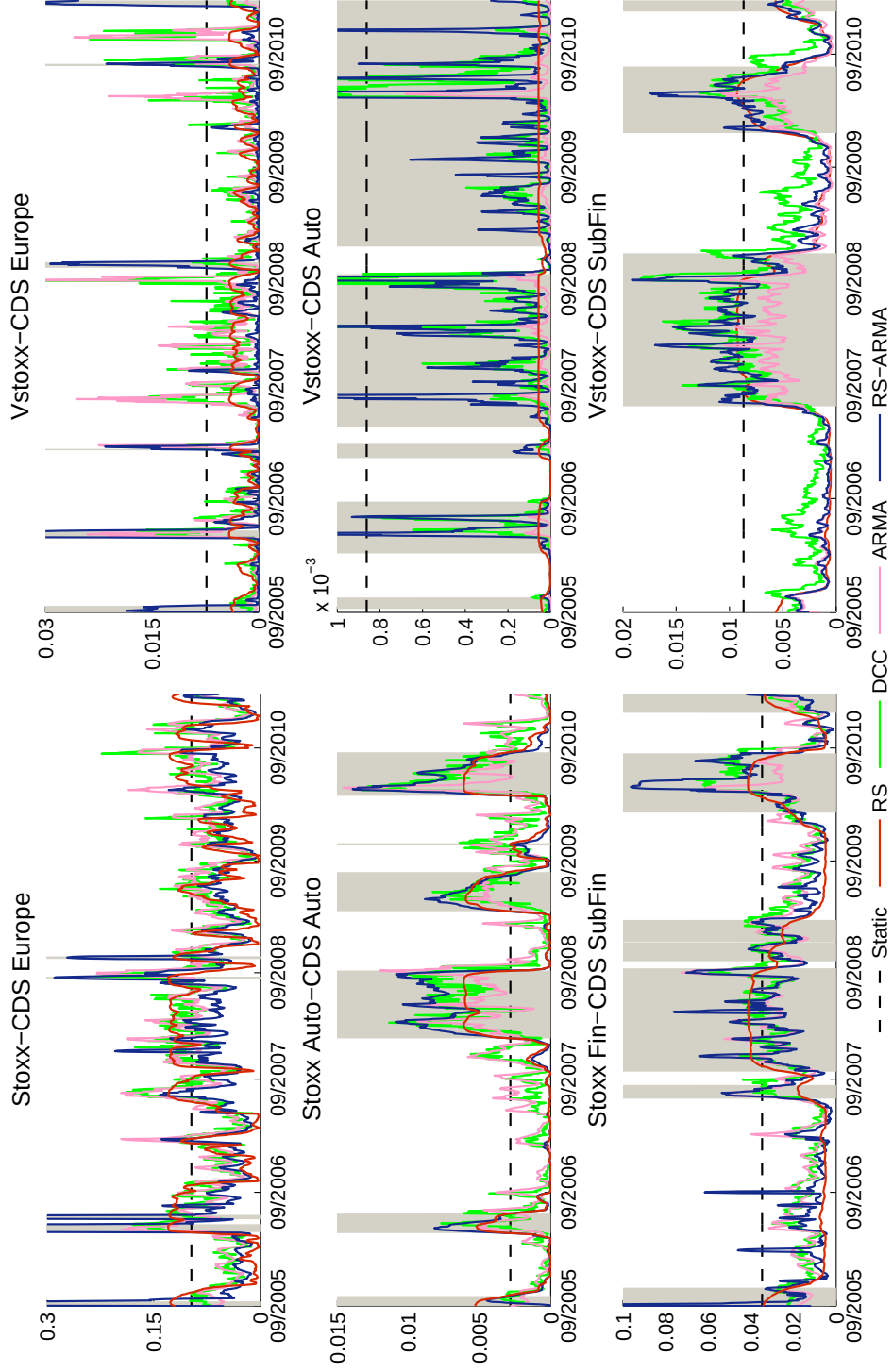
²¹The relatively long *high* dependence regime identified for CDS Auto and VStoxx is not surprising given that the European auto sector has suffered various setbacks in recent years, *e.g.* energy crisis and sharp fall in consumer demand.

Figure 3.3: Rank Correlation Estimates



The graphs show the dynamics of *rank correlation* (Kendall's τ) estimated by Student's t copula for pairs of Stoxx-CDS indices (left column) and Vstoxx-CDS indices (right column). For RS and RS-ARMA copula models, the graphs show the weighted average $\tau_t^H p_t^H + \tau_t^L p_t^L$ where the weights are given by the smoothed probability of each regime, *i.e.* $p_t^s = P[S_t = s | I_T]$, $s \in \{H, L\}$, computed as detailed in Appendix 3E.

Figure 3.4: Tail Dependence Estimates



The graphs show the dynamics of *tail dependence* (λ) estimated by Student's t copula for pairs of Stoxx-CDS indices (left column) and Vstoxx-CDS indices (right column). For RS and RS-ARMA copula models, the graphs show the weighted average $\lambda_t^H p_t^H + \lambda_t^L p_t^L$ where the weights are given by the smoothed probability of each regime, *i.e.* $p_t^s = P[S_t = s | I_T]$, $s \in \{H, L\}$, computed as detailed in Appendix 3.E.

which reflect the presence of two different CDS-equity bivariate distributions corresponding, respectively, to crisis and normal episodes. The intuition behind this finding is that CDS spreads react more vigorously to ‘extreme’ bad news in crises than in normal periods, namely, the degree of tail dependence exacerbates during periods of market stress. While the tail dependence estimates may seem small, they are broadly aligned with those in Garcia and Tsafack (2011) for European equity-bond pairs and with those in Jondeau and Rockinger (2006) for cross-country equity market pairs. The *high* tail dependence regime is most apparent for CDS SubFin and Stoxx Fin returns which may indicate that financials are particularly sensitive to extreme market events. Regardless of the level of tail dependence, the graphs endorse the RS-ARMA copula as more adequate for capturing sudden shifts in tail dependence and confirm the biases arising from the use of non-regime-switching copula, which tend to smooth out the degree of dependence over time. The latter effectively implies overestimation of the extent of dependence in normal periods and underestimation during crisis periods. The upshot is that using an implausible model of asset dependence that does not permit sudden changes of regime or that constrains the within-regime dependence to be constant could be costly from a risk management perspective. This question is addressed in the next section.

3.4.2 Out-of-sample copula forecasts for risk management

The economic value of the proposed *regime-switching dynamic* copulas is demonstrated via a Monte Carlo simulation to set 1-day-ahead Value at Risk (VaR) trading limits for portfolios of equity and CDS instruments.²² Since the 1996 Market Risk Amendment (MRA) to the Basel Accord, the VaR measure has played a central role in regulatory capital assessments and remains one of the most common portfolio risk control tools in banks and insurance firms. The MRA stipulates that banks should internally compute VaR on a daily basis for backtesting purposes although regulators

²²In this chapter, we chose as evaluation method for the out-of-sample copula forecasts the economic loss function implicit in VaR backtesting together with explicit regulatory-driven loss functions. However, other purely statistical methods are feasible to compare the accuracy of out-of-sample copula density forecasts; *e.g.*, see Diks and van Dijk (2010).

(usually, the central bank) requires 10-day-ahead VaR to be reported for establishing the minimum capital requirement, possibly to mitigate the costs of too frequent monitoring. The reason why the prescribed horizon for backtesting purposes is 1 day is that it increases the number of observations; in 1-day VaR the number of observations is 252 per year whereas with 2-week VaR it reduces to 26. By now the commercial banking industry has settled on the 1-day horizon.

The 1-day-ahead VaR is an α -quantile prediction of the future portfolio profit and loss (P/L) distribution. It provides a measure of the maximum future losses over a time span $[t, t+1]$, which can be formalized as

$$P [R_{t+1} \leq VaR_{t+1}^\alpha | I_t] = \alpha \quad (3.13)$$

where R_{t+1} denotes the portfolio return on day $t + 1$, and I_t is the information set available on day t . The nominal coverage $0 < \alpha < 1$ is typically set at 0.01 or 0.05 for long trading positions (*i.e.*, left tail) meaning that the risk manager seeks a high degree of statistical confidence, 99% and 95%, respectively, that the portfolio loss on trading day $t+1$ will not exceed the VaR extracted from information up to day t .

VaR can be estimated using various methods, ranging from non-parametric (simulation), semi-parametric (CAViaR) to fully parametric (location-scale) and optimal combinations thereof; e.g., Kuester et al. (2006) and Fuertes and Olmo (2012). Large banks and financial institutions require multivariate VaR models for capturing appropriately the asset dependence structure in their trading portfolios. We adopt a Monte Carlo copula-based approach via the Cholesky decomposition of the correlation matrix for simulating the portfolio value changes, or P/L distribution, and estimating the VaR at any confidence level; see Appendix 3.G.

Various backtesting methods can be used for assessing the accuracy of VaR forecasts. Let H_{t+1} denote a *hit* or exception, namely, a day when the *ex post* portfolio return falls below the out-of-sample VaR forecast (*i.e.*, larger loss than the maximum

loss anticipated). Formally, the hit sequence (0s or 1s) is given by

$$H_{t+1} = \begin{cases} 1 & \text{if } R_{t+1} < VaR_{t+1}^\alpha \\ 0 & \text{otherwise} \end{cases},$$

Kupiec (1995)'s *unconditional coverage* (UC) test is designed to assess whether the expected *hit* rate is equal to the nominal coverage rate, namely, the hypotheses are $H_0 : \mathbb{E}(H_{t+1}) = \alpha$ versus $H_A : \mathbb{E}(H_{t+1}) \neq \alpha$. Since the random variable H_{t+1} is binomial, the expected probability of observing N exceptions over an evaluation period of T_1 trading days is $(1 - \alpha)^{T_1 - N} \alpha^N$ under H_0 . The corresponding likelihood ratio statistic is

$$LR_{UC} = -2 \ln \left(\frac{(1 - \alpha)^{T_1 - N} \alpha^N}{(1 - \hat{\alpha})^{T_1 - N} \hat{\alpha}^N} \right) \stackrel{asy}{\sim} \chi_1^2 \quad (3.14)$$

where $\hat{\alpha} = \frac{N}{T_1}$ is the observed *hit* rate. A weakness of this test is its unconditional nature, *i.e.* it only “counts” hits but disregards how clustered they are. A well-specified risk management model should efficiently exploit all the available information I_t so that VaR exceptions are unpredictable, *i.e.* $\mathbb{E}(H_{t+1}|I_t) = \mathbb{E}(H_{t+1}) = \alpha$.

Christoffersen (1998)'s *conditional coverage* (CC) test overcomes this drawback. Its aim is to assess whether the correct out-of-sample VaR specification property, $\mathbb{E}(H_{t+1}|I_t) = \alpha$ is met. An implication of this property is that H_{t+1} should be *iid* binomial with mean α . Hence, this is essentially a test of the joint hypothesis of correct unconditional coverage and independence of the hits via the LR statistic

$$LR_{CC} = LR_{UC} + LR_{Ind} = -2 \ln \left(\frac{(1 - \alpha)^{T_1 - N} \alpha^N}{(1 - \hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1 - \hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}}} \right) \stackrel{asy}{\sim} \chi_2^2 \quad (3.15)$$

where n_{10} denotes the number of transitions or instances when an exception occurred on day t and not on day $t - 1$ and $\hat{\pi}_{10} = \frac{n_{10}}{n_{10} + n_{11}}$ is the estimated probability of having an exception on day t conditional on not having an exception on day $t - 1$. Thus the test can detect if the probability of observing an exception, under the assumption of independence, is equal to α which amounts to testing that $\pi_{01} = \pi_{11} = \alpha$.

However, the condition of correct VaR specification $\mathbb{E}(H_{t+1}|I_t) = \alpha$ is stronger

than what Christoffersen (1998)'s CC test can detect. The out-of-sample hits H_{t+1} should be uncorrelated with *any* variable in I_t , meaning that H_{t+1} should be a completely unpredictable process. Christoffersen (1998)'s test can only detect autocorrelation of order one because it is built upon a first-order Markov chain assumption for the hits.

Engle and Manganelli (2004)'s *dynamic quantile* (DQ) test for conditional coverage was developed to address this shortcoming. This is essentially a Wald test for the overall significance of a linear probability model $\mathbf{H} - \alpha \mathbf{1} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where $\mathbf{H} - \alpha \mathbf{1}$ with $\mathbf{H} = (H_{t+1})$ the demeaned hit variable, $\mathbf{1}$ a vector of ones, $\mathbf{X} = (H_t, \dots, H_{t-k}, VaR_{t+1}^\alpha)'$ the regressor vector, and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_{k+2})'$ the corresponding slope coefficients. The null hypothesis is $H_0 : \boldsymbol{\beta} = 0$ and it can be tested using the Wald type test statistic

$$DQ = \frac{\hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{X} \hat{\boldsymbol{\beta}}}{\alpha(1 - \alpha)} \stackrel{asy}{\sim} \chi_{k+2}^2. \quad (3.16)$$

Below we employ $k = 4$ as in Kuester et al. (2006).

One drawback of these common backtesting approaches is that they cannot provide a ranking of VaR models. According to the requirements of the Basel Committee on Banking Supervision, the magnitude as well as the number of exceptions are a matter of regulatory concern. The quadratic loss function suggested by Lopez (1998) takes into account both aspects by adding a penalty based on the size of the exceptions

$$L_{t+1}^Q = \begin{cases} 1 + (R_{t+1} - VaR_{t+1}^\alpha)^2 & \text{if } R_{t+1} < VaR_{t+1}^\alpha \\ 0 & \text{otherwise} \end{cases}, \quad (3.17)$$

and thus, larger tail losses get a disproportionately heavier penalty. However, the above loss function can be subject to the criticism that squared monetary returns lack financial intuition. Blanco and Ihle (1999) suggest focusing on the relative size

of exceptions (percentage) via the loss function

$$L_{t+1}^{\%} = \begin{cases} \frac{R_{t+1} - VaR_{t+1}^{\alpha}}{VaR_{t+1}^{\alpha}} \times 100 & \text{if } R_{t+1} < VaR_{t+1}^{\alpha} \\ 0 & \text{otherwise} \end{cases} . \quad (3.18)$$

The average losses $\bar{L}^Q = \frac{1}{T_1} \sum_{t=1}^{T_1} L_{t+1}^Q$ and $\bar{L}^{\%} = \frac{1}{T_1} \sum_{t=1}^{T_1} L_{t+1}^{\%}$ contain additional information on how good the VaR model is for predicting tail behavior of the portfolio P/L distribution. Therefore, they can rank those VaR models that pass the initial backtesting stage according to their potential cost to the risk manager.

The out-of-sample or holdout period for the VaR forecast evaluation is March 11, 2010 to March 11, 2011 (last sample year, $T_1 = 256$ days) which conforms with the Basel committee recommendation of using an evaluation period of at least 250 business days. The exercise is conducted for the six (CDS-equity) portfolios studied in the chapter at two nominal coverage levels $\alpha = \{0.05, 0.01\}$. The simulation copula-based VaR approach outlined in Appendix 3.G is deployed sequentially over a rolling window of fixed-length ($T_0 = 1,124$ days). Thus, the first estimation window runs from September 10, 2005 to March 10, 2010 and the corresponding one-step-ahead downside risk forecast VaR_{t+1}^{α} is for March 11, 2010. This rolling window approach to generate VaR forecasts offers some "shield" for the simple static copula model against changing market conditions. However, as the results below suggest, the dependence forecasts obtained from models that explicitly capture regime-switching behavior are able to adapt faster and more effectively to changing market conditions. Table 3.5 summarizes the performance of VaR forecasts stemming from various formulations of the Student's t copula function. Since the two dynamic formulations, ARMA and DCC, did not produce markedly different results to save space we only report the results for the former. The hit rate is above the desired nominal coverage α in all four copula formulations but less so with RS-ARMA. For all portfolios, Kupiec's UC test and Christoffersen's CC test are comfortably passed by the four copula-based VaR models considered. In contrast, we observe various rejections of the relatively tough Engle and Manganelli's DQ test, however, none of them is associated with the

RS-ARMA copula formulation. For the Stoxx-CDS portfolio, the null hypothesis of correct VaR model specification is rejected by the DQ test in the static, conventional RS and dynamic ARMA copula formulations.

Taking into account now both the frequency and magnitude of exceptions, both the *quadratic loss* function (3.17) and, more clearly, the *percentage loss* function (3.18) confirm that there is economic value in modeling CDS-equity portfolio risk with regime-switching dynamic copula. For all portfolios, the largest reduction in average out-of-sample losses relative to the static copula is attained by the RS-ARMA copula which also improves upon the conventional RS copula. The economic benefit of flexibly modeling portfolio risk using regime-switching dynamic copula is most clearly seen for the financial CDS-equity portfolios with a *percentage loss* reduction of over 40% and 70% for VaRs at the 0.05 and 0.01 nominal coverages, respectively.

The risk management exercise has thus far relied on the Student's t copula. We now consider the Gumbel copula which also captures tail dependence but in an asymmetric manner. Our subsequent VaR analysis is based on the Gumbel copula that describes dependence on the “adverse” tail; *i.e.*, large CDS returns together with low equity returns or with high equity volatility. In the dynamic ARMA formulation, the average level of tail dependence λ_t inferred from Gumbel copula is about 0.25 for the six CDS-equity pairs and is strongly significant in each case whereas the tail dependence inferred from Student's t copula is very low (order of magnitude 10^{-3}). This notable contrast is likely to have an impact on the VaR forecasts, that is, Gumbel copula forecasts can be expected to yield more conservative VaRs than Student's t copula forecasts. Table 3.6 summarizes the VaR forecasting performance for the Gumbel copula.

Like-for-like comparisons reveal that the Gumbel copula leads to a more reliable risk management model than the Student's t copula. For all portfolios, the DQ test is unable to reject the null hypothesis of correct VaR model specification using the Gumbel copula, irrespective of whether it is formulated in a purely static, dynamic or regime-switching framework. In line with our expectations based on the tail dependence estimates, out-of-sample VaR forecasts from Gumbel copula are more

Table 3.5: Value-at-Risk Portfolios of Stoxx and CDS Indices (Student's t copula)

	Static t copula		RS t copula		ARMA t copula		RS-ARMA t copula	
	5% VaR	1% VaR	5% VaR	1% VaR	5% VaR	1% VaR	5% VaR	1% VaR
Stoxx - CDS Europe								
Total exceptions	6.61%	2.33%	6.61%	2.33%	6.61%	2.33%	6.23%	2.33%
LR_{UC} test	0.257	0.067	0.257	0.067	0.257	0.067	0.384	0.067
LR_{CC} test	0.486	0.158	0.486	0.158	0.146	0.158	0.642	0.158
DQ test	0.000	0.000	0.029	0.021	0.046	0.000	0.556	0.213
Quadratic loss	2.011	0.985	1.942	0.878	1.542	0.356	1.412	0.480
Percentage loss	6.493%	2.527%	5.76%	1.68%	4.531%	1.038%	4.469%	1.038%
Stoxx Auto - CDS Auto								
Total exceptions	7.00%	1.95%	6.61%	1.56%	6.61%	1.95%	6.23%	1.56%
LR_{UC} test	0.163	0.177	0.257	0.407	0.257	0.177	0.384	0.407
LR_{CC} test	0.090	0.358	0.486	0.655	0.355	0.358	0.642	0.655
DQ test	0.090	0.145	0.267	0.892	0.149	0.323	0.836	0.558
Quadratic loss	1.789	0.539	1.860	0.548	1.963	0.576	1.743	0.448
Percentage loss	5.95%	1.68%	4.82%	1.44%	5.94%	1.52%	4.78%	1.40%
Stoxx Fin - CDS SubFin								
Total exceptions	6.23%	1.56%	5.84%	1.95%	5.45%	1.56%	5.06%	1.56%
LR_{UC} test	0.384	0.407	0.548	0.177	0.745	0.407	0.966	0.407
LR_{CC} test	0.407	0.655	0.424	0.358	0.864	0.655	0.473	0.655
DQ test	0.246	0.833	0.248	0.646	0.883	0.800	0.868	0.439
Quadratic loss	1.305	0.386	1.233	0.368	1.204	0.215	1.129	0.126
Percentage loss	4.01%	1.00%	3.68%	0.78%	2.44%	0.38%	2.40%	0.28%
Vstoxx - CDS Europe								
Total exceptions	5.45%	1.95%	5.06%	1.95%	5.06%	1.56%	5.06%	1.17%
LR_{UC} test	0.745	0.177	0.966	0.177	0.966	0.407	0.966	0.793
LR_{CC} test	0.399	0.358	0.473	0.358	0.871	0.655	0.473	0.921
DQ test	0.378	0.282	0.341	0.074	0.832	0.415	0.457	0.721
Quadratic loss	2.376	0.441	2.074	0.462	2.097	0.429	2.137	0.361
Percentage loss	3.53%	0.73%	3.04%	0.73%	3.17%	0.75%	3.03%	0.61%
Vstoxx - CDS Auto								
Total exceptions	5.45%	1.95%	5.06%	1.17%	5.06%	1.17%	5.06%	1.17%
LR_{UC} test	0.745	0.177	0.966	0.793	0.966	0.793	0.966	0.793
LR_{CC} test	0.399	0.358	0.473	0.921	0.473	0.921	0.473	0.921
DQ test	0.783	0.001	0.358	0.934	0.319	0.996	0.779	0.979
Quadratic loss	2.578	0.355	2.544	0.459	3.761	1.352	3.246	0.981
Percentage loss	2.77%	0.70%	2.78%	0.62%	2.64%	0.68%	2.55%	0.60%
Vstoxx - CDS SubFin								
Total exceptions	5.45%	1.17%	5.45%	1.17%	5.45%	1.17%	5.06%	1.17%
LR_{UC} test	0.745	0.793	0.745	0.793	0.745	0.793	0.966	0.793
LR_{CC} test	0.399	0.921	0.864	0.921	0.399	0.921	0.338	0.921
DQ test	0.468	0.997	0.769	0.990	0.417	0.992	0.399	0.977
Quadratic loss	1.822	0.559	1.259	0.237	1.689	0.422	0.823	0.252
Percentage loss	1.77%	0.42%	1.72%	0.36%	1.89%	0.39%	1.55%	0.30%

Notes: This table reports the results of the comparison of out-of-sample VaR forecasts based on three Student's t copula formulations (static, dynamic and regime-switching) for portfolios of Stoxx/Vstoxx and CDS indices. Total exceptions are percentage of days in out-of-sample period when the actual portfolio loss exceeds the VaR forecast. p -values are reported for Kupiec's unconditional coverage (UC), Christoffersen's conditional coverage (CC) and Engle & Manganelli's dynamic quantile (DQ) tests. The quadratic and percentage losses reported are the out-of-sample averages of Eq. (3.17) and Eq. (3.18), respectively.

Table 3.6: Value-at-Risk Portfolios of Stoxx and CDS Indices (Gumbel Copula)

	Static Gumbel		RS Gumbel		ARMA Gumbel		RS-ARMA Gumbel	
	5% VaR	1% VaR	5% VaR	1% VaR	5% VaR	1% VaR	5% VaR	1% VaR
Stoxx - CDS Europe								
Total exceptions	4.28%	1.56%	4.28%	1.56%	4.28%	1.56%	4.28%	1.56%
LR_{UC} test	0.588	0.407	0.588	0.407	0.588	0.407	0.588	0.407
LR_{CC} test	0.504	0.655	0.504	0.655	0.504	0.655	0.504	0.655
DQ test	0.600	0.595	0.608	0.609	0.621	0.620	0.667	0.648
Quadratic loss	1.254	0.452	1.250	0.447	1.245	0.436	1.227	0.432
Percentage loss	3.06%	0.94%	3.05%	0.92%	3.01%	0.90%	2.93%	0.88%
Stoxx Auto - CDS Auto								
Total exceptions	4.67%	1.56%	4.67%	1.56%	4.28%	1.56%	3.89%	1.17%
LR_{UC} test	0.806	0.407	0.806	0.407	0.588	0.407	0.397	0.793
LR_{CC} test	0.512	0.655	0.512	0.655	0.504	0.655	0.447	0.921
DQ test	0.767	0.455	0.770	0.491	0.808	0.479	0.825	0.879
Quadratic loss	1.557	0.538	1.555	0.550	1.564	0.543	1.443	0.540
Percentage loss	3.54%	0.95%	3.49%	0.91%	3.53%	0.90%	3.35%	0.92%
Stoxx Fin - CDS SubFin								
Total exceptions	5.06%	1.17%	4.28%	1.17%	4.67%	1.17%	4.28%	1.17%
LR_{UC} test	0.966	0.793	0.588	0.407	0.806	0.793	0.588	0.793
LR_{CC} test	0.871	0.921	0.644	0.921	0.792	0.921	0.644	0.921
DQ test	0.920	0.995	0.901	0.987	0.861	0.996	0.790	0.997
Quadratic loss	0.789	0.110	0.774	0.109	0.776	0.105	0.803	0.103
Percentage loss	1.95%	0.27%	1.91%	0.27%	1.94%	0.27%	1.89%	0.26%
Vstoxx - CDS Europe								
Total exceptions	5.45%	1.95%	5.45%	1.95%	5.45%	1.95%	5.06%	1.56%
LR_{UC} test	0.745	0.177	0.745	0.177	0.745	0.177	0.966	0.407
LR_{CC} test	0.399	0.358	0.399	0.358	0.399	0.358	0.473	0.655
DQ test	0.414	0.123	0.422	0.130	0.422	0.128	0.821	0.260
Quadratic loss	1.681	0.316	1.609	0.299	1.638	0.300	1.641	0.301
Percentage loss	7.88%	2.94%	7.46%	2.53%	7.84%	2.92%	7.30%	2.38%
Vstoxx - CDS Auto								
Total exceptions	5.45%	1.17%	5.45%	1.17%	4.28%	1.17%	4.28%	0.78%
LR_{UC} test	0.745	0.793	0.745	0.793	0.588	0.793	0.588	0.710
LR_{CC} test	0.399	0.921	0.399	0.921	0.504	0.921	0.504	0.911
DQ test	0.441	0.996	0.445	0.995	0.777	0.994	0.778	0.980
Quadratic loss	2.766	0.360	2.639	0.331	2.661	0.272	2.217	0.211
Percentage loss	2.44%	0.49%	2.39%	0.45%	2.18%	0.33%	1.80%	0.28%
Vstoxx - CDS SubFin								
Total exceptions	5.45%	0.78%	5.06%	0.78%	5.06%	0.78%	5.06%	0.78%
LR_{UC} test	0.745	0.710	0.966	0.710	0.966	0.710	0.966	0.710
LR_{CC} test	0.864	0.911	0.871	0.911	0.871	0.911	0.871	0.911
DQ test	0.660	0.999	0.735	0.998	0.784	0.999	0.779	0.999
Quadratic loss	1.491	0.266	1.476	0.257	1.466	0.263	1.529	0.220
Percentage loss	1.56%	0.21%	1.54%	0.19%	1.51%	0.21%	1.46%	0.18%

Notes: See note to Table 3.5.

conservative (higher) than those from Student's t copula. In fact, while the actual coverage levels of the Student's t VaRs always exceeded the nominal levels (Table 3.5) those for Gumbel-based VaRs tend to be slightly below. Relatively, the average out-of-sample losses in excess of VaR lessen in the Gumbel-based framework. Hence, relaxing the assumption of symmetric tail dependence can improve risk management practice. In this sense, our analysis is consistent with Okimoto (2008) who documents for international (U.S. and U.K.) equity index portfolios that ignoring asymmetric tail-dependence effects during bear market conditions tends to underestimate VaR.

Finally, we can see that when the underlying copula function of choice is Gumbel the RS-ARMA formulation still remains superior to the static, dynamic and conventional RS models, according to the average portfolio losses. Overall, it seems fair to conclude that flexibly modeling the CDS-equity portfolio risk by allowing not only for sudden regime-changes but also mean-reversion in dependence within each regime can entail economic benefits from a risk management perspective.

3.5 Conclusion

Accurately describing the bivariate distribution of CDS and equity instruments is of relevance to risk managers for setting VaR trading limits, to traders for hedging the market risk of their CDS positions, and to regulators and economic policymakers in order to set minimum capital levels. Sudden changes from a low or “normal” dependence regime to a high or “crash” dependence regime can occur as a reflection of important systemic shocks. We propose flexible copula models that explicitly capture regime-switching behavior and allow for mean-reversion in dependence within each regime. By means of the proposed Markov-switching dynamic copulas and simpler (nested) versions, we provide a comprehensive study of the dependence structure in CDS-equity markets. The evaluation and comparison of copulas is conducted both in-sample using common goodness-of-fit measures and out-of-sample using Value at Risk (VaR) forecast accuracy measures.

The proposed models confirm the presence of significant negative comovement

between CDS returns and stock returns, and significantly positive comovement between CDS returns and stock return volatility over the period from September 2005 to March 2011. They also indicate that asset dependence is time-varying and nonlinear. Significant regime-switching dependence is revealed not only in the central part of the bivariate distributions but also in the tails; namely, *low* and *high* dependence periods alternate over time. The latter broadly coincide with the automotive crisis, the subprime mortgage crisis and the Greek and European sovereign debt crises. The findings suggest that during periods of stress systematic factors play a stronger role as drivers of default and volatility shocks have longer lasting effects than return shocks. Inadequately modeling the bivariate distribution by ignoring the time-variation in dependence within each regime or altogether neglecting regime-switching effects implies too “smooth” rank correlation and tail dependence measures that under(over)-estimate the comovement of CDS and equity markets in crisis (normal) periods.

Our assessment of the competing copula models in a risk-management exercise leads to two main conclusions. First, neglecting regime-switching effects in dependence can be costly because it tends to underestimate the maximum potential losses of CDS-equity portfolios. Second, relaxing the assumption that the dependence structure remains constant within each regime can be beneficial for improving the accuracy of out-of-sample VaR forecasts and producing smaller average regulatory losses. Lastly, the study provides yet another example of a disconnect between in-sample fit and out-of-sample predictability; namely, the Student’s t copula function is strongly supported by common in-sample statistical criteria but the Gumbel copula which focuses on the adverse tail leads to more conservative 1-day-ahead VaR trading limits.

These findings are relevant to institutional investors using CDS contracts to hedge their equity holdings and for capital structure arbitrageurs, particularly, those predominantly exposed to financial and auto sectors. The flexible copula models proposed could prove useful for banks in order to address recommendations from Basel III to carry out more rigorous stress testing of the trading book. Our study offers important insights into the regime-switching dependence dynamics of CDS spreads and tradeable systematic risk factors. In this context, extending the flexible regime-

switching copula models here proposed to ascribe a role to exogenous structural variables as drivers of the regime-transition may be an interesting avenue of further research.

Appendix

3.A Merton's Structural Model and Extensions

Pioneered by Merton (1974) and Black and Scholes (1973), structural (or asset value) model is one of the two primary classes of credit risk modeling approaches.²³ It assumes that at time t a firm with risky assets A_t is financed by equity E_t and zero-coupon debt D_t of face value K maturing at time $T > t$: $A_t = E_t + D_t$.

When the firm's asset is valued more than its debt $A_T \geq K$ at time T , the debt holders will be paid the full amount K and the shareholders' equity will be $(A_T - K)$. On the other hand, when the firm fails to repay (therefore defaults on) the debt at T , the debt holders can only recover $A_T < K$ and the shareholders will get nothing. The equity value at time T can be represented as an European call option on asset A_t with strike price K maturing at T , $E_T = \max(A_T - K, 0)$. The asset value is assumed to follow a geometric Brownian motion process, with risk-neutral dynamics given

$$dA_t = rA_t dt + \sigma_A A_t dW_t \quad (3.19)$$

where r denotes the risk-free interest rate, σ_A is the volatility of asset's returns, and W_t is a Brownian motion under the risk-neutral measure. Applying Black-Scholes formula would give

$$E_t = A_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2)$$

where $d_1 = \frac{1}{\sigma_A \sqrt{T-t}} \left[\ln\left(\frac{A_t}{K}\right) + \left(r + \frac{\sigma_A^2}{2}\right)(T-t) \right]$, $d_2 = d_1 - \sigma_A \sqrt{T-t}$, and $\Phi(\cdot)$ denotes the standard normal *cdf*. The probability of default at time T is given by $P(A_T < K) = \Phi(-d_2)$.

A typical strategy of debt holders to protect themselves from the credit risk is to long a put option P_t on A_t with strike K maturing at T . The put option will be valued at $(K - A_T)$ if $A_T < K$, and worth nothing if $A_T > K$. Purchasing the put option guarantees that the credit risk of the loan is hedged completely as the

²³The other one is the reduced form model.

debt holder's payoff equals K at maturity no matter if the obligor defaults or not. It therefore forms a risk-free position

$$D_t + P_t = Ke^{-r(T-t)}. \quad (3.20)$$

The price of put option P_t is determined by applying Black-Scholes formula as

$$P_t = Ke^{-r(T-t)}\Phi(-d_2) - A_t\Phi(-d_1). \quad (3.21)$$

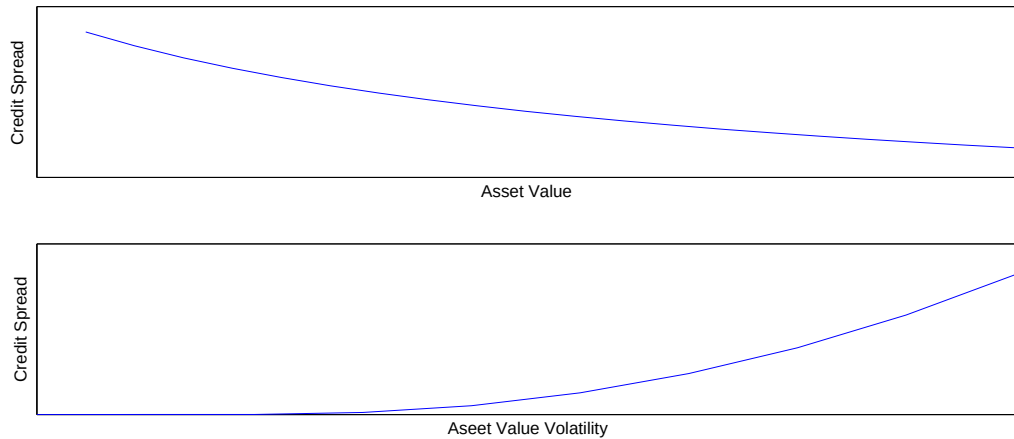
Taking account the credit risk spread (risk premium) s , the value of the risky bond is

$$D_t = Ke^{-(r+s)(T-t)}. \quad (3.22)$$

Combining Eq.(3.20) – (3.22) gives a closed-form formula for the credit spread

$$s = -\frac{1}{T-t} \ln \left[\Phi(d_2) - \frac{A_t}{K} e^{r(T-t)} \Phi(-d_1) \right]$$

where $\frac{A_t}{K}$ represents the firm's leverage. Note that s depends only on A_t and σ_A which is in line with the economic intuition. Their nonlinear relationship can be observed from the below figures.



Many approaches have been proposed to improve the classical Merton's model. The first passage model introduced by Black and Cox (1976) allows the firm may

default at any time before the debt maturity. Jones et al. (1984) suggest to introduce stochastic interest rates to improve the model's performance. Longstaff and Schwartz (1995) employ a Vasicek process for the interest rate, $dr_t = (a - br_t)dt + \sigma_t dW_t^{(r)}$, while Kim et al. (1993) consider a CIR process, $dr_t = (a - br_t)dt + \sigma_t \sqrt{r_t} dW_t^{(r)}$, and Briys and De Varenne (1997) treat the interest rate following a generalized Vasicek process, $dr_t = (a(t) - b(t)r_t)dt + \sigma_t(t)dW_t^{(r)}$. By comparing the Merton's model and its four extensions Eom et al. (2004) find substantial spread predication errors that four models underestimate the spread observed from the market while the other one overestimate it.

3.B Dependence Measures

Pearson linear correlation Let X_1 and X_2 denote two continuous random variables representing the returns of asset 1 and asset 2, respectively. Pearson correlation $\rho \in [-1, 1]$ is defined by

$$\rho = \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)}\sqrt{Var(X_2)}}$$

where $Cov(X_1, X_2) = \mathbb{E}(X_1, X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2)$ is the covariance, and $Var(X_1) = Cov(X_1, X_1)$ denotes the variance. ρ is an appropriate measure of dependence only if the true joint distribution of the asset returns is *elliptical*, and it is scale invariant only under strictly increasing linear transformations.

Kendall's τ rank correlation Kendall's $\tau \in [-1, 1]$ is a rank correlation measure based on the *concordance* notion: two observation pairs (x_{1t}, x_{2t}) and (x_{1s}, x_{2s}) are concordant if $(x_{1t} - x_{1s})(x_{2t} - x_{2s}) > 0$ and discordant if $(x_{1t} - x_{1s})(x_{2t} - x_{2s}) < 0$, for $t, s = 1, \dots, T$. The rank correlation of (X_1, X_2) is a measure of the probability of concordance minus the probability of discordance as given by

$$\tau = \frac{\# \text{concordant pairs} - \# \text{discordant pairs}}{\frac{1}{2}T(T-1)}$$

where T is the number of observations for x_1 and x_2 , and the denominator indicates the total number of pairs $(x_{1t}, x_{2t}), (x_{1s}, x_{2s})$. Kendall's τ is scale invariant under strictly increasing (non)linear transformations.

Tail dependence Tail dependence $\lambda \in [0, 1]$ measures the concordance in the tails, or extreme values of two random variables, X_1 and X_2 , and is defined in terms of limiting conditional probabilities of α -quantile exceedances as $\alpha \rightarrow 1$. Formally, for two positively correlated variables, *upper* tail dependence denoted λ^U is the probability that X_2 is above its α -quantile ($X_{2,\alpha}$) given that X_1 is above its α -quantile ($X_{1,\alpha}$), *i.e.* $\lambda^U \equiv P(X_2 > X_{2,\alpha} | X_1 > X_{1,\alpha})$, while *lower* tail dependence denoted λ^L is the probability that X_2 is smaller than its $(1 - \alpha)$ -quantile given that X_1 is smaller than its $(1 - \alpha)$ -quantile, *i.e.* $\lambda^L \equiv P(X_2 \leq X_{2,(1-\alpha)} | X_1 \leq X_{1,(1-\alpha)})$. For two *negatively correlated* variables, *upper tail dependence* is the probability that X_2 is above its α -quantile given that X_1 is below its $(1 - \alpha)$ -quantile, *i.e.* $\lambda^U \equiv P(X_2 > X_{2,\alpha} | X_1 \leq X_{1,(1-\alpha)})$, while *lower tail dependence* is the probability that X_2 is below its $(1 - \alpha)$ -quantile given that X_1 is above its α -quantile, *i.e.* $\lambda^L \equiv P(X_2 \leq X_{2,(1-\alpha)} | X_1 > X_{1,\alpha})$. Like the rank correlation, tail dependence depends only on the copula of X_1 and X_2 thus the roles of X_1 and X_2 are interchangeable.²⁴

3.C Static Copulas

Gaussian copula (elliptical) The bivariate Gaussian copula *pdf* can be written as follows

$$c(u_1, u_2; \boldsymbol{\theta} = \rho) = \frac{1}{\sqrt{1 - \rho^2}} \exp \left(-\frac{1}{2} \Psi' (R^{-1} - I) \Psi \right)$$

where R is the correlation matrix with off-diagonal element the Pearson correlation ρ , and $\Psi = (\Phi^{-1}(u_1), \Phi^{-1}(u_2))'$ with Φ^{-1} the inverse of the univariate standardized

²⁴For further discussion on dependence measures see Cherubini et al. (2004), Ch.3.1 or McNeil et al. (2005) Ch. 5.2.

Gaussian *cdf*. Tail dependence is not captured.

Student's t copula (elliptical) The bivariate Student's t copula *pdf* can be written as follows

$$c(u_1, u_2; \boldsymbol{\theta} = (\nu, \rho)) = \frac{1}{\sqrt{1-\rho^2}} \frac{\Gamma\left(\frac{\nu+2}{2}\right) \Gamma\left(\frac{\nu}{2}\right) \prod_{i=1}^2 \left(1 + \frac{1}{\nu} (t_{\nu_i}^{-1}(u_i))^2\right)^{\frac{\nu+1}{2}}}{\left(\Gamma\left(\frac{\nu+1}{2}\right)\right)^2 \left(1 + \frac{1}{\nu} \Psi' R^{-1} \Psi\right)^{\frac{\nu+2}{2}}}$$

where R is the correlation matrix with ρ as off-diagonal element, $\Psi = (t_{\nu_1}^{-1}(u_1), t_{\nu_2}^{-1}(u_2))'$ and $t_{\nu}^{-1}(u)$ is the inverse *cdf* of the Student's t with $\nu > 2$ degrees of freedom. Student's t copula captures tail dependence but imposes the restriction of upper and lower tail symmetry. It converges to Gaussian copula as $\nu \rightarrow \infty$.

Gumbel copula (Archimedean) The bivariate Gumbel copula *pdf* can be written as follows:

$$c(u_1, u_2; \boldsymbol{\theta} = \eta) = \frac{(\log u_1 \log u_2)^{e^\eta}}{u_1 u_2} \exp \left\{ - \left[\sum_{n=1}^2 (-\log u_n)^{e^\eta+1} \right]^{\frac{1}{e^\eta+1}} \right\} \\ \left[e^\eta + \left[\sum_{n=1}^2 (-\log u_n)^{e^\eta+1} \right]^{\frac{1}{e^\eta+1}} \left[\sum_{n=1}^2 (-\log u_n)^{e^\eta+1} \right]^{-2+\frac{1}{e^\eta+1}} \right]$$

where $\eta \in [1, +\infty)$. For positively correlated variables (*e.g.*, CDS and VStoxx), the Gumbel copula can capture upper tail dependence but 180°-anticlockwise-rotated (or survival) Gumbel can capture lower tail dependence. For negatively correlated variables (*e.g.*, CDS and Stoxx), the 90°(270°)-anticlockwise-rotated Gumbel copula captures upper (lower) tail dependence.

SJC copula (Archimedean) The bivariate SJC copula *pdf* can be written as follows:

$$c(u_1, u_2; \boldsymbol{\theta} = (\lambda^U, \lambda^L)) = \frac{1}{2} \left[\frac{\partial^2}{\partial u_1 \partial u_2} C_{JC}(u_1, u_2 | \lambda^U, \lambda^L) + \frac{\partial^2}{\partial u_1 \partial u_2} C_{JC}(1 - u_1, 1 - u_2 | \lambda^U, \lambda^L) \right] \quad (3.23)$$

where

$$\frac{\partial^2}{\partial u_1 \partial u_2} C_{JC}(u_1, u_2; \boldsymbol{\theta} = (\lambda^U, \lambda^L)) = A - B; \quad (3.24)$$

$$\begin{aligned} A &= \frac{\gamma \kappa \left(1 - \frac{1}{Z^{1/\gamma}}\right)^{-2+\frac{1}{\gamma}} \cdot \left(\frac{1}{\gamma} - 1\right) \cdot (1 - u_1)^{\kappa-1} \cdot (1 - u_2)^{\kappa-1}}{(1 - (1 - u_1)^\kappa)^{1+\gamma} \cdot (1 - (1 - u_2)^\kappa)^{1+\gamma} \cdot Z^{2+\frac{2}{\gamma}}}; \\ B &= \frac{\kappa \left(1 - \frac{1}{Z^{1/\gamma}}\right)^{-2+\frac{1}{\gamma}} \cdot \left(\frac{1}{\kappa} - 1\right) \cdot (1 - u_1)^{\kappa-1} \cdot (1 - u_2)^{\kappa-1}}{(1 - (1 - u_1)^\kappa)^{1+\gamma} \cdot (1 - (1 - u_2)^\kappa)^{1+\gamma} \cdot Z^{2+\frac{2}{\gamma}}}; \\ Z &= \frac{1}{(1 - (1 - u_1)^\kappa)^\gamma} + \frac{1}{(1 - (1 - u_2)^\kappa)^\gamma} - 1, \end{aligned}$$

and λ^U (λ^L) is the upper (lower) tail dependence. C_{JC} is the *cdf* of the Joe-Clayton copula as

$$C_{JC}(u_1, u_2; \boldsymbol{\theta} = (\lambda^U, \lambda^L)) = 1 - \left\{ 1 - \frac{1}{\left[\frac{1}{(1 - (1 - u_1)^\kappa)^\gamma} + \frac{1}{(1 - (1 - u_2)^\kappa)^\gamma} - 1 \right]^{\frac{1}{\gamma}}} \right\}^{\frac{1}{\kappa}}$$

with $\kappa = \frac{1}{\log_2(2 - \lambda^U)}$, $\gamma = \frac{1}{\log_2 \lambda^L}$.

The following table summarizes how the rank correlation and tail dependence are obtained:

Copula	Kendall's τ	Lower tail dependence (λ^L)	Upper tail dependence (λ^U)
Gaussian	$\frac{2}{\pi} \arcsin \rho$	0	0
Student's t	$\frac{2}{\pi} \arcsin \rho$	$2t_{\nu+1} \left[-\sqrt{\nu+1} \frac{\sqrt{1-\rho}}{\sqrt{1+\rho}} \right]$	$2t_{\nu+1} \left[-\sqrt{\nu+1} \frac{\sqrt{1-\rho}}{\sqrt{1+\rho}} \right]$
Gumbel	$1 - \frac{1}{\eta}$	-	$2 - 2^{\frac{1}{\eta}}$
SJC	-	λ^L	λ^U

Elliptical copulas permit both positive dependence (*i.e.*, $\tau > 0$) and negative dependence (*i.e.*, $\tau < 0$). Rotating the Archimedean copulas by counterclockwise 90 and 270 degrees accommodates negative dependence, which is not possible with the standard non-rotated versions. In particular, the distribution functions C_{90° and C_{270° of a copula C rotated by counterclockwise 90 and 270 degrees, respectively, are given as $C_{90^\circ}(u_1, u_2) = u_2 - C(1 - u_1, u_2)$ and $C_{270^\circ}(u_1, u_2) = u_1 - C(u_1, 1 - u_2)$.

3.D Log-likelihood Functions of Static Copulas

Copula parameters are estimated via maximum likelihood approaches. We list below the log-likelihood functions of the copula functions used in this chapter.

Gaussian Copula (elliptical) The log-likelihood function is given by

$$\mathcal{L}_{\text{Gaussian}}(R; u_1, u_2) = -\frac{1}{2} \sum_{t=1}^T (\log(1 - \rho^2) + \Psi'_t(R^{-1} - I) \Psi_t). \quad (3.25)$$

where R is the correlation matrix with ρ as off-diagonal element, $\Psi_t = (\Phi^{-1}(u_{1t}), \Phi^{-1}(u_{2t}))$ and $\Phi^{-1}(u)$ is the inverse *cdf* of the Gaussian.

Student t Copula (elliptical) The log-likelihood function of the Student- t copula is

$$\begin{aligned}
\mathcal{L}_{Student-t}(\rho, \nu; u_1, u_2) &= \sum_{t=1}^T \log c(u_{1t}, u_{2t}; \nu, \rho) \\
&= \sum_{t=1}^T \left\{ -\frac{1}{2} \log(1 - \rho^2) + \log \Gamma\left(\frac{\nu+2}{2}\right) + \log \Gamma\left(\frac{\nu}{2}\right) \right. \\
&\quad \left. - 2 \log \Gamma\left(\frac{\nu+1}{2}\right) + \frac{\nu+1}{2} \sum_{i=1}^2 \log\left(1 + \frac{\Psi_i^2}{\nu}\right) \right. \\
&\quad \left. - \frac{\nu+2}{2} \log\left(1 + \frac{\Psi' R^{-1} \Psi}{\nu}\right) \right\} \\
&= -\frac{1}{2} \sum_{t=1}^T \log(1 - \rho^2) + T \log \Gamma\left(\frac{\nu+2}{2}\right) + T \log \Gamma\left(\frac{\nu}{2}\right) \\
&\quad - 2T \log \Gamma\left(\frac{\nu+1}{2}\right) + \frac{\nu+1}{2} \sum_{t=1}^T \sum_{i=1}^2 \log\left(1 + \frac{\Psi_{it}^2}{\nu}\right) \\
&\quad - \frac{\nu+2}{2} \sum_{t=1}^T \log\left(1 + \frac{\Psi'_t R^{-1} \Psi_t}{\nu}\right) \tag{3.26}
\end{aligned}$$

where R is the correlation matrix with ρ as off-diagonal element, $\Psi = (t_{\nu_1}^{-1}(u_1), t_{\nu_2}^{-1}(u_2))'$ and $t_{\nu}^{-1}(u)$ is the inverse *cdf* of the Student's t with $\nu > 2$ degrees of freedom.

Gumbel Copula (Archimedean) The log-likelihood function of the Gumbel copula is

$$\mathcal{L}_{Gumbel}(\eta; u_1, u_2) = \sum_{t=1}^T (\log(f_1) + \log(f_2) + \log(f_3) + \log(f_4) + \log(f_5)) \tag{3.27}$$

where $f_1 = \eta^{-1}$, $f_2 = e^{-\frac{x}{\eta}}$, $f_3 = x^{-2+\frac{1}{\eta}}$, $f_4 = \frac{1}{\eta}(x-1) + 2$, $f_5 = \left(\frac{\eta}{u_1}(-\log u_{1t})^{\eta-1}\right) \left(\frac{\eta}{u_2}(-\log u_{2t})^{\eta-1}\right)$ and $x = (-\log u_{1t})^{\eta} + (-\log u_{2t})^{\eta}$.

SJC Copula (Archimedean) The log-likelihood function for estimating the parameters of the SJC copula is

$$\begin{aligned}
\mathcal{L}_{SJC}(\lambda^U, \lambda^L; u_1, u_2) &= \sum_{t=1}^T \log \left\{ \frac{\partial^2}{\partial u_1 \partial u_2} \frac{1}{2} [C_{JC}(u_{1t}, u_{2t} | \lambda^U, \lambda^L) \right. \\
&\quad \left. + C_{JC}(1 - u_{1t}, 1 - u_{2t} | \lambda^U, \lambda^L) + u_{1t} + u_{2t} - 1] \right\} \\
&= \sum_{t=1}^T \log \left\{ \frac{1}{2} \left[\frac{\partial^2}{\partial u_1 \partial u_2} C_{JC}(u_{1t}, u_{2t} | \lambda^U, \lambda^L) \right. \right. \\
&\quad \left. \left. + \frac{\partial^2}{\partial u_1 \partial u_2} C_{JC}(1 - u_{1t}, 1 - u_{2t} | \lambda^U, \lambda^L) \right] \right\}. \quad (3.28)
\end{aligned}$$

where the second derivatives are derived from Eq.(3.24).

3.E Estimation of Regime-Switching Copula

First, we follow the Hamilton (1989)'s filtering algorithm to obtain $P[S_t = s | I_t]$, the filtered probability of unobservable regime S_t given the available information set, via a two-step process iterated from $t = 1, \dots, T$ starting from an initial value at $t = 0$. The parameters are estimated by ML. Next, we adopt Kim (1994)'s algorithm to obtain the *smoothed probabilities*, $P[S_t = s | I_T]$, or probabilities of each regime given the full sample information set I_T , starting from the last filtered probability at $t = T$ as initial value and iterating backwards from $t = T - 1$ to $t = 1$.

The *filtering* algorithm involves the following two sequential steps:

1. Inference about the current state given the past values of the observed variable

$$P[S_t = s | I_{t-1}] = \sum_{k=0}^1 P[S_t = s | S_{t-1} = k, I_{t-1}] P[S_{t-1} = k | I_{t-1}]$$

where $s = \{0, 1\}$ denotes regimes $\{H, L\}$, respectively; $I_t = [u_{1t}, u_{2t}, I_{t-1}]$ is the time t information set; and the migration probabilities $P[S_t = s | S_{t-1} = k, I_{t-1}]$ are the entries of matrix π in Eq. (3.10).

2. Filtering of S_t in order to generate future forecasts on the prevailing state

$$P[S_t = s | I_t] = \frac{c_t^{S_t}(u_{1t}, u_{2t} | S_t = s, I_{t-1}) P[S_t = s | I_{t-1}]}{\sum_{k=0}^1 c_t(u_{1t}, u_{2t} | S_t = k, I_{t-1}) P[S_t = k | I_{t-1}]}$$

where $c_t^{S_t}(\cdot)$ is the *pdf* of RS-ARMA or RS-DCC copula models.

The log-likelihood function of the RS-copula is

$$\begin{aligned} \mathcal{L}(u_1, u_2 | \boldsymbol{\theta}) &= \mathcal{L}_C(\boldsymbol{\theta}, \hat{\boldsymbol{\phi}}) + \sum_{i=1}^2 L_{X_i}(\phi_i) \\ &= \sum_{t=1}^T \log \left(\sum_{S_t=1}^2 c_t^{S_t}(u_{1t}, u_{2t} | S_t, I_{t-1}) \Pr[S_t | I_{t-1}] \right) \\ &\quad + \sum_{t=1}^T \sum_{i=1}^2 \log(f_{X_i}(x_{i,t}; \phi_{i,t})). \end{aligned}$$

The parameters can be estimated either by the IFM or CML methods.

Once the parameters are estimated and all filtered probabilities $P[S_t = s | I_t]$ for $s \in \{0, 1\}$ and $t = 1, 2, \dots, T$, are obtained, we follow Kim's algorithm to obtain the *smoothed probabilities*

$$P[S_t = s | I_T] = \sum_{k=0}^1 \frac{P[S_{t+1} = s | S_t = k, I_t] P[S_t = s | I_t] P[S_{t+1} = k | I_T]}{\sum_{j=0}^1 P[S_{t+1} = s | S_t = k, I_t] P[S_t = k | I_t]}, t = T-1, \dots, 1$$

starting from $P[S_T = s | I_T]$ and iterating backwards for $t = T-1, T-2, \dots, 1$.

3.F Timeline of late 2000s Crises

Credit Crunch

- **July 10, 2007:** S&P announces it may cut ratings on \$12bn of subprime debt.
- **August 9, 2007:** ECB pumps 95bn euros into the banking system to improve liquidity.

- **October 1, 2007:** UBS announces \$3.4bn losses from sub-prime related investments.
- **October 30, 2007:** Merrill Lynch unveils \$7.9bn exposure to bad debt.
- **January 19, 2008:** World Bank predicts slowdown of global economic growth in 2008.
- **January 21, 2008:** Global stock markets suffer their largest fall since September 2001.
- **February 17, 2008:** UK government nationalizes Northern Rock.
- **March 17, 2008:** Wall Street's 5th largest bank, Bear Stearns, is acquired by JP Morgan Chase.
- **April 8, 2008:** IMF warns that potential losses from the credit crunch could reach \$1tn.
- **September 7, 2008:** Large US mortgage lenders Fannie Mae and Freddie Mac are nationalized.
- **September 15, 2008:** Lehman Brothers files for Chapter 11 bankruptcy protection.
- **September 16, 2008:** US Fed announces \$85bn rescue package for AIG.
- **September 17, 2008:** Lloyds TSB announces takeover of largest British mortgage lender HBOS.
- **October 13, 2008:** UK government announces £37bn injection to RBS, Lloyds TSB and HBOS.
- **November 6, 2008:** Bank of England cuts base interest rate to lowest level since 1955.

Energy Crisis

- **March 5, 2005:** Crude oil prices rose to new highs above \$50 per barrel (bbl).
- **September 2005:** US hurricane Katrina pushes gasoline prices to a record high.
- **August 11, 2005:** Crude oil prices broke the psychological barrier of \$60 bbl.
- **July 13, 2006:** Israeli attacks on Lebanon pushed crude oil prices to historical highs above \$78.40 bbl.
- **October 19, 2007:** US light crude rose to \$90.02 bbl.
- **March 5, 2008:** OPEC accused the US of economic "mismanagement" responsible for oil prices.
- **March 12, 2008:** Oil prices surged above \$110 bbl.

Automotive Industry Crisis

- **May 5, 2005:** S&P cut the debt ratings of GM and Ford to “junk” status.
- **February 12, 2008:** GM announced its operating loss was \$2bn.
- **October 7, 2008:** SEAT cut production at its Martorell plant by 5%.
- **November 20, 2008:** PSA Peugeot Citroen predicts sales volumes would fall by at least 10% in 2009, following a 17% drop in the current quarter.
- **November 23, 2008:** Jaguar Land Rover was seeking a \$1.5bn loan from the government.
- **December 11, 2008:** The Swedish government injected \$3.5bn to rescue its troubled auto markers, Volvo and Saab.
- **December 19, 2008:** US government said it would use up to \$17.4bn to help the big three US carmakers, General Motors, Ford and Chrysler.

- **December 20, 2008:** GM and Chrysler receive CA\$4bn government loans from Canada and the province of Ontario.
- **January 8, 2009:** Nissan UK announced it was to shed 1200 jobs from its factories in North East England.
- **January 22, 2009:** Fiat announces a 19% drop in revenues for 2008 Q3.
- **February 11, 2009:** PSA Peugeot Citroen announced it would cut 11,000 jobs world wide.
- **February 12, 2009:** Renault announces a 78% drop in profits for 2008.
- **April 22, 2009:** GM admits it will default on a \$1bn bond debt payment due in June.
- **April 30, 2009:** Chrysler files for Chapter 11 bankruptcy protection.
- **June 1, 2009:** GM files for Chapter 11 bankruptcy protection.

European sovereign debt crisis

- **October 10, 2008:** Fitch downgrades Iceland Sovereign debt from A+ to BBB-.
- **December 8, 2009:** Fitch ratings agency downgraded Greece's credit rating from A- to BBB+.
- **April 23, 2010:** Greek PM calls for eurozone-IMF rescue package. FTSE falls more than 600p.
- **May 18, 2010:** Greece gets first bailout of \$18bn from EFSF, IMF and bilateral loans
- **November 29, 2010:** Ireland receives \$113bn bailout from EU, IMF and EFSF
- **January 5, 2010:** S&P downgrades Iceland's rating to junk grade.

Sources: news.bbc.co.uk; www.reuters.com; www.bloomberg.com.

3.G Copula VaR Simulation

The portfolio profit and loss (P/L) distribution is simulated using copula as follows:

1. Obtain the 2×2 rank correlation matrix forecast Σ using data up to day t . In static copula, the off-diagonal entry is $\hat{\tau}_{t+1} = \hat{\tau}_t = \hat{\tau}$. In dynamic copula, $\hat{\tau}_{t+1}$ is the 1-day-ahead projection of the ARMA Eq. (3.8) or DCC Eq. (3.9). In regime-switching copulas, the forecast hinges on the filtered probabilities and estimated migration matrix π in Eq.(3.10) as $\hat{\tau}_{t+1} = (\text{P}[S_t = H | I_t], \text{P}[S_t = L | I_t]) \pi (\hat{\tau}_{t+1}^H, \hat{\tau}_{t+1}^L)'$.
2. Simulate two independent standard normal random variates $\mathbf{z} = (z_1, z_2)'$.
3. Simulate a random variate s from a $\chi_{\hat{\nu}}^2$ distribution, independent of $\mathbf{z}$, where $\hat{\nu}$ is the degree-of-freedom estimated using data up to day t .
4. Form the vectors $\mathbf{b} = \mathbf{A}\mathbf{z}$ and $\mathbf{c} = \frac{\sqrt{\hat{\nu}}}{\sqrt{s}}\mathbf{b}$ where $\mathbf{c} = (c_1, c_2)'$ and $\mathbf{A}$ is the Cholesky decomposition of Σ .
5. Determine the components $(u_1, u_2) = (t_{\hat{\nu}}(c_1), t_{\hat{\nu}}(c_2))$ of copula where $t_{\hat{\nu}}$ is the *cdf* of Student's t distribution with degrees-of-freedom parameter $\hat{\nu}$.
6. Obtain the standardized asset log-returns: $\mathbf{Q} = (q_1, q_2) = (\hat{F}_1^{-1}(u_1), \hat{F}_2^{-1}(u_2))$, where $\hat{F}_n^{-1}$ is the inverse empirical *cdf* of standardized residuals, x_n , $n = 1, 2$, of the in-sample data.
7. Relocate and rescale the returns as $(r_{1,t+1}, r_{2,t+1}) = (\hat{\mu}_{1,t+1} + q_1 \sqrt{\hat{\sigma}_{1,t+1}}, \hat{\mu}_{2,t+1} + q_2 \sqrt{\hat{\sigma}_{2,t+1}})$ with $\hat{\mu}_{t+1}$ and $\hat{\sigma}_{t+1}$ denoting ARMA-GARCH-skT forecasts of conditional mean and variance made at t .
8. Obtain the 1-day-ahead P/L forecast of the equally-weighted portfolio as $r_{t+1} = 0.5r_{1,t+1} + 0.5r_{2,t+1}$.

Repeat $J = 100,000$ times the above steps to obtain the empirical or simulated 1-day-ahead P/L distribution $\{r_{t+1,j}\}_{j=1}^J$ from which any α -quantiles (VaR) can be measured.

An Extreme Risk Management Framework for Portfolio optimization

4.1 Introduction

In Modern Portfolio Theory, pioneered by Markowitz (1952), investment decisions are a tradeoff between expected return and variance. The resulting mean-variance (MV) portfolio optimization framework provides the investor with a set of portfolios along the *efficient frontier* that represent the best possible returns for a target variance or the lowest variance for a given target return. By relying on this framework, investors focus solely on the first two moments of the return distribution. Risk is solely defined in terms of *variance* or how much the returns deviate from the expected or *mean* return. But it is well known that variance is not a good measure of risk. Exclusive reliance on the variance of the return distribution implies assuming Gaussianity. Thus the MV strategy implicitly assumes symmetry, namely, it rewards attractive abnormal profits as much as it punishes undesirable acute losses. The common wisdom is instead that investors are particularly concerned about downside risk. It is well known also that financial returns are fat-tailed, leptokurtic and asymmetric. Investors who infer that their portfolios are well diversified under the Gaussian tenet can suffer

catastrophic losses during financial market downturns.

As an effort to include investor preferences going beyond mean and variance in portfolio selections, Baumol (1963) introduced Value-at-Risk (VaR) as alternative risk measure. VaR gives the worst loss over a target horizon at a given confidence level. Since JP Morgan published its RiskMetrics Technical Document in 1994, VaR has established itself as the most common risk management tool in industry but it came in for severe criticism with the onset of the record financial losses of 2008 and 2009. Artzner et al. (1997, 1999) formalized the main properties that a coherent risk measure must satisfy, and showed that VaR is not coherent because subadditivity is not fulfilled for non-normal distributions. This means that the sum of the VaRs of individual assets might be less than the portfolio VaR.¹

An coherent risk measure known as Conditional VaR (CVaR) or expected shortfall was popularized by Artzner et al. (1999) and is gaining prominence among banks and regulators because it is more sensitive to the shape of the tail distribution in a bell curve of potential losses.² CVaR gives the expected total loss exceeding VaR. Rockafellar and Uryasev (2000, 2002) and Andersson et al. (2001) introduced portfolio optimization based on CVaR as relevant risk measure, namely, the mean-CVaR (MCVaR) framework.³ (C)VaR of a portfolio can be estimated nonparametrically by historical simulation or parametrically by a covariance approach. Historical simulation (C)VaR simple requires estimating empirical quantiles based on available past data but, unfortunately, it may be noisy due to sparsity of observations in the tails. In the parametric framework, (C)VaR is directly affected by the assumptions made with regard to the shape of the multivariate distribution that generates the portfolio returns. Extreme value theory (EVT) can be useful tools in this context because it focus only on modeling the tails of the portfolio return distribution without requiring any assumptions on the high-density (central) part. EVT provides the distribution

¹Subadditivity is a property of a mathematical function by which the value of the function at point $A + B$ is less or equal than the sum of the function's values at points A and B , that is, $q(A + B) \leq q(A) + q(B)$. See Appendix 4.A

²It is also called Average Value at Risk (AVaR) or Expected Tail Loss (ETL).

³The Basel committee on Banking Supervision is recently encouraging banks to improve their downside risk modeling, for instance, using stressed VaR or expected shortfall instead of VaR; see Basel Committee on Banking Supervision (2012).

of a variable conditionally upon its values exceeding a certain threshold and, not surprisingly, it has excelled in out-of-sample VaR forecasting competitions; *e.g.*, Gencay and Selcuk (2004), Chan and Gray (2006) and Kuester et al. (2006). EVT has been employed to calculate VaR since Embrechts et al. (1999), but it has been barely utilized for portfolio optimization. To the best of our knowledge, the only exception is Haque et al. (2007) where EVT is applied to the problem of optimizing bivariate portfolios of U.S. and Mexican equity during the 1994 peso crisis.⁴

Another limitation of the MV framework is that by assuming that individual asset returns are Gaussian, it immediately follows that portfolio returns are assumed multivariate Gaussian. Thus the return dependence structure is represented by the Pearson (linear) correlation coefficient which neglects important nonlinear features such as tail dependence. Relying on the multivariate Gaussian distribution implies neglecting *tail dependence* defined as the probability of extreme losses (*e.g.*, large negative returns) in one asset given that extreme losses had occurred in another asset. To cope with this issue and give more realistic description of the fat tails, some exponential- and polynomial-tail distributions have been suggested by researchers for portfolios with different asset classes, for example, Yu et al. (2009) introduce multivariate variance gamma process; Hu (2009) propose a multivariate generalized hyperbolic distribution (GHD) framework; Xiong and Idzorek (2011) optimize portfolio with a multivariate Lévy stable distribution. Besides the high computational cost, all these studies imply that the portfolio and its asset components must have the same distribution. For example, if a portfolio follows a multivariate GHD process, then its marginal distribution at asset level is the univariate GHD.⁵

⁴Instead of minimizing CVaR, Haque et al. (2007) adopt the safety-first portfolio selection strategy. The safety-first criterion is a risk management technique that allows an investor to select one portfolio rather than another based on the criterion that the probability of the portfolio's return falling below a minimum desired threshold is minimized.

⁵For example, generalized hyperbolic distribution can be represented as a normal mean-variance mixture where the mixture variable is generalized inverse Gaussian distributed. The multivariate GHD can be given as

$$f(\mathbf{x}) = c \frac{K_{\lambda-0.5d} \left(\sqrt{(\chi + (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})) (\psi + \boldsymbol{\gamma}' \Sigma^{-1} \boldsymbol{\gamma})} \right) \exp \{ (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} \boldsymbol{\gamma} \}}{(\chi + (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})) (\psi + \boldsymbol{\gamma}' \Sigma^{-1} \boldsymbol{\gamma})^{0.5d-\lambda}}$$

Copula is a powerful approach to construct flexible multivariate joint distributions that can reproduce various, possibly nonlinear, dependence structures between asset returns. Built upon the Sklar (1959) theorem, a copula is essentially a function that puts together unidimensional marginal distributions (or margins) to construct a multivariate distribution and its main benefit versus canonical multivariate distributions is that it is flexible enough to permit any margin. For example, Demarta and McNeil (2005) construct the meta- t distributions, i.e. distributions with a Student's t copula and univariate margins other than student t or the univariate t margins with different degrees of freedom. Shaw and Lee (2007) claim that both multivariate normal and multivariate Student's t distribution are a special case of Gaussian and Student's t copula respectively that the copulas have no requirement on the type or shape of univariate margins. The empirical study by Breymann et al. (2003) shows that the Student t copula is generally superior to the Gaussian copula when fitting financial data for the reason that the Student's t -copula can capture better phenomenon of dependent extreme values, which is often observed in financial data. In portfolio allocation applications, Di Clemente and Romano (2004) advocate static copula models, i.e. Gaussian copula, the Student t copula, the grouped t copula and the Clayton copula, to capture tail dependence for optimal loan portfolio allocation in the MCVaR framework and document that ignoring tail dependence leads to underestimating the likelihood of extreme portfolio losses, a finding confirmed by Boubaker and Sghaier (2012). Bai and Sun (2007) adopt multivariate Archimedean copula with constant dependence for optimizing portfolio based on CVaR measures. While these studies represent a significant step forward in adopting copulas to characterize the multivariate portfolio return distribution, their main limitation is that they focus on static (or unconditional) formulations that do not allow for time variation in the dependence structure among asset constituents. And yet there is ample evidence that the corre-

where the normalizing constant is $c = \frac{(\sqrt{\chi\psi})^{-\lambda} \psi^\lambda (\psi + \boldsymbol{\gamma}' \Sigma^{-1} \boldsymbol{\gamma})^{0.5d-\lambda}}{(2\pi)^{0.5d} |\Sigma|^{0.5} K_\lambda(\sqrt{\chi\psi})}$, $\{\lambda, \chi, \psi\}$ are parameters of generalized inverse Gaussian distribution, d is the rank of the covariance matrix Σ , K_λ is a modified Bessel function of the third kind with index λ , $\{\boldsymbol{\mu}, \boldsymbol{\gamma}\}$ are parameter vectors of normal mean-variance mixture distribution, and $|\cdot|$ denotes the determinant. After calibrating the GHD using in-sample data, one can simulate a large number of scenarios and then optimize the portfolio allocation by minimizing the chosen risk measure, e.g. variance or (C)VaR.

lation between asset returns, both in the central part of their distribution and in the tails, becomes stronger during crash episodes; e.g., see Patton (2006), Jondeau and Rockinger (2006) and the empirical results from Chapter 3. Such realistic features can be captured via *dynamic* or conditional copula models. However, no paper as yet has suggested the application of this flexible multivariate modeling framework in the empirical portfolio optimization literature.⁶

Our study contributes to the literature in two ways. First, we advocate a novel modeling approach that combines extreme value analysis and *dynamic* copulas in order to characterize the joint density function of portfolio assets. This allows us to demonstrate the advantages of allowing for time-varying dependence structures versus the static copula framework used thus far in the portfolio optimization literature. More specifically, we apply the EVT peaks-over-threshold approach to model the tails of the marginal distributions of individual asset returns, and then build the portfolio multivariate return density from those margins using dynamic copula. The dynamic aspect of copula is relevant because it allows investors to capture increased dependence in crisis periods. Second, we confront two portfolio optimization strategies: the traditional MV approach that uses variance as the relevant risk measure, and the MCVaR approach that relies instead on CVaR or expected shortfall. A set of out-of-sample Monte Carlo experiments are conducted to simulate the two portfolio optimization strategies assuming various data generating processes which include the best in-sample-fitted GARCH-EVT-copula model. The analysis is based on credit risk hedging portfolios formed by combining seven sectoral CDS indices. Such portfolios allow institutional investors such as investment banks and hedge funds to hedge their exposure to marketwide credit risk.

As a preview of our findings, we demonstrate empirically that the properties of optimal portfolios hinge on how asset returns and their dependence structure are modeled, as well as the choice of relevant risk measure to minimize. In particular, two modeling decisions – the selection of *distribution* (Gaussian, Student’s t or Gen-

⁶Boubaker and Sghaier (2012) allow for time-varying return dependence structures for portfolio optimization in a rather *ad hoc* way that requires re-estimating static copula over different blocks of data of arbitrary length.

eralized Error Distribution) to model individual returns and portfolio returns, and the assumption of static versus *dynamic* (copula) dependence structures – are shown to have a significant impact on the efficient frontier of portfolio optimization. Our findings confirm that MV models underestimate the risk of rare extreme losses and thus can only generate suboptimal portfolios for a given target return compared to the MCVaR approach. The simulations show that the MCVaR strategy facilitates a more attractive risk-return tradeoff than the standard MV approach. Even though portfolio managers may invest “efficiently” on the basis of the MV model, they may suffer catastrophically large losses. This makes the MCVaR framework very appealing. These findings have implications for risk hedgers and portfolio managers that are seeking to allocate optimally their risk capitals to achieve the optimal risk-return balance while minimizing the probability of severe tail losses.

The remainder of the paper is organized as follows. In Section 4.2, we introduce the EVT-copula framework adopted to characterize the multivariate distribution of the portfolio. Section 4.3 outlines the two optimization strategies based on the classical MV approach and the MCVaR approach, respectively. Section 4.4 presents the data and empirical results. Section 4.5 concludes.

4.2 GARCH-EVT-Copula Modeling of Portfolio Returns

Portfolio optimization requires estimating the joint probability distribution of asset returns and choosing a risk metric to focus upon. This section discusses the former task while the issue of choosing a risk metric is discussed in Section 4.3. Essentially, we advocate an approach where the portfolio density function is built in two steps. First, by applying EVT special efforts are aimed at modeling the lower tail of the univariate return distributions. Second, the margins thus obtained are joined using dynamic copula models that allow for time-variation in dependence structures and nest static copula models.

EVT is formulated for independent and identically distributed (*iid*) processes but there is ample evidence for various asset classes that financial returns exhibit autocorrelation and volatility clustering. The typical approach is to apply EVT to appropriately ARMA-GARCH filtered returns; *e.g.*, Diebold et al. (2000) and McNeil and Frey (2001). In what follows, uppercase $F(\cdot)$ denotes a cumulative distribution function (*cdf*) and lowercase $f(\cdot)$ denotes a probability density function (*pdf*).

4.2.1 ARMA-GARCH filtering of asset returns

Let $\{r_t\}$, $t = 1, \dots, T$, denote daily logarithmic returns for an asset. We conceptualize the underlying probability distribution as belonging to a location-scale family of the form

$$r_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t x_t$$

where location μ_t and scale σ_t are ARMA and GARCH processes, respectively. The ARMA-GARCH model is

$$r_t = a_0 + \sum_{i=1}^p a_i r_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t \quad (4.1)$$

$$\sigma_t^2 = c_0 + \sum_{i=1}^m c_i \sigma_{t-i}^2 + \sum_{j=1}^n d_j \varepsilon_{t-j}^2 \quad (4.2)$$

where the standardized (filtered) returns $x_t \equiv \frac{r_t - \mu_t}{\sigma_t} = \frac{\varepsilon_t}{\sigma_t}$ are *iid* with zero-location and unit-scale probability density $f(x; \boldsymbol{\theta})$. Various candidate distributions are considered in our analysis: Gaussian, Student's t , Generalized Error Distribution (GED), and skewed versions thereof. The ARMA-GARCH model parameters are estimated by Quasi Maximum Likelihood.

GED is a symmetric unimodal member of the exponential family with density function given by

$$f(x; \mu, \beta, \kappa) = \frac{\kappa e^{-\frac{1}{2} \left| \frac{x - \mu}{\beta} \right|^\kappa}}{2^{1 + \frac{1}{\kappa}} \beta \Gamma(\kappa^{-1})}. \quad (4.3)$$

which is characterized by three parameters, *location* $\mu \in (-\infty, +\infty)$, *scale* $\beta \in (0, \infty)$

and *shape* $\kappa \in (0, \infty)$; and $\Gamma(\cdot)$ is the gamma function. Since GED is symmetric and unimodal, both the median and the mean are given by μ . Compared with the Student's t density, GED has a sharper peak at μ with the sharpness (and thickness of tails) increasing as κ decreases. GED nests important distributions: for $\kappa = 2$, it becomes the Gaussian distribution, $N(\mu, \beta^2)$; for $\kappa = 1$, it becomes the Double Exponential or Laplace distribution, $L(\mu, 4\beta^2)$; as $\kappa \rightarrow \infty$, it converges to the Uniform distribution $U(\mu - \beta, \mu + \beta)$.

The central moments of GED are given by

$$\mathbb{E}[(x - \mu)^n] = \frac{1}{2^{1+\frac{1}{\kappa}}\beta\Gamma(\frac{1}{\kappa} + 1)} \int_{-\infty}^{\infty} (x - \mu)^n e^{-\frac{1}{2}\left|\frac{x-\mu}{\beta}\right|^\kappa} dx. \quad (4.4)$$

Symmetry implies zero odd ($n = 1, 3, \dots$) moments. Even moments can be rewritten as

$$\begin{aligned} \mathbb{E}[(x - \mu)^n] &= \frac{2^{\frac{n}{\kappa}}\beta^n}{\Gamma(\kappa^{-1})} \int_0^{\infty} t^{\frac{1}{\kappa}(n+1)-1} e^{-t} dt. \\ &= 2^{\frac{n}{\kappa}}\beta^n \frac{\Gamma(\kappa^{-1}(n+1))}{\Gamma(\kappa^{-1})}, \quad n = 2, 4, \dots \end{aligned} \quad (4.5)$$

Thus the variance and kurtosis are given by $\mathbb{E}[(x - \mu)^2] \equiv \sigma^2 = 2^{2\kappa^{-1}}\beta^2 \frac{\Gamma(3\kappa^{-1})}{\Gamma(\kappa^{-1})}$ and $\mathbb{E}[(x - \mu)^4] = \frac{\Gamma(5\kappa^{-1})\Gamma(\kappa^{-1})}{\Gamma^2(3\kappa^{-1})}$, respectively. The standardized GED density function can be expressed as

$$f(z) = f\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma} \frac{\kappa e^{-\frac{1}{2}\left|\frac{x-\mu}{\sigma\lambda}\right|^\kappa}}{\lambda 2^{1+\kappa^{-1}}\Gamma(\kappa^{-1})} \quad (4.6)$$

where $\lambda = \sqrt{\frac{\Gamma(\kappa^{-1})}{2^{2/\kappa}\Gamma(3\kappa^{-1})}}$, and z has unit standard deviation $Var(z) = 1$. The *cdf* is given by as

$$F(z) = F\left(\frac{x - \mu}{\sigma}\right) = \begin{cases} 0.5(1 + \Gamma_{|c_0 z|^\kappa}(\kappa^{-1})) & x \geq \mu \\ 0.5(1 - \Gamma_{|c_0 z|^\kappa}(\kappa^{-1})) & x < \mu \end{cases} \quad (4.7)$$

where $c_0 = \sqrt{\Gamma(3\kappa^{-1})/\Gamma(\kappa^{-1})}$.

Fernandez and Steel (1998) introduce asymmetry into unimodal densities (*e.g.*, Gaussian, Student's t and GED) in an elegant manner by adding a skewness parameter $\xi \in \mathbb{R}^+$ as follows

$$f(x; \boldsymbol{\theta}, \xi) = \frac{2}{\xi + \xi^{-1}} [f(\xi x) H(-x) + f(\xi^{-1} x) H(x)]$$

where $\boldsymbol{\theta}$ is the parameter vector of the original density and ξ is the additional skewness parameter, and $H(\cdot)$ is the Heaviside function. The original density corresponds to $\xi = 1$. The first two moments are given by

$$\begin{aligned} \mathbb{E}[x] &\equiv \mu = M_1 (\xi - \xi^{-1}) \\ \mathbb{E}[(x - \mu)^2] &\equiv \text{Var}(x) = (M_2 - M_1^2) (\xi^2 + \xi^{-2}) + 2M_1^2 - M_2 \end{aligned}$$

where $M_k = 2 \int_0^\infty x^k f(x) dx$ with $k = 1, 2$.

4.2.2 Extreme Value Theory

Analogous with the central limit theorem, where the normal distribution acts the limit for the distribution of the mean of a large number *i.i.d.* random variables, the extreme value theory investigates the limit distribution of the sample maximum. Empirical models of financial returns based on distributional assumptions such as Gaussian, Student's t and GED are often chosen based on their ability to fit data near the mode given that only a few observations fall in the distribution tails by definition. But effective risk management requires accurate estimation of the likelihood of rare events that could trigger catastrophic losses. Extreme value theory can be useful for this purpose because it is specifically aimed at modeling tail behavior without requiring assumptions on the entire distribution, *i.e.* it provides a semi-parametric model for the tails of distribution functions.⁷ *Block maxima* and *peaks-over-threshold*

⁷EVT was pioneered by Fisher and Tippett (1928) and Jenkinson (1955). Early applications of EVT, dating back to the 1950s, were aimed at addressing environmental questions for insurance analysis. It is rather more recently that EVT has been utilized in finance; see Embrechts et al. (1997) for a comprehensive review of financial applications.

are the two main EVT modeling methodologies which we describe next.

Let $\{x_t\}$, $t = 1, \dots, T$, denote an *iid* process with distribution $F(x)$. The maximum of a block of $n < T$ observations, called *block maximum* and denoted $M_n = \max(x_1, \dots, x_n)$, follows asymptotically the probability distribution

$$\mathbb{P} \left[\frac{M_n - b_n}{a_n} \leq y \right] = F^n(a_n y + b_n) \rightarrow G(y), \quad n \rightarrow +\infty \quad (4.8)$$

as $n \rightarrow +\infty$ for all $y \in \mathbb{R}$, where $a_n > 0$ and b_n are appropriate constants, $F^n(\cdot)$ is $F(\cdot)$ raised to power of n , and $G(\cdot)$ is a non-degenerate distribution function. According to the Extremal Types Theorem, the *block maxima* distribution $G(\cdot)$ must be either Frechet, negative Weibull or Gumbel; these three distributions can be cast as members of the Generalized Extreme Value (GEV) distribution with *cdf* given by

$$G(y) = \begin{cases} \exp \left\{ - \left(1 + \xi \frac{y - \mu}{\beta} \right)^{-1/\xi} \right\} & \xi \neq 0 \\ \exp \left\{ -e^{-\frac{y - \mu}{\beta}} \right\} & \xi = 0 \end{cases}, \quad (4.9)$$

where $\mu, \beta > 0$ and ξ are *location*, *scale* and *shape* parameters, respectively. GED becomes the Frechet distribution for $\xi > 0$, the negative Weibull distribution for $\xi < 0$, and the Gumbel distribution for $\xi = 0$.

Let $\{x_t - u\}$ $t = 1, \dots, T$, denote the *exceedances* or *peaks-over-threshold* process where $x_t > u$ and u denotes a threshold loss. The *exceedances* distribution can be formalized as

$$\Pr[x_t - u \leq y \mid x_t > u] = \frac{F(y + u) - F(u)}{1 - F(u)} \rightarrow H(y), \quad t = 1, \dots, T.$$

According to the Pickands-Balkema-de-Haan Theorem, for a sufficiently large threshold loss u , the *exceedances* distribution can be approximated by the Generalized

Pareto Distribution (GPD) as⁸

$$H(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta}\right)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp\left\{-\frac{y}{\beta}\right\} & \xi = 0 \end{cases}, \quad (4.10)$$

where $\beta > 0$ and ξ are *scale* and *shape* parameters, respectively. GPD nests the exponential distribution ($\xi = 0$), the heavy-tailed Pareto Type I distribution ($\xi > 0$) and the short-tailed Pareto Type II distribution ($\xi < 0$).

In sum, asymptotic EVT suggests modeling either *block maxima* using a GEV distribution, or *peaks-over-threshold* using a GPD. However, the former approach makes inefficient use of data when many extreme observations occur intensively in a given block of time, given that only one observation (*i.e.*, block maximum) is recorded.⁹ In our empirical analysis below, we adopt the *peaks-over-threshold* approach to characterize the behavior of (filtered) return exceedances. The parameters of GPD are estimated by maximizing the corresponding log-likelihood function

$$\begin{aligned} \ln \mathfrak{L}(y_1, \dots, y_{N_u}; \beta, \xi) &= \sum_{j=1}^{N_u} \ln h(y_j; \beta, \xi) \\ &= -N_u \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{j=1}^{N_u} \ln \left(1 + \xi \frac{y_j}{\beta}\right) \end{aligned}$$

where N_u is the total number of observed exceedances $y_j \equiv x_j - u$ for given threshold u .

For financial applications with emphasis on the occurrence of extraordinary events, McNeil and Frey (2001) highlight as main advantages of extreme value analysis that: 1) It builds upon a sound statistical theory and lends itself as a parametric modeling approach for the tail distribution, 2) It circumvents the need for assumptions regarding the entire return distribution, *i.e.* the central part of the distribution can

⁸See Balkema and de Haan (1974) and Pickands (1975).

⁹Take the problem of modeling the “heat wave” phenomenon as an example. Very high temperatures are often recorded over consecutive days in the summer. By focusing on *block maxima*, most of those observations will be neglected and instead, irrelevant maximum temperatures recorded in the winter will be included.

be estimated empirically, and therefore model uncertainty is considerably reduced.

4.2.3 Dynamic Copula

Copula has become an increasingly popular approach to model dependence among asset returns and this is mainly due to its flexibility. Multivariate distributions can be nested as special cases of copulas. However, a problem of conventional copula models is that they ignore the dynamics of dependence. Only recently, Sklar's theorem was proven in the context of conditional distributions by Patton (2006) who then laid the theoretical foundations for time-varying (or dynamic) copula. We follow Jondeau and Rockinger (2006)'s dynamic copula formulation which is inspired from Engle (2002) dynamic conditional covariance (DCC) model. The DCC type dynamic copula model¹⁰ has been discussed in Section 3.2.2 of Chapter 3 and assume that the evolution of the dependence structure of a k -dimensional portfolio in an elliptical copula (e.g. Gaussian or Student's t ¹¹) is characterized by the following process:

$$\mathbf{Q}_t = (1 - \varphi - \psi) \bar{\mathbf{Q}} + \varphi \mathbf{Q}_{t-1} + \psi \epsilon_{t-1} \cdot \epsilon'_{t-1}, \quad \varphi + \psi < 1; \varphi, \psi \in (0, 1) \quad (4.11)$$

$$\mathbf{R}_t = \tilde{\mathbf{Q}}_t^{-1} \mathbf{Q}_t \tilde{\mathbf{Q}}_t^{-1} \quad (4.12)$$

where $\bar{\mathbf{Q}}$ is the unconditional covariance of $\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, \dots, \epsilon_{k,t})'$ estimated by its sample analogue $\bar{\mathbf{Q}} = T^{-1} \sum_{t=1}^T \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}'_t$, where $\epsilon_i = F_i^{-1}(u_i)$ and u are the transformed standardized residuals via the empirical *cdf* transformation. $\mathbf{Q}_t$ is the auxiliary matrix driving the rank correlation dynamics. $\tilde{\mathbf{Q}}_t$ is a $k \times k$ diagonal matrix with diagonal elements the square root of those of $\mathbf{Q}_t$ and $\mathbf{R}_t$ is a $k \times k$ rank correlation *matrix*. DCC copulas have been shown to be quite successful in modeling the dependence of financial returns; see Jondeau and Rockinger (2006) and the discussion in Chapter 3.

¹⁰Considering the difficulty of applying ARMA fashion dynamic copula to high dimensional contexts, we only adopt the DCC dynamic copula in this empirical study.

¹¹Appendix (4.B) outlines the *pdf* of the multi-dimensional Gaussian and Student's t copulas.

4.3 Portfolio optimization Methods

One of the most important elements of Markowitz (1952)'s Modern Portfolio Theory is the notion of *efficient frontier* in the mean-variance space. A classical mean-variance (MV) portfolio strategy consists of minimizing the portfolio risk, proxied by the variance of the joint distribution, subject to a target portfolio return. Formally,

$$\begin{aligned} \min_{\boldsymbol{\omega}} Var &= \min_{\boldsymbol{\omega}} \boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega} \\ \text{subject to } \boldsymbol{\omega}'\boldsymbol{\mu} &= g \\ \boldsymbol{\omega}'\boldsymbol{I} &= 1 \end{aligned}$$

where $\boldsymbol{\Sigma}$ is the estimated covariance matrix of asset returns, $\boldsymbol{\mu}$ is the estimated vector of expected asset returns, $\boldsymbol{I}$ is a vector of ones, g is the *a priori* chosen portfolio target return, and $\boldsymbol{\omega}$ is the resulting optimal vector of weights. The efficient frontier is constructed by solving the problem for different values of g .

Adopting the variance as portfolio risk measure has the advantage of computational simplicity, but it may lead to unsatisfactory outcomes for a risk manager because it implicitly assumes symmetry or equal probabilities of losses and profits, and it underestimates the chance of rare adverse events. Partly in recognition of its widespread use by banks, since 1996 the Basel Committee for banking supervision began to state some of the risk management requirements in terms of percentiles, namely, Value-at-Risk (VaR) of loss distributions. Current regulations impose capital requirements on banks and financial institutions proportional to the VaR of a portfolio. VaR has established itself as the most popular risk metric for determining the largest size of losses in trading books at a given confidence level. Thus, for instance, 99% VaR is an estimate of the maximum portfolio loss which is exceeded with 5% probability. However, an important shortcoming of VaR is that it is not coherent because, for non-normal distributions, it fails to satisfy the subadditivity property. VaR is thus inappropriate for portfolio optimization; see Appendix 4.A. An alternative coherent risk metric proposed by Rockafellar and Uryasev (2000) is

called Conditional Value at Risk (CVaR) or Expected Shortfall. CVaR is defined as the expected loss exceeding VaR and thus it represents an upper bound for VaR. Formally,

$$CVaR^\alpha \equiv \mathbb{E}(-r > VaR^\alpha), \quad (4.13)$$

where VaR^α denotes the maximum loss at confidence level $\alpha \in (0, 1)$ typically chosen at 0.99 or 0.95 and r denotes the portfolio loss. It follows from $CVaR^\alpha \geq VaR^\alpha$ that, if the risk manager can control CVaR then he can control VaR but not the other way round.

A thorny issue with CVaR is that it is difficult to compute. Let $r(\boldsymbol{\omega}, \boldsymbol{\mu})$ be a portfolio return function where $\boldsymbol{\omega}$ and $\boldsymbol{\mu}$ are vectors of weights and expected asset returns, respectively. We can rewrite (4.13) as follows

$$CVaR^\alpha(\boldsymbol{\omega}) = \frac{1}{\alpha} \int_{-r(\boldsymbol{\omega}, \boldsymbol{\mu}) > VaR^\alpha(\boldsymbol{\omega}, \boldsymbol{\mu})} r(\boldsymbol{\omega}, \boldsymbol{\mu}) f(\boldsymbol{\mu}) d\boldsymbol{\mu}$$

where $f(\boldsymbol{\mu})$ denotes the multivariate *pdf* of asset returns. Rockafellar and Uryasev (2000) propose an alternative simpler function

$$F^\alpha(\boldsymbol{\omega}, d) = d + \frac{1}{\alpha} \int_{-r(\boldsymbol{\omega}, \boldsymbol{\mu}) > d} (-r(\boldsymbol{\omega}, \boldsymbol{\mu}) - d) f(\boldsymbol{\mu}) d\boldsymbol{\mu}. \quad (4.14)$$

and demonstrate that $F^\alpha(\boldsymbol{\omega}, d)$ is a convex function with respect to d , and that VaR is a minimum point of this function with respect to d . So in the mean-CVaR framework, where variance is replaced by CVaR as the relevant risk metric, the optimization problem becomes $\min_{\boldsymbol{\omega}, d} F^\alpha(\boldsymbol{\omega}, d) = \min_{\boldsymbol{\omega}} CVaR^\alpha(\boldsymbol{\omega})$. Rockafellar and Uryasev (2000) and Andersson et al. (2001) suggest to approximate (4.14) by Monte Carlo simulation as follows

$$F^\alpha(\boldsymbol{\omega}, d) = d + \frac{1}{\alpha N} \sum_{i=1}^N (-r(\boldsymbol{\omega}, \boldsymbol{\mu}_i) - d)^+, \quad (4.15)$$

where N denotes the number of replications, and $z^+ = \max(0, z)$. This optimization

can be approached as a linear programming problem

$$\begin{aligned}
& \min_{\boldsymbol{\omega}, z, d} d + \frac{1}{\alpha N} \sum_{i=1}^N z_i & (4.16) \\
& \text{subject to } z_i \geq -r(\boldsymbol{\omega}, \boldsymbol{\mu}_i) - d; \\
& z_i \geq 0; \\
& \boldsymbol{\omega}' \mathbf{I} = 1; \\
& \boldsymbol{\omega}' \boldsymbol{\mu}_i = g.
\end{aligned}$$

where $\boldsymbol{\omega}$ is the MCVaR optimal vector of weights.

4.4 Data and Empirical Results

4.4.1 Data description and preliminary statistics

We consider portfolios comprising seven iTraxx sectoral CDS Europe indices: Automobile (Auto), Consumers (Cons), Energy, Industrials (Inds), Telecommunications, media and technology (TMT), Financials Senior (FinSnr) and Financials Subordinated (FinSub). The observations from *Bloomberg* are daily midpoint index CDS spread quotes over the period running from 20 September, 2005 to 11 March, 2011 ($T = 1371$ observations). All index spread quotes are based on five-year maturity single-name CDS contracts which are typically the most liquid in the market.

Over the past two decades the global credit derivative market has seen extraordinary growth and has provided ample trading opportunities for asset managers. Credit derivatives and, in particular, CDS indices allow investors and speculators to monitor the performance of the market and take credit positions without owning the underlying credit assets (loans or bonds). A portfolio consisting of sectoral CDS indices would allow institutional investors to insure themselves against marketwide credit risk. Moreover, should the correlation between sectoral index CDS returns be low it would make such a portfolio desirable for diversification purposes. The sample correlations bear this out. On average the Pearson (linear) correlation is 0.214 ranging

from a low of 0.040 between Cons and TMT to a high of 0.627 between SnrFin and SubFin. Likewise, the average Kendall (nonlinear) rank correlation is 0.237 ranging from a low of 0.162 for Inds and SubFin, to a high of 0.519 for SnrFin and SubFin.

Figure 4.1, Panel (a), plots the index CDS spreads normalized to a basis of 1.¹² A clear comovement of indices can be observed. A fast increase in CDS spreads for all sectors begins to gather pace in mid-2007 when the US housing bubble bursts. CDS spreads peak towards the end of 2008 following the disastrous collapse of Lehman Brothers. Thereafter the spreads gradually decrease roughly to pre-Lehmann levels until a second and third waves of dramatic increases in CDS spreads occurred during 2010 and 2011, respectively, reflecting the Greek and European sovereign debt crises. Figure 4.1, Panel (b), plots the daily returns for the sectoral CDS indices. All returns are calculated as logarithmic midpoint-to-midpoint spread quote ratios. Heteroskedasticity is clearly apparent as volatility clustering is clearly visible.

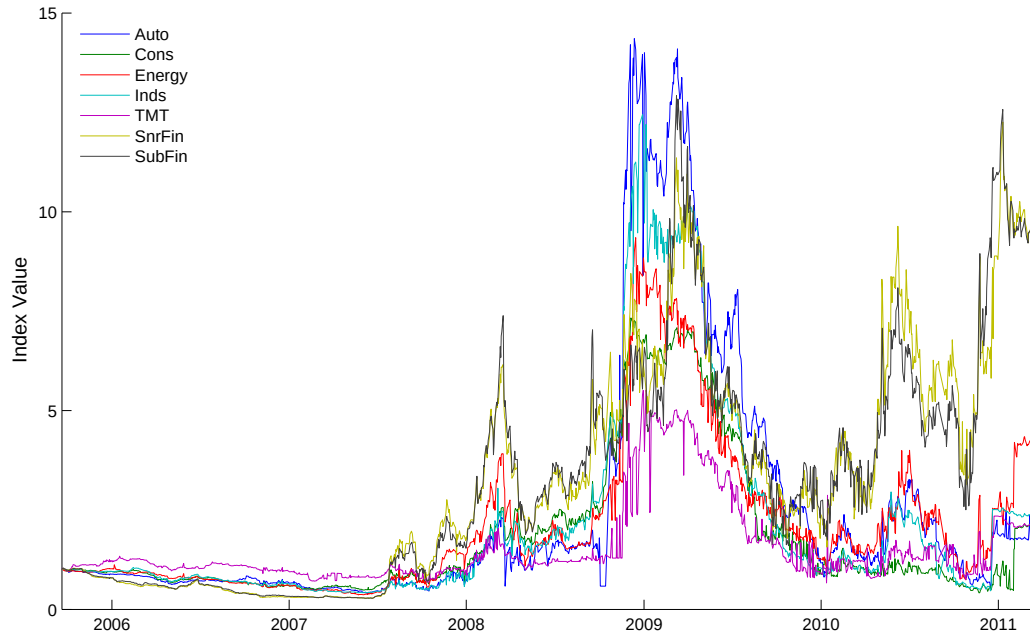
Summary statistics for the daily returns series are reported in Table 4.1. FinSnr CDS contracts have the largest mean return while Auto CDS contracts have the lowest. TMT CDS contracts are the most volatile. All CDS return distributions show heavy tails, particularly, Autos and Industrials stand out and also show relatively large positive skewness. The Jarque-Bera test confirms significant departures from non-normality for all sectoral CDS index returns.

Both the Ljung-Box Q test and Engle's ARCH LM test reject the null of no autocorrelation for lags up to ten in returns and squared returns confirming, respectively, serial dependence and heteroskedasticity. Figure 4.2 displays quantile-to-quantile (QQ) plots for CDS returns which represent the sample quantiles against theoretical quantiles from a normal distribution. The plots are essentially symmetric around the 45° line (i.e., $y = x$) but the type of deviations observed at both extremes - left end is below the line and right end is above the line - reflect the strong fat-tailedness of the CDS return distribution.

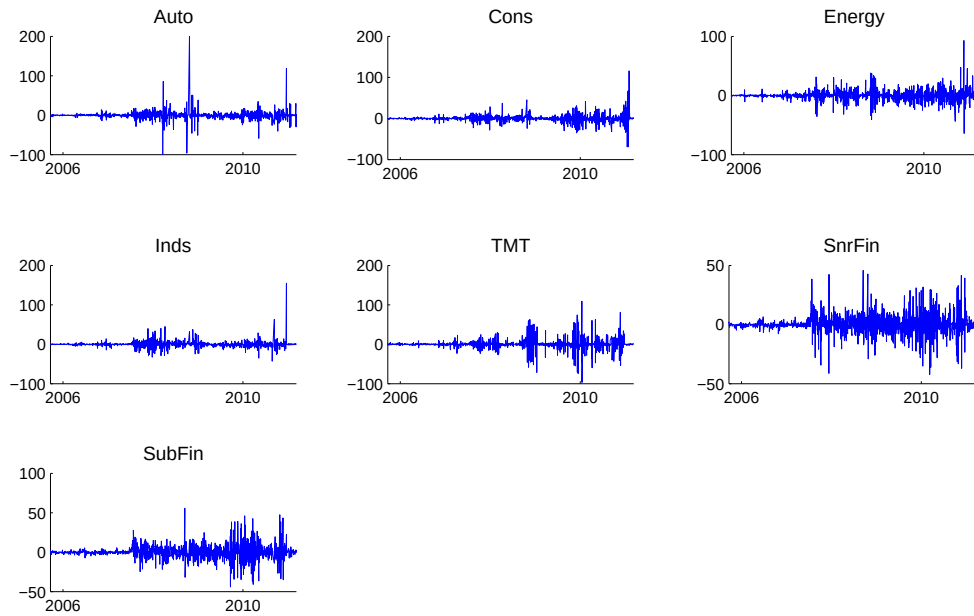
¹²The empirical analysis is performed using MATLAB 2012a.

Figure 4.1: Sectoral iTraxx Europe Indices

(a) *Relative Daily Index Mid-points*



(b) *Daily Logarithmic Returns of CDS Spreads*



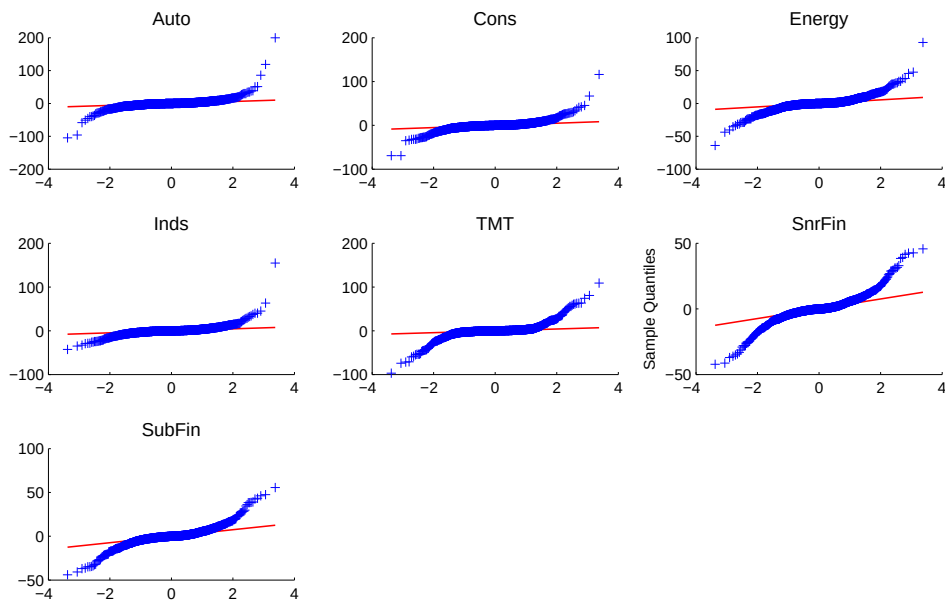
Notes: The top figure plots the daily levels of iTraxx CDS indices with all series normalized to start at 1. The bottom figure plots the daily logarithmic returns.

Table 4.1: Summary Statistics of Daily Logarithmic Return

	Auto	Cons	Energy	Inds	TMT	SnrFin	SubFin
Mean	0.043	0.057	0.108	0.065	0.057	0.168	0.167
Median	-0.009	-0.051	0.000	-0.097	0.000	-0.033	-0.055
Max	199.852	116.015	92.803	154.846	108.893	45.735	55.641
Min	-104.745	-69.126	-64.028	-42.751	-96.817	-42.202	-43.987
Std. Dev.	10.801	8.388	8.293	7.930	12.436	7.996	8.358
Skewness	4.650	1.657	0.837	5.929	0.364	0.206	0.485
Kurtosis	114.039	42.502	21.604	115.302	20.623	10.079	11.149
Observation	1370	1370	1370	1370	1370	1370	1370
Jarque-Bera	708751	89702	19917	727951	17760	2870	3845
<i>p</i> -value	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Ljung-Box(10)	57.580	152.430	62.241	21.782	205.577	59.512	61.490
<i>p</i> -value	0.000	0.000	0.000	0.001	0.000	0.000	0.000
ARCH(10)	47.953	165.736	113.140	45.941	227.697	121.797	162.844
<i>p</i> -value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: This table presents summary statistics of the logarithmic daily return series (%) used over period from 20 September, 2005 to 11 March, 2011.

Figure 4.2: QQ Plot of Normal Distribution



Notes: This figure displays the quantile-quantile plot of the sample quantiles versus theoretical quantiles from a normal distribution.

4.4.2 Joint probability density function

We begin by identifying appropriate lags for the ARMA-GARCH filter, assuming Gaussian innovations, in order to remove serial correlation and heteroskedasticity in the CDS returns. The filtered returns are essentially *iid* as borne out in the autocorrelation functions plotted in Figure 4.3.

Using the lag orders thus identified, we now compare the ARMA-GARCH model under various assumptions for the innovations, Gaussian, Student's t , GED and their skewed versions.¹³ The maximized log-likelihood value (LL), the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are employed as gauge of fit. As shown in Table 4.2, the best univariate return distributions or marginals are achieved with GED. Unsurprisingly, the Gaussian distribution provides a very poor fit. The parameter estimates of the best fit ARMA-GARCH-GED model are shown in Table 4.3.¹⁴ The GED tail-thickness parameter κ is significantly below 2 for all CDS indices implying heavier tails than the normal. The Ljung-Box Q(10) test and ARCH LM(10) test indicate no serial correlation or volatility clustering, respectively, in the filtered returns.

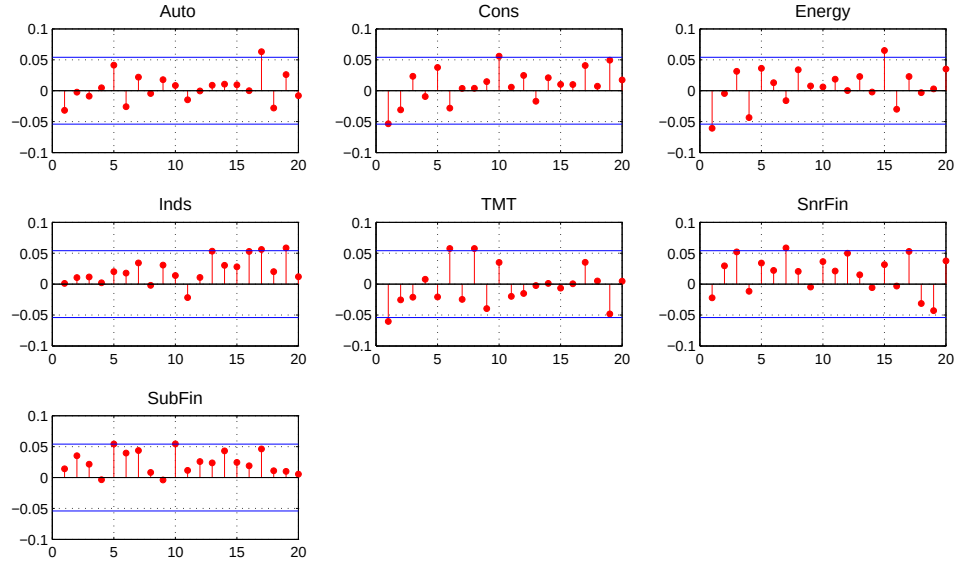
The QQ plot is employed to visually check whether the filtered returns come from the GED distribution. Figure 4.4 shows the quantiles of empirical distribution against the hypothesized GED distribution indicating that the central quantiles fit fairly well in all return series. It can be seen that, while a fairly good fit is achieved in the high-density (central) area for all filtered returns, this is not so in the low-density (tail) areas. The deviations at far-ends indicate that the tails may have different characteristics with the central areas of distribution and ignoring this fact by simply using the same theoretical distribution to describe the whole sample may result in misspecification.

¹³The identified ARMA lag orders are $p = q = 0$ for all sectors except Cons and SnrFin for which $p = 0$, $q = 1$, and $p = 2$, $q = 0$, respectively. The identified GARCH lag orders are $m = n = 1$ throughout.

¹⁴Although the skewed GED fits the Auto and financial CDS index returns slightly better according to the AIC, given the much higher computational costs of this distribution, it is reasonable to focus on the GED throughout.

Figure 4.3: Correlograms of ARMA-GARCH filtered returns

(a) *Levels*



(b) *Squares*

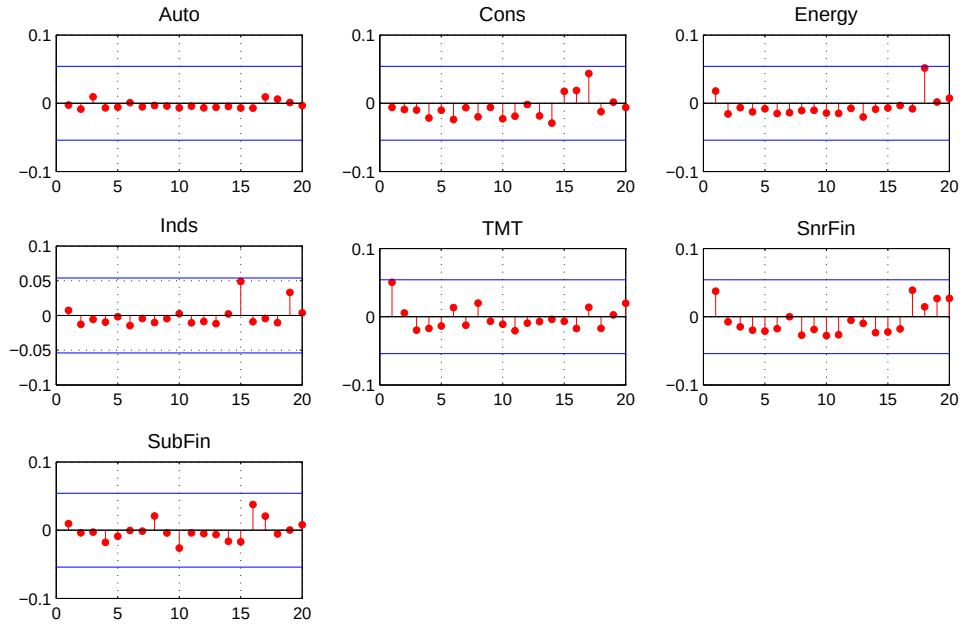


Table 4.2: Univariate Model Selection

		Auto	Cons	Energy	Inds	TMT	SnrFin	SubFin
Normal	LL	-4488.55	-4098.73	-4380.40	-4126.77	-4515.24	-4248.02	-4162.53
	AIC	8987.10	8207.47	8770.80	8261.53	9040.78	8508.05	8337.07
	BIC	9013.21	8233.58	8796.92	8282.42	9066.59	8539.38	8368.40
Skew Normal	LL	-4498.16	-4096.97	-4424.82	-4112.46	-4554.16	-4247.68	-4164.53
	AIC	9006.33	8205.95	8859.64	8234.91	9118.33	8509.35	8339.07
	BIC	9032.44	8237.28	8885.75	8261.02	9144.44	8545.91	8365.18
Student's t	LL	-4030.65	-3881.12	-4060.32	-3727.89	-4030.11	-4093.18	-4058.75
	AIC	8071.30	7774.24	8132.65	7465.77	8076.21	8200.37	8131.49
	BIC	8097.41	7805.58	8163.98	7491.89	8107.55	8236.93	8168.05
Skew Student's t	LL	-4029.81	-3881.12	-4059.62	-3727.44	-4030.96	-4091.97	-4057.62
	AIC	8070.62	7776.23	8133.24	7466.89	8075.92	8199.95	8127.26
	BIC	8102.96	7812.79	8169.80	7498.22	8112.48	8241.73	8158.60
GED	LL	-4004.18	-3844.70	-4005.89	-3708.94	-3866.20	-4089.02	-4048.21
	AIC	8018.36	7701.40	8021.77	7427.88	7742.40	8192.05	8110.42
	BIC	8044.47	7732.74	8047.88	7453.99	7768.52	8228.61	8146.98
Skew GED	LL	-4004.48	-3844.70	-4005.91	-3708.92	-3887.20	-4086.18	-4046.24
	AIC	8020.95	7703.40	8025.81	7429.85	7748.40	8188.37	8108.49
	BIC	8052.29	7739.96	8062.37	7461.18	7784.96	8230.15	8150.27

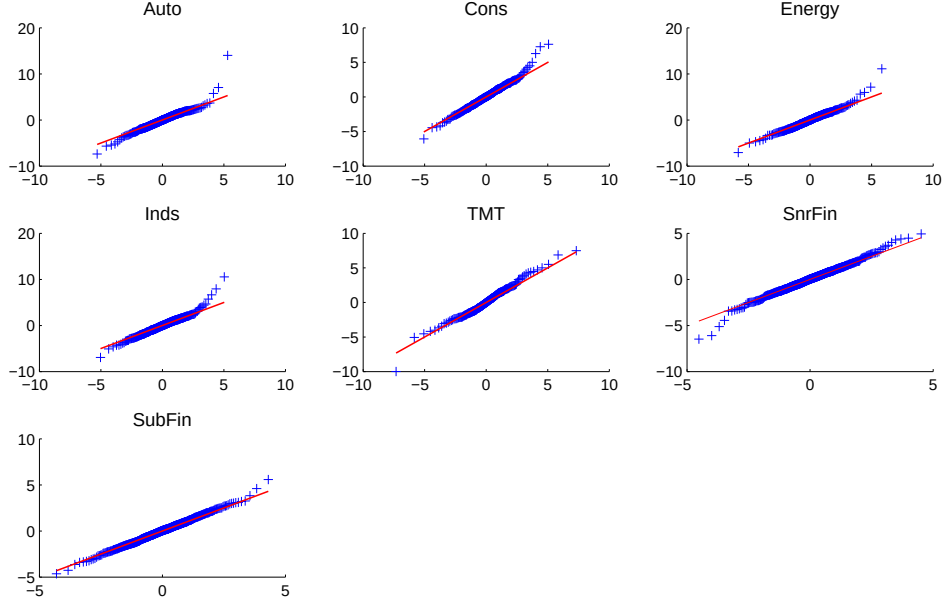
Notes: This table compares five competing marginal distributions fitted to filtered returns via three statistical measures: log-likelihood (LL) value, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) at the optimum. The bold numbers denote the best selection.

Table 4.3: Estimated Parameters for ARMA-GARCH-GED model

	Auto	Cons	Energy	Inds	TMT	SnrFin	SubFin
<i>Conditional Mean Equation:</i>							
<i>Intercept</i>	-0.103** (0.018)	-0.106** (0.009)	0.000 (0.000)	-0.051** (0.008)	0.000 (0.000)	-0.141** (0.014)	-0.286** (0.021)
<i>AR1</i>	-	-	-	-	-	-0.026** (0.001)	-
<i>AR2</i>	-	-	-	-	-	0.061** (0.001)	-
<i>MA1</i>	-	-0.044** (0.000)	-	-	-	-	-
<i>Conditional Variance Equation:</i>							
<i>Intercept</i>	0.096 (0.077)	0.248 (0.255)	0.175 (0.107)	0.111 (0.097)	1.069 (0.824)	0.079* (0.041)	0.080** (0.035)
<i>ARCH1</i>	0.152** (0.029)	0.203** (0.072)	0.137** (0.035)	0.223** (0.050)	0.161** (0.027)	0.133** (0.030)	0.155** (0.022)
<i>GARCH1</i>	0.847** (0.039)	0.796** (0.091)	0.862** (0.046)	0.776** (0.064)	0.839** (0.069)	0.866** (0.031)	0.844** (0.022)
κ	0.722** (0.065)	0.782** (0.044)	0.613** (0.055)	0.789** (0.057)	0.398** (0.075)	0.952** (0.054)	1.038** (0.064)
<i>Ljung-Box test on standardized residuals:</i>							
<i>Q(10)</i>	0.772	0.061	0.176	0.920	0.053	0.054	0.057
<i>Q(15)</i>	0.952	0.338	0.095	0.704	0.062	0.072	0.073
<i>Q(20)</i>	0.773	0.361	0.059	0.097	0.059	0.076	0.129
<i>Ljung-Box test on squared standardized residuals:</i>							
<i>Q(10)</i>	1.000	0.954	0.993	1.000	0.429	0.599	0.971
<i>Q(15)</i>	1.000	0.971	0.999	0.922	0.753	0.770	0.997
<i>Q(20)</i>	1.000	0.640	0.953	0.970	0.883	0.739	0.997
<i>ARCH LM test:</i>							
<i>ARCH(10)</i>	0.948	0.920	0.676	0.860	0.067	0.363	0.930
<i>ARCH(15)</i>	0.997	0.952	0.951	0.993	0.135	0.651	0.983
<i>ARCH(20)</i>	1.000	0.963	0.992	1.000	0.434	0.786	0.994

Notes: The estimated parameters correspond to the ARMA-GARCH(1,1) return evolution process with GED residuals according to Eq.(4.1), Eq.(4.2), and Eq.(4.6). κ denotes shape parameter of GED. The Standard errors are in parentheses. ** and * denote statistical significance at 5% and 10% confidence level, respectively.

Figure 4.4: QQ Plot of Generalized Error Distribution



Notes: This figure displays the quantile-quantile plot of the sample quantiles versus theoretical quantiles from a General Error distribution.

To improve how well the marginal models characterize the CDS index returns at the tails, we resort to EVT and deploy the *peaks-over-threshold* modeling approach described in Section 4.2.2. Setting a left (right) tail threshold at the 5th (95th) quantile¹⁵ of the filtered returns enables about 70 exceedances at each tail to estimate the GPD parameters. The threshold values and GPD parameter estimates are set out in Table 4.4.

The shape parameter ξ is significantly positive suggesting that the tails of filtered returns conform to a heavy-tailed Pareto Type I distribution. Moreover, the GPD parameters differ at the upper and lower tails which indicates asymmetry. The Welch's *t*-test rejects the null hypothesis of symmetry in the scale parameter $H_0 : \beta^U = \beta^L$ at the 5% or 1% in all but one case (Energy) and in the shape parameter $H_0 : \xi^U = \xi^L$ in all cases at the 1% level. By comparing the empirical *cdf* of the tails to the fitted *cdf* in Figure 4.5 we can conclude that the estimated GPD fits quite well the filtered

¹⁵A rule of thumb of determining u outlined by Hull and White (2006) is that u should be approximately equal to the 5th (95th) percentage of the empirical distribution. Here We follow McNeil and Frey (2001).

Table 4.4: Estimates of EVT Model of Upper and Lower Tail Behavior

	Auto	Cons	Energy	Inds	TMT	SnrFin	SubFin
u^L	-1.819	-1.706	-1.768	-1.892	-1.413	-1.681	-1.740
ξ^L	0.262*	0.085	0.288*	0.150	0.326**	0.187**	0.021
	(0.145)	(0.137)	(0.147)	(0.141)	(0.153)	(0.074)	(0.142)
β^L	0.857**	0.833**	0.723**	0.838**	0.741**	0.801**	0.731**
	(0.174)	(0.153)	(0.136)	(0.156)	(0.142)	(0.153)	(0.138)
u^U	1.655	1.755	1.592	1.640	1.527	1.678	1.794
ξ^U	0.511**	0.385**	0.479**	0.590**	0.190	0.072	0.201
	(0.158)	(0.172)	(0.188)	(0.198)	(0.180)	(0.142)	(0.126)
β^U	0.555**	0.686**	0.694**	0.629**	1.070**	0.856**	0.619**
	(0.107)	(0.141)	(0.151)	(0.139)	(0.232)	(0.161)	(0.108)

Notes: This table reports the estimated parameters of GPD fitted to approximated i.i.d. residuals obtained from the ARMA-GARCH filter with GED innovations. $u^L(u^U)$ denotes the threshold at the 5% (95%) quantile of the empirical distribution of the residuals. ξ and β denote the scale and shape parameter of GPD from Eq.(4.10) respectively. Subscripts (U/L) indicate the upper/lower tail. Standard errors are in parentheses. ** and * denote statistical significance at 5% and 10% confidence level, respectively.

returns.

Thus far we have devoted all modeling efforts to the first stage which is aimed at obtaining each univariate or marginal CDS index return distribution. The second-stage is aimed at characterizing the portfolio multivariate return distribution. For this purpose, we resort to copula functions whose inputs are the standardized residuals of the marginal models transformed to uniformly distributed observations via the integral transform. The estimation method is a semiparametric *cdf* transformation which can be thought of as a modification of the *Canonical Maximum Likelihood* (CML) method: the high-density (central) area of the margins is transformed via the non-parametric empirical *cdf* while the low-density (tail) areas are transformed via the parametric GPD *cdf* obtained at the first stage. Comparing with the standard CML method, our approach provides a further flexibility for capturing fat tails observed in practice. Since there is ample evidence that the degree of dependence among financial asset returns is not constant over time but instead it increases during crash periods, we implement the copulas both in a static and dynamic (DCC) framework. Four

Figure 4.5: Generalized Pareto Distribution Fit at Tails

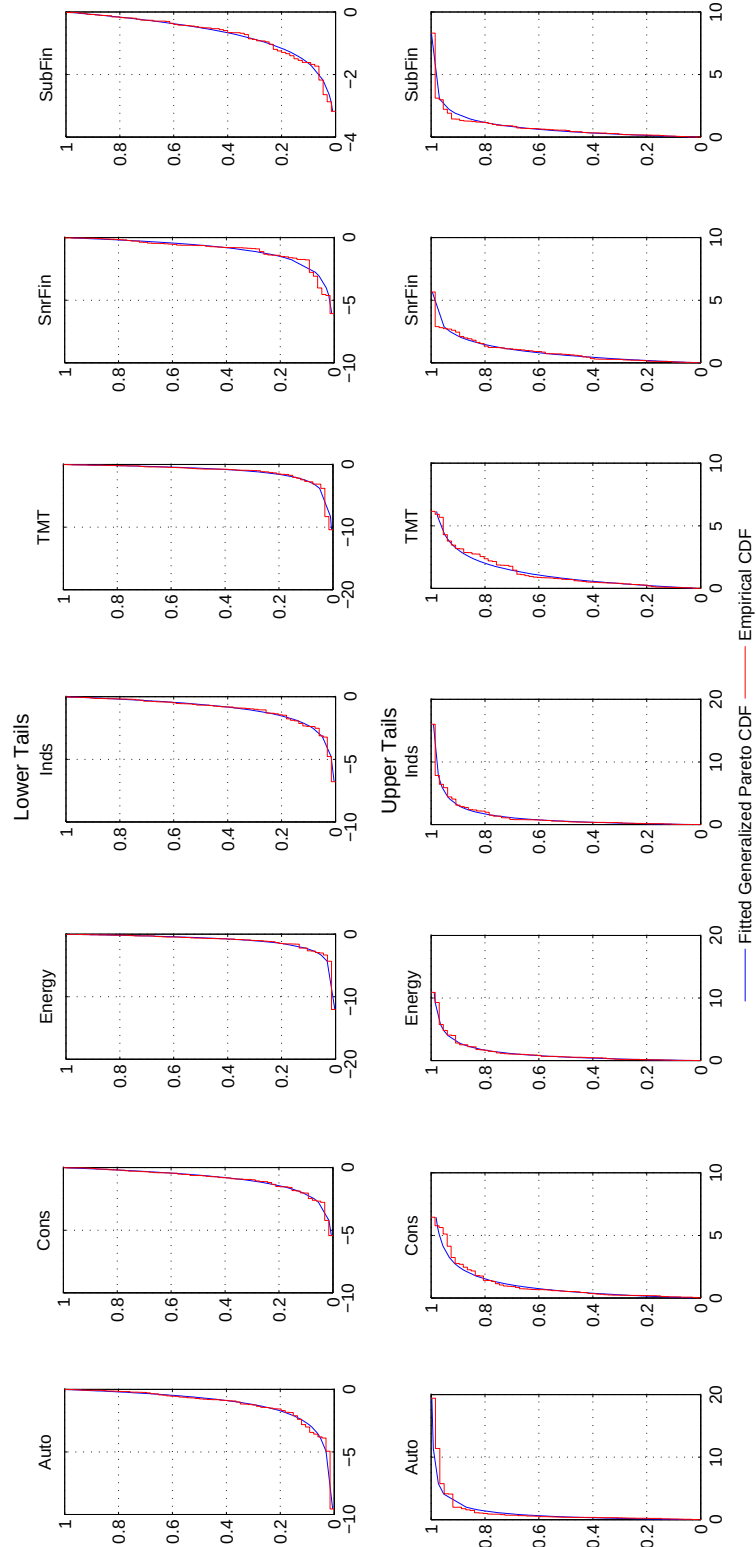


Table 4.5: Estimation Result of Copula Models

	Static- t	DCC- t	Static-Gaussian	DCC-Gaussian
ν	5.524** (0.341)	5.966** (0.369)	-	-
ψ	-	0.064** (0.009)	-	0.054** (0.008)
φ	-	0.855** (0.294)	-	0.857** (0.031)
LL	1424.921	1635.285	1066.612	1317.927
AIC	-2847.944	-3264.569	-2131.224	-2631.854

Notes: This table presents the estimated parameters of copula models, i.e. Gaussian and Student's t , with two different dependence structures, i.e. static and DCC-type dynamic copula via Eq.(4.11). Standard errors are in parentheses. AIC denotes Akaike information criterion and LL denotes the log-likelihood. ** and * denote statistical signification at 5% and 10% confidence level, respectively.

copula models are thus considered: Gaussian and Student's t in two formulations, static and dynamic. Table 4.5 sets out the results.

Gaussian copulas underperform Student's t copulas irrespective of the formulation, static or dynamic, as borne out by lower LL and higher AIC values. This can be attributed to the fact that Gaussian copulas characterize the multivariate distribution tails poorly and, in particular, they neglect tail dependence. A likelihood ratio test of the hypothesis $H_0 : 1/\nu = 0$ is strongly significant at the 1% level implying that the Student's t copula is not statistically equivalent to the Gaussian. Both LL and AIC favor the dynamic copula formulation rather than the static which is further corroborated by the statistical significance of the DCC copula parameters φ and ψ in Eq. (4.11). A likelihood ratio test rejects the hypothesis $H_0 : \varphi + \psi = 1$ implying that the pairwise rank correlations are stationary.

In sum, for the univariate marginals the best fit is provided by the ARMA-GARCH-GED-EVT model and for the multivariate distribution by DCC Student's t copula. A portfolio manager uses a model estimated with information available up to time t in order to forecast (i.e., form an expectation regarding) asset returns at time $t+1$, and then deploys either the MV strategy or MCVaR strategy to optimally allocate her money into a portfolio at the close of day t (or open of day $t+1$). At

the end of day $t+1$ she can evaluate the realized daily return of the portfolio thus formed. But what will happen if her model wrongly assumes that the distribution is Gaussian or Student's t ? Or wrongly assumes that the joint density is a Gaussian copula instead of a Student's t copula? Or the model wrongly ignores the dependence (or time-variation) in asset return dependence? Next section is aimed at addressing these questions.

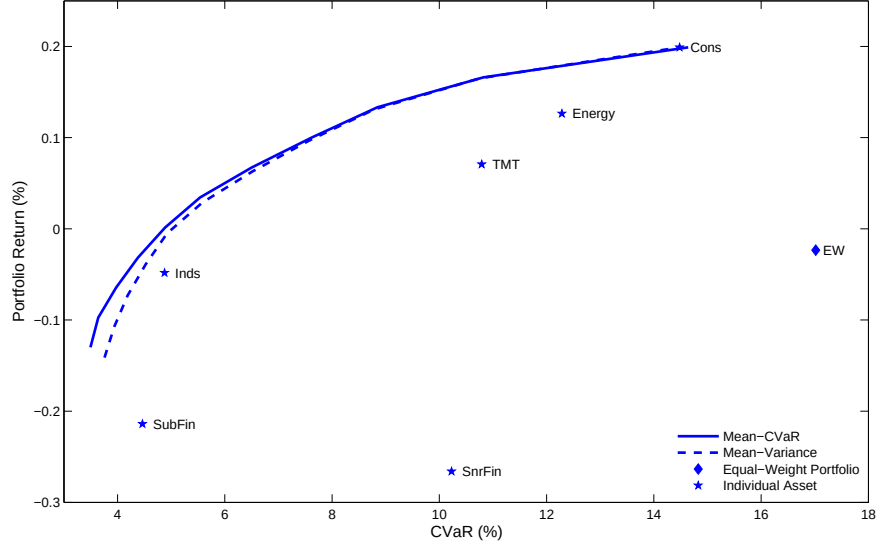
4.4.3 Performance of portfolio optimization strategies

To gain some insights into the properties of the proposed MCVaR optimization methodology compared with the traditional MV approach, we conduct a set of Monte Carlo simulation experiments. We firstly draw $N=10,000$ uniform distributed independent random numbers and introduce dependence into them via the in-sample estimated static or dynamic copulas. Next in order for these random numbers to mimic the properties of the sample observations as much as possible, we transform them by inverting the *cdf*, namely, applying the interpolation method to invert the empirical *cdf* and the inverse distribution function for the Generalized Pareto *cdf*. Finally, each series of random numbers are used to fit univariate (marginal) ARMA-GARCH models from which one-day-ahead returns are predicted. See Appendix 4.C for further details.

4.4.3.1 One-step-ahead forecasts using the entire sample

We begin by assuming that the portfolio manager employs the overall best in-sample model, namely GARCH-GED-EVT models for the margins together with a DCC- t copula for the multivariate distribution, in order to simulate the returns. The models are calibrated in-sample using the full sample of index CDS returns and the portfolio optimization approach is then deployed based on the expected or one-day-ahead predicted (ex ante) returns for March 12, 2011. For simplicity, short-selling is not allowed. The resulting curves are shown in the “mean return versus 5% CVaR” space (or mean-CVaR space for short) in Figure 4.6. One curve is the MV efficient

Figure 4.6: Optimal Portfolio Efficient Frontier at 5% Confidence Level



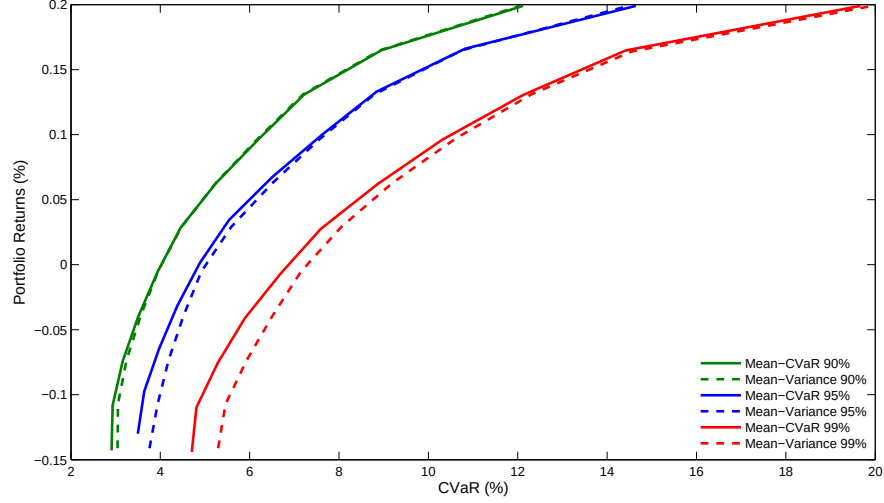
Notes: This graph shows the efficient frontiers generated by the MCVaR (solid line) and the MV (dash line) models with risk measure CVaR at 5% confidence level. The stars denote the individual index in the return-CVaR space. Auto has the lowest return and highest CVaR and is far beyond the range of the graph. The diamond denotes the equal-weight portfolio.

frontier (dash line) and the other curve is the MCVaR efficient frontier (solid line).¹⁶

The stars in Figure 4.6 denote the individual CDS index constituents in the return-CVaR space. It is clear that both the MV and the MCVaR portfolio allocation strategies make investors better off than the naive one (diamond) that invests equally in each CDS index and can only generate a negative return with high tail risk. For each return level, the MV optimized portfolio has larger CVaR than the MCVaR optimized portfolio which, in turn, implies that investors who follow the mean-variance strategy are more vulnerable to extreme tail-event losses. As we move upwards alongside the return-CVaR efficient frontiers, the MV and MCVaR optimized portfolios become closer and eventually coincide because all portfolios converge to the single-asset portfolio (Cons CDS index in our setting) with the largest return. The findings are in line

¹⁶Since the efficient frontier of MV is originally in the mean-variance space while that of MCVaR is in the mean-CVaR space, in order to make the two directly comparable we map the former onto the same mean-CVaR space. To do this, using the simulated returns at time $t+1$ we obtain the optimal weights of portfolios on the MV efficient frontier and then calculate the 5% CVaR of the thus simulated optimal portfolios.

Figure 4.7: Portfolio Optimal Frontier: Confidence Level of CVaR



Notes: This graph shows the efficient frontiers generated by the MCVaR (solid line) and the MV (dash line) models with risk measure CVaR at 10% (in green), 5% (in blue) and 1% (in red) confidence levels.

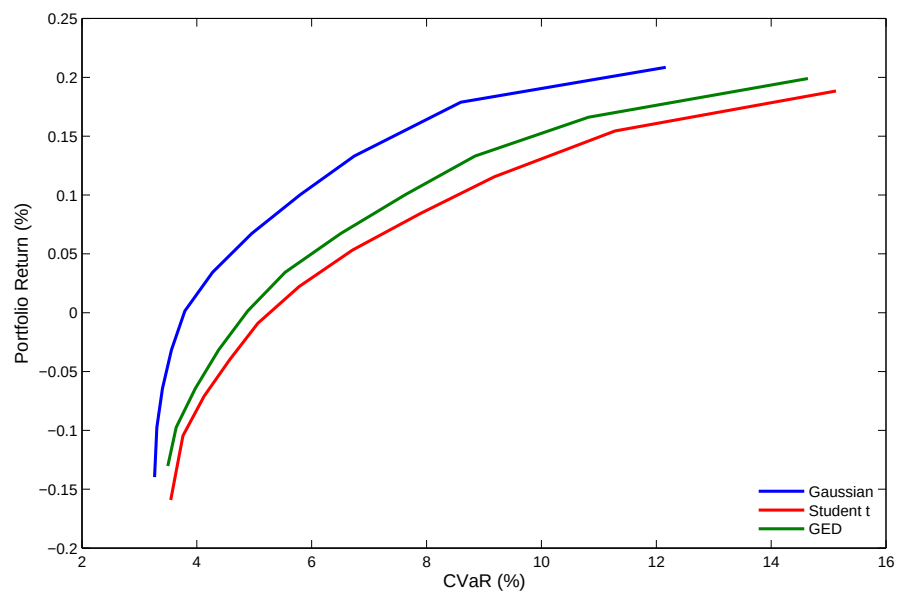
with the extant evidence suggesting that a MV-optimal portfolio is “near optimal” in a CVaR sense; see Krokmal et al. (2002).

We then explore the impact on portfolio allocation of choosing different confidence levels $\alpha \in \{0.99, 0.95, 0.90\}$ for the tail risk measure CVaR. The portfolio manager switches from $\alpha = 0.90$ to $\alpha = 0.99$ if she is concerned with more extreme losses. The results are shown in Figure 4.7. The graph confirms that the difference in the MV and MCVaR efficient frontiers magnifies as α increases.

Next we investigate the impact of making different distributional assumptions regarding the margins. In particular, we presume that the portfolio manager resorts to Gaussian, Student’s t or GED assumptions in the in-sample modeling exercise to construct expected (or forecasted) returns at $t+1$. The results reported in Figure 4.8 reveal clear differences in the three MCVaR efficient frontiers.

For a given return level, the 95% CVaRs of optimized portfolios lie in ascending order for the Gaussian, GED and Student’s t marginal models. Since GED provides the best data fit, as shown in Section 4.4.2, the analysis suggests that investors who optimize their portfolios wrongly assuming a Gaussian distribution may suffer large

Figure 4.8: Optimal Portfolio Efficient Frontier at 5% Confidence Level: Distribution Comparison



Notes: This graph compares the efficient frontiers generated by the MCVaR models with different margin distributions, e.g. Gaussian (blue line), Student's t (red line), and GED (green line), with risk measure CVaR at 5% confidence level.

tail-event losses. This is because the Gaussian distribution portrays short tails and hence, underpredicts the true downside tail risk. In contrast, the Student's t distribution (red line) has polynomial tails and it is heavier-tailed than GED. Therefore, when the true distribution is GED, the optimized portfolio under the assumption of Student's t margins will instead overestimate the 95% CVaR, as Figure 4.8 illustrates.

4.4.3.2 Performance of realized portfolios with optimal weights

We then extend the simulation analysis conducted in Section 4.4.3.1 by iterating the simulations over rolling windows of one year (250 days) in order to gauge the potential of the two portfolio construction strategies for *real* risk management. The methodology is similar to that employed in Gaivoronski and Pflug (2005), but instead of forecasted one-day-ahead return we incorporate realized returns in our exercise.

Specifically, over each rolling window of fixed length (1,118 days), the portfolio manager calibrates her model in-sample, e.g. the first window spans the period from September 20, 2005 to March 11, 2010 and the ARMA-GARCH models with Gaussian or GED innovations and EVT for univariate margins and the static/dynamic DCC Gaussian or Student's t copula models for joint density¹⁷ (totally eight combinations) are estimated. She then uses it to simulate one-day-ahead returns, e.g. the first out-of-sample day on March 12, 2010, following the algorithm described in Appendix 4.C. The exercise is repeated $N=10,000$ times, i.e. generating N scenarios, and those forecasted returns are used to generate the MV and MCVaR efficient frontiers. It reflects the portfolio manager's judgments on the optimal portfolio allocation at the end of each window. The optimal weights at the minimum risk level, i.e. those in the left end of the frontier, associated with the observed returns of the sectoral CDS index constituents are then used to calculate the portfolio's realized returns. As a result, she collects 250 days of portfolio's realized returns under different investment strategies from March 12, 2010 to March 11, 2011.

¹⁷In order to avoid too many modeling approaches which makes the reporting of results cumbersome, we constrain the comparison of distributional assumptions for the marginal models to the Gaussian (in the spirit of the mean-variance framework) and the GED that provides the best data fit.

The statistics reported in Table 4.6 include the mean, standard deviation, α -VaR and α -CVaR, for $\alpha = 0.95$, of the realized optimal portfolios under the MV or MCVaR strategies over the out-of-sample period. It is apparent that at the minimum risk level the mean returns of MCVaR portfolios are slightly higher than those of MV portfolios. Similarly, the lower standard deviation and 95% CVaR suggest that the MCVaR portfolios are relatively less risky. Furthermore, the superior return-risk tradeoff associated with the optimal MCVaR portfolios is robust regardless of the choice of in-sample model, i.e. in seven out of eight in-sample models we can observe higher return-risk ratios of the MCVaR portfolios than those associated with the MV portfolios.¹⁸ For instance, for portfolios modeled assuming GED margins and dynamic Student's t copula dependence, we obtain return-risk ratios of 2.395% (mean/standard deviation) and 1.291% (mean/CVaR) for the MV optimal portfolio, whereas the corresponding ratios for MCVaR portfolios are 2.574% and 1.396% respectively. We find that the in-sample model does not seem to play an important role here as for a given model specification the MCVaR strategy can generate more attractive returns than the MV approach does at the same risk level. Since the MCVaR strategy aims to avoid extreme large losses, it in turn reduces the volatility of portfolio returns and produces a better return-risk tradeoff as expected.

Regarding the impact of the assumptions made in simulating the one-day-ahead returns, we find that the resulting optimized realized portfolios using GED margins have higher return-risk ratios than when Gaussian margins are assumed. Since GED has heavier tails than the Gaussian, this implies that the MCVaR strategy becomes more efficient in asset allocation when the heavier tails are introduced. The copula dependence structure assumed is another key aspect. When the portfolio multivariate distribution is characterized by a Student's t copula the resulting return-(C)VaR frontier is shifted to the left relative to that obtained using a Gaussian copula, which

¹⁸There is one exception, the portfolio with GED margins and DCC Gaussian copula for which the return-risk ratio is lower for standard deviation and VaR but higher for CVaR. But the later metric is gaining prominence among bank risk managers who are seeking to avoid catastrophic losses because it is more sensitive than VaR (and standard deviation) to the shape of the tail distribution in a bell curve of potential losses. Reassuringly, the return-CVaR ratio confirms a better performance of the MCVaR model than the MV model throughout.

Table 4.6: Statistics of Realized Optimal Portfolio Returns

	MV	$\frac{return}{risk}$	MCVaR	$\frac{return}{risk}$	MV	$\frac{return}{risk}$	MCVaR	$\frac{return}{risk}$
	GED Margins				Gaussian Margins			
<i>Static Gaussian Copula</i>								
Mean	0.125	-	0.131	-	0.079	-	0.104	-
Std Dev	8.469	1.474	8.404	1.557	8.071	0.977	7.724	1.351
$VaR_{5\%}$	8.072	1.546	8.106	1.614	7.393	1.067	7.505	1.391
$CVaR_{5\%}$	14.334	0.871	14.027	0.933	13.157	0.600	12.903	0.809
<i>Static t Copula</i>								
Mean	0.136	-	0.137	-	0.108	-	0.120	-
Std Dev	8.440	1.617	8.405	1.632	8.456	1.277	8.144	1.473
$VaR_{5\%}$	8.956	1.524	8.059	1.702	7.715	1.400	7.634	1.572
$CVaR_{5\%}$	14.630	0.933	14.597	0.940	13.336	0.810	12.941	0.927
<i>DCC Gaussian Copula</i>								
Mean	0.141	-	0.140	-	0.104	-	0.125	-
Std Dev	6.891	2.044	6.861	2.041	6.759	1.532	6.705	1.863
$VaR_{5\%}$	8.343	1.689	8.604	1.627	7.967	1.300	7.897	1.582
$CVaR_{5\%}$	13.240	1.064	12.866	1.088	12.369	0.837	12.319	1.014
<i>DCC t Copula</i>								
Mean	0.176	-	0.188	-	0.120	-	0.135	-
Std Dev	7.350	2.395	7.299	2.574	7.229	1.662	7.219	1.865
$VaR_{5\%}$	8.327	2.114	8.564	2.193	7.524	1.597	7.371	1.827
$CVaR_{5\%}$	13.636	1.291	13.453	1.396	13.028	0.922	12.959	1.039

Notes: This table presents the statistics of realized optimal portfolio based on the MV and MCVaR strategies with risk measure of CVaR at 5% confidence level over 12 March, 2010 to 11 March, 2011. All realized portfolio at $t+1$ are constructed based on the optimal asset weights of the minimum risk MV and MCVaR portfolios with simulated returns ($N = 10,000$) at $t+1$ where $t+1$ denotes the day point over 12 March, 2010 to 11 March, 2011. The simulations are based on the models calibrated in-sample with 250 rolling windows. Two innovation distribution, i.e. Gaussian and GED, and two copulas, i.e. Gaussian and Student's t , in static and dynamic DCC settings are considered for the in-sample models. Mean, Std Dev, $VaR_{5\%}$ and $CVaR_{5\%}$ denote the expected mean, standard deviation, VaR and CVaR at 5% level of realized portfolios. All figures denote percentage.

stems from the fact that the latter ignores the tail dependence. Moreover, adopting dynamic copulas as opposed to static copulas enhances the mean realized returns and return-risk ratios of the optimized portfolios. For example, the mean return (return-CVaR ratio) of optimized MCVaR portfolios via DCC t copulas is 0.188% (1.396%) compared with 0.131% (0.940%) using static t copulas. Comparing with static/constant parameter copulas, a crucial advantage of dynamic copulas is that their conditional (up to current time t information) specification makes them more able to reflect changes in dependence structure among assets so that the portfolio manager can quickly adjust her assets allocation to avoid large risks. To sum up, an accurate description of the portfolio's multivariate return distribution which, in turn, hinges on a good characterization of the marginal distributions has an impact on portfolio optimization.

4.5 Conclusion

Even though there is ample consensus that asset returns deviate from the Gaussian assumption and have time-varying and nonlinear interdependence features such as tail dependence, many financial applications still rely on simple models that overlook these stylized facts and, as a result, underestimate the probability of catastrophic losses. The focus of this study is portfolio optimization. We advocate the combination of a heavy-tailed ARMA-GARCH filter and an EVT peaks-over-threshold approach to model the univariate marginal distributions, and a dynamic Student's t copula model to obtain a flexible multivariate portfolio return distribution that accommodates time-varying tail dependence. Portfolios thus characterized are simulated to investigate the performance of two allocation strategies, the classical mean-variance strategy that represents the cornerstone of modern portfolio theory and the mean-CVaR strategy that relies on conditional Value-at-Risk (or expected shortfall) as relevant risk measure. The analysis is based on both out-of-sample forecasted returns and realized returns.

The findings suggest that mean-CVaR optimized portfolios offer a more attrac-

tive return-risk tradeoff than mean-variance optimized portfolios. The discrepancy between the two optimization strategies increases with the confidence level associated to the CVaR measure. We show that portfolio management is influenced by the asset return distributions (margins) and the dependence structure (multivariate distribution) assumed in the one-day-ahead returns simulations. The superiority of the mean-CVaR optimized portfolio relative to the mean-variance optimized portfolio is more strongly revealed when we introduce in the Monte Carlo simulation of one-day-ahead returns the heavy tailedness of asset return distributions and the dynamic tail dependence in portfolio returns. An important implication arising out of these conclusions is that risk managers are more capable of averting catastrophic losses while maintaining the same desired return level if they rely on the mean-CVaR portfolio allocation approach instead of the classical mean-variance strategy.

Appendix

4.A Coherent Risk Measures

Artzner et al. (1999) advocated that good risk measures need to have below list of properties.

Theorem (Risk Measure). *Consider a set of random variables X . A function $g : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is called a risk measure if it satisfies:*

- (i) *Monotonous: if $A \leq B$, $A, B \in X$ then $g(A) \leq g(B)$;*
- (ii) *Subadditive: if $A, B, A + B \in X$ then $g(A + B) \leq g(A) + g(B)$;*
- (iii) *Positively homogeneous: if $A \in X, \alpha > 0$ then $g(\alpha A) = \alpha g(A)$;*
- (iv) *Translation invariant: if $A \in X, \alpha \in \mathbb{R}$ then $g(A + \alpha) = g(A) - \alpha$.*

VaR satisfies all but subadditivity. In other words it might be possible to have a portfolio with higher VaR than the sum of the individual VaR of portfolio components.

4.B Multivariate Elliptical Copula

Gaussian copula A n -dimensional Gaussian copula is equivalent to the multivariate Gaussian joint function with *cdf*:

$$C(u_1, u_2, \dots, u_n; R) = \Phi_R(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n))$$

where $\Phi^{-1}(u)$ denotes the inverse of the Gaussian *cdf* and R is the correlation matrix. The *pdf* of this is

$$\begin{aligned} c(\Phi(x_1), \Phi(x_2), \dots, \Phi(x_n); R) &= \frac{f^{Ga}(x_1, x_2, \dots, x_n)}{\prod_{i=1}^n f_i^{Ga}(x_i)} \\ &= \frac{\frac{1}{(2\pi)^{\frac{n}{2}} |R|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} X' (R^{-1} - I) X\right\}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} x_i^2\right\}} \end{aligned}$$

where f^{Ga} denotes the *pdf* of multivariate Gaussian and f_i^{Ga} is the marginal Gaussian density. If we make $u = \Phi(x)$ and $\Psi = (\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n))'$ then the

above equation can be written as

$$c(u_1, u_2, \dots, u_n; R) = \frac{1}{|R|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \Psi' (R^{-1} - I) \Psi \right\}.$$

Student-*t* copula A n -dimensional Student's t copula C is a n -dimensional distribution function on $[0, 1]^n$ with standard uniform marginal distributions and cdf as follows

$$\begin{aligned} C(u_1, u_2, \dots, u_n; \nu, R) &= T_{R, \nu}(t_{\nu}^{-1}(u_1), t_{\nu}^{-1}(u_2), \dots, t_{\nu}^{-1}(u_n)) \\ &= \int_{-\infty}^{t_{\nu}^{-1}(u_1)} \cdots \int_{-\infty}^{t_{\nu}^{-1}(u_n)} \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2}) \sqrt{(\pi\nu)^n \det(R)}} \left(1 + \frac{X'R^{-1}X}{\nu}\right)^{-\frac{\nu+n}{2}} dx \end{aligned}$$

where $t_{\nu}^{-1}(u)$ denotes the inverse of the Student's t cdf , ν is the degree of freedom and R is the correlation matrix. The corresponding pdf can be derived as

$$\begin{aligned} c(u_1, u_2, \dots, u_n; \nu, R) &= \frac{f^t(x_1, x_2, \dots, x_n)}{\prod_{i=1}^n f_i^t(x_i)} \\ &= |R|^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2})} \left[\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})} \right]^n \frac{(1 + \frac{1}{\nu} \Psi' R^{-1} \Psi)^{-\frac{\nu+n}{2}}}{\prod_{i=1}^2 (1 + \frac{1}{\nu} \Psi_i^2)^{-\frac{\nu+1}{2}}} \end{aligned}$$

where $\Psi = (t_{\nu_1}^{-1}(u_1), t_{\nu_2}^{-1}(u_2), \dots, t_{\nu_n}^{-1}(u_n))'$.

4.C Monte Carlo simulation for 1-day-ahead returns

The empirical distribution of 1-day-ahead returns is obtained as follows:

1. Find the Cholesky decomposition $\mathbf{A}_{K \times K}$ of the forecasted rank correlation matrix $\hat{R}_{t+1}$ with 1s in the diagonal and with pairwise 1-day-ahead Kendall's rank correlations as off-diagonal elements from a copula model estimated using in-sample data for the K -asset portfolio.
2. Simulate K independent standard normal random variates $\mathbf{z} = (z_1, \dots, z_K)'$.
3. Form the vector $\mathbf{b} = \mathbf{A}\mathbf{z}$ where $\mathbf{b} = (b_1, \dots, b_K)'$ are dependent random variables

with dependence introduced through the matrix $\mathbf{A}$ obtained in Step 1.

4. Determine the components $(u_1, \dots, u_K) = (\Phi(b_1), \dots, \Phi(b_K))$ of Gaussian copula where Φ is the normal *cdf*. The components of Student's *t* copula are obtained as $(u_1, \dots, u_K) = (t_{\hat{\nu}}(c_1), \dots, t_{\hat{\nu}}(c_K))$ where $t_{\hat{\nu}}$ is the Student's *t* *cdf* with $\hat{\nu}$ degrees of freedom estimated in sample and $(c_1, \dots, c_K)' = \frac{\sqrt{\hat{\nu}}}{\sqrt{s}} \mathbf{b}$ where s is a random variable simulated from a $\chi_{\hat{\nu}}^2$ distribution and s is independent of $\mathbf{z}$.
5. Obtain the standardized asset returns $(q_1, \dots, q_K) = (\hat{F}_1^{-1}(u_1), \dots, \hat{F}_K^{-1}(u_K))$, where $\hat{F}_k^{-1}$ is either the inverse empirical *cdf* of standardized residuals for the high-density (or central) area or the inverse GPD *cdf* of the in-sample data x_k , $k \in \{1, \dots, K\}$, for the tails.
6. Rescale and relocate the standardized returns by using the ARMA-GARCH forecasts of conditional mean or location $\hat{\mu}_{t+1}$ and conditional variance or scale $\hat{\sigma}_{t+1}$ based on the in-sample data as $(r_{1,t+1}, \dots, r_{K,t+1}) = (\hat{\mu}_{1,t+1} + q_1 \sqrt{\hat{\sigma}_{1,t+1}}, \dots, \hat{\mu}_{K,t+1} + q_K \sqrt{\hat{\sigma}_{K,t+1}})$.
7. Repeat $N = 10,000$ times the above steps 1 to 6 to obtain the empirical (simulated) distribution of one-day-ahead returns $\{r_{1n,t+1}, \dots, r_{Kn,t+1}\}_{n=1}^N$ for each of the K assets.

Conclusions and Suggestions for Future Research

This chapter summarizes the entire thesis and offers suggestions for future research. The main subject of this thesis was to provide further insight into some important topics of modern risk management practice by deploying advanced quantitative techniques.

5.1 Concluding Remarks

As one of the code activities of banks, Basel Accords – Basel I (1988), Basel II (2004) and the looming Basel III (2011) – have changed and are still changing the way that banks address the management of risk. Under the Basel Accord frameworks banks are required to hold adequate risk-sensitive minimum capital in order to safeguard its solvency and overall economic stability. As a response to the recent global financial catastrophes, the crisis-driven Basel III aims to make the banking sector more resilient to stress market conditions. This thesis grouped with three empirical studies provides the economic relevance to the Basel III's macroprudential goal from three directions: in Chapter 2, we proposed a MMC model of credit rating migration to promote countercyclical capital buffers in order to dampen procyclicality; in Chapter

3, a flexible regime switching copula model is developed for more accurately characterizing portfolio dependence. The regime-switching dynamic dependence enable more realistic stress testing scenarios with important implications for the determination of bank capital levels; in Chapter 4, we advocate a novel EVT-dynamic copula MCVaR approach to reduce portfolio's large downside risk for improving its efficiency and resilience to extreme market environment. The detailed findings of each chapter are set out below.

5.1.1 Business-cycle adjusted credit migration for calculating capital requirement

In Chapter 2 we explored the concept of credit rating migration which plays a key role in calculating the risk of banks' loan books as well as determining banks' risk capital allocation under the Basel framework. We firstly briefed the two classical estimation approaches, the discrete-time cohort method and its continuous-time extension, the hazard rate model, that both assume rating migrations are time-homogeneous and follow a Markov chain process. We then loosened the strong assumption by allowing time-heterogeneity so that our advocated MMC model of credit rating migration explicitly recognized the stochastic evolution of business cycles. We examined the performance of the MMC estimator against the naive cyclical counterpart and the classical through-the-cycle estimators in three different frameworks: the purely statistical framework with emphasis on in-sample estimation accuracy; the forward-looking framework evaluating forecast ability through loss functions; and the economic implication of capital attribution.

The analysis was based on 26-years of S&P marketwide and sectoral rating data and concluded that the MMC approach yields more reliable default risk measures than the naive cyclical estimator does, especially in economy contraction. The forecast accuracy gains of the MMC estimator over its naive counterpart become more prominent as the time horizon lengthens. The economic application revealed that MMC and naive cyclical approaches both suggest a higher level of buffer than those

from classical through-the-cycle estimators to against risks in economic contraction, but it was statistically and economically overestimated by the naive estimator compared with the MMC approach. We also found banks using the MMC estimator would counter-cyclically increase capital by 6% during economic expansion and free up to 17% capital for lending during downturns compared to the naive estimator. Thus, the MMC estimator is well aligned with the Basel III macroprudential initiative to dampen procyclicality by reducing the recession-versus-expansion gap in capital buffers.

5.1.2 Regime switching copula for characterizing CDS-equity dependence

Chapter 3 investigated the key concern of risk management – the dependence – in the context of dynamic linkages between credit risk swap index and the corresponding equity market. The accurate measurement of the dependence between the two markets is of importance to risk managers for setting trading limits, or traders for hedging the market risk of their credit portfolio positions, or policymakers to set capital rules in stressed market conditions. It is also closely relevant to the Basel II Accord and new Basel III Accord for better calculation of risk capital requirements of credit portfolios.

In this chapter we proposed a time-varying regime switching copula model, whose parameters are allowed not just to conditionally depend on historical return comovements but also to vary between regimes, in order to investigate the importance of various dynamic patterns in the dependencies between the iTraxx CDS market and the underlying stock market return and volatility. The regime-switching copula model was compared with purely dynamic and static copula models. Using daily data for both marketwide and sectoral indices from the two markets, we found the CDS market is negatively correlated with stock return and positively correlated with stock return volatility. We also documented asymmetric behavior, namely, a regime of high dependence during "crisis" periods characterized by extreme adverse comovements in the two markets alternates with a regime of low dependence during more "normal"

periods. The high dependence regime coincides with the credit crunch and the European sovereign debt crisis. In-sample statistical analysis reveals the relevance of the regimes at the center and at the tails of the joint distribution. Ignoring such effects is shown to lead to underestimation of dependence in periods of crisis.

An economic evaluation framework was also deployed to examine the performance of the aforementioned copula models by determining whether they could produce accurate VaR estimates via Monte Carlo simulations. The superiority of the regime-switching approach over purely dynamic or static copula is underlined via out-of-sample VaR backtesting relevant for risk management. Our results have important implications for banks' in-house calculations of capital requirements and point into a clear direction for improvement of banks' stress testing platforms.

5.1.3 Downside extreme risk for portfolio optimization

Chapter 4 studied the impact of the choice of risk measure on optimal portfolio selection with a focus on downside extreme risks. The lessons from the recent financial crisis indicate that some undesirable extreme events, which are naturally rare, at the far-end tails will largely contribute to the potential losses, therefore a good risk management framework should control for such events.

In this chapter we advocated a hybrid method, combining a heavy-tailed GARCH filter with an EVT approach for providing more realistic descriptions of stylized facts such as leptokurtosis, asymmetry, auto-correlation and heteroskedasticity of univariate asset returns. A dynamic t copula model was then employed to glue these assets in a portfolio setting and to provide timely explorations of the correlation evolution especially the dependence at tails. We optimized risk portfolio with CVaR, which only concerns the downside losses, as the alternative risk measure versus the conventional standard deviation measure of the old-school Markowitz's mean-variance framework. To investigate the performance of the two portfolio allocation approaches, i.e. MV and MCVaR, a set of Monte Carlo out-of-sample simulation experiments was conducted.

The results suggested the traditional MV approach can only produce suboptimal protection against extreme losses thus it should not be solely used, especially when

prevention of large downside risk is a central concern. The performance of portfolio optimization is jointly determined by the assumption of the asset return distribution, the dependent structure and the choice of risk measures. It implies that a more realistic model with the ability to give better description on the asset components and their comovements can always benefit making proper decision of portfolio selection. Finally the horse-racing experiment of risk-return allocation on realized portfolios provided evidence that MCVaR models incorporating heavy tails in asset level and tail dependence in portfolio level can significantly increase the performance of portfolio optimization. We concluded that models taking care of the non-normality that we observed in the real-world may reduce the portfolio's large downside risk, improve its efficiency and resilience to extreme market environment, therefore it has meaningful empirical implications for risk hedging and portfolio investments.

5.2 Further Research

In the final section of this thesis, we list below a couple of possible directions for future extensions.

One of the potential limitations of the MMC method developed in Chapter 2 is that we implicitly assumed the regime-depended rating migration process is immediately triggered upon the change of business cycle. But these two Markov process might not work simultaneously, namely there might be some lag reaction of one process relative to the other. Thus, a reasonable extension to cope with this limitation would be to identify such lag effects before employing the MMC model. Another interesting extension would be some attempts to add non-Markovian features, such as rating momentum, on the current time heterogeneity content. But one can expect that the problem setting would definitely become more complicated.

One of the potential weaknesses of the models proposed in the Chapter 3 is the highly computational complexity. Since the models involve a large number of parameters, it takes average 2.5 minutes for calibration which is about fifteen (two hundreds)

times longer than that for dynamic (static) copulas.¹ But this speed could be notably shortened by either rewriting our MATLAB code in C/C++, or rewriting some of the functions in C/C++ and calling them from MATLAB. Secondly, as we have mentioned in the Chapter 3, one of the challenges the dynamic copula models face is how to extend them to a high-dimensional setting. It is straightforward for DCC-type copulas, but is not clear for ARMA-type copulas. Another limitation of the methods employed here is that we implicitly assume the dependence structure is identical for all pairs within the portfolio and they enter the high dependence regime simultaneously. For instance, the pair dependencies of A-B, B-C and C-A in a three asset (A, B and C) portfolio are assumed to be described by the same copula function. In real life we might need a more flexible model to allow for different dependence patterns. For example, the pair of A-B can be characterized by a Student's t copula, whereas a Clayton copula is for the pair of B-C and each of them has individual parameters to control the timing of entering the high dependence regime. It is the idea of the so-called regime-switching Vine copula proposed very recently, but still limited in the static context. Hence it will be interesting to introduce dynamics somehow to the Vine copulas.

Finally, the high computational cost once again becomes one of disadvantages of the methods proposed in Chapter 4. The computational complexity increases rapidly when the new risk measure, i.e. CVaR, is employed or when the size of samples and portfolio components becomes large. The computational burden may be eased by deploying the smoothing algorithms developed by Alexander et al. (2006) and Zhu et al. (2009) that their approaches are shown to be computationally significantly more efficient than the linear programming method for the CVaR optimization problem. Secondly, the Chapter can be extended to allow investors to have multiple holding periods and have no short-selling restrictions. The extended model should be able to answer how the portfolio allocation strategy should be constructed to reflect investor's predication on the market's long-term or the short-term behavior. Another

¹The results are based on our programs running with MATLAB 2010b on an Intel Core i3-380M laptop. The optimization function *fmincon* took 83% of the overall computational time in order to find a global minimum value.

possible avenue of future work inspired by extreme value theory may be to model the dependence of extreme events separately from that at the center of the joint distribution. Very recently a small number of pioneering studies has started working on this area by using the method of extreme value copula. This field is still at the very early stages of development and, of course, only confined to a static setting. Hence, it will be interesting to extend this idea into a dynamic framework and delve into whether such an approach would benefit risk management.

References

- Aas, K., Haff, I., 2006. The Generalized Hyperbolic Skew Student's t Distribution. *Journal of Financial Econometrics* 4 (2), 275–309.
- Alexander, C., Kaeck, A., 2008. Regime Dependent Determinants of Credit Default Swap Spreads. *Journal of Banking & Finance* 31 (6), 1008–1021.
- Alexander, S., Coleman, T. F., Li, Y., 2006. Minimizing CVaR and VaR for a Portfolio of Derivatives. *Journal of Banking & Finance* 30 (2), 583–605.
- Altman, E. I., Kao, D. L., 1992. Rating Drift of High Yield Bonds. *The Journal of Fixed Income* 1 (4), 15–20.
- Andersson, F., Mausser, H., Rosen, D., Uryasev, S., 2001. Credit Risk Optimization with Conditional Value-at-Risk Criterion. *Mathematical Programming* 89 (2), 273–291.
- Ang, A., Chen, J., 2002. Asymmetric Correlations of Equity Portfolios. *Journal of Financial Economics* 63 (3), 443–494.
- Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent Measures of Risk. *Mathematical Finance* 9 (3), 203–228.
- Artzner, P., Delbaenn, F., Eber, J. M., Heath, D., 1997. Thinking Coherently. *Risk* 10 (11), 68–71.

- Bai, M., Sun, L., 2007. Application of Copula and Copula-CVaR in the Multivariate Portfolio Optimization. In: *Combinatorics, Algorithms, Probabilistic and Experimental Methodologies*. Springer, pp. 231–242.
- Baker, K., Mansi, S. A., 2002. Assessing Credit Rating Agencies by Bond Issuers and Institutional Investors. *Journal of Business Finance & Accounting* 29 (9-10), 1367–1398.
- Balkema, A., de Haan, L., 1974. Residual Life Time at Great Age. *Annals of Probability* 2 (5), 792–804.
- Bangia, A., Diebold, F. X., Schuermann, T., Co, O. W., 2002. Ratings Migration and the Business Cycle, With Application to Credit Portfolio Stress Testing. *Journal of Banking & Finance* 26 (2-3), 445–474.
- Basel Committee on Banking Supervision, 1988. International Convergence of Capital Measurement and Capital Standards. Basel Committee on Banking Supervision, BIS, Basel, Switzerland.
- Basel Committee on Banking Supervision, 2004. International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Basel Committee on Banking Supervision, BIS, Basel, Switzerland.
- Basel Committee on Banking Supervision, 2011a. Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems. Basel Committee on Banking Supervision, BIS, Basel, Switzerland.
- Basel Committee on Banking Supervision, 2011b. Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems. Bank for International Settlements.
- Basel Committee on Banking Supervision, May 2012. Fundamental Review of the Trading Book. Consultative document, Basel Committee on Banking Supervision.

- Batta, G., 2011. The Direct Relevance of Accounting Information for Credit Default Swap Pricing. *Journal of Business Finance & Accounting* 38 (9-10), 1096–1122.
- Baumol, W. J., 1963. An Expected Gain-Confidence Limit Criterion for Portfolio Selection. *Management Science* 10 (1), 174–182.
- Berge, T. J., Jordà, O., 2009. The Classification of Economic Activity. Working Paper - University of California, Davis.
URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1500024
- Billio, M., Caporin, M., 2005. Multivariate Markov Switching Dynamic Conditional Correlation GARCH Representations for Contagion Analysis. *Statistical Methods and Applications* 14 (2), 145–161.
- Black, F., Cox, J. C., 1976. Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *Journal of Finance* 31, 351–367.
- Black, F., Scholes, M., 1973. The Pricing of OOption and Corporate Liabilities. *Journal of Political Economy* 81, 637–654.
- Blanco, C., Ihle, G., 1999. How Good is Your VaR? Using Backtesting to Assess System Performance. *Financial Engineering News*, August, 1–2.
- Blanco, R., Brennan, S., Marsh, I. W., 2005. An Empirical Analysis of the Dynamic Relation between Investment-grade Bonds and Credit Default Swaps. *Journal of Finance* 60 (5), 2255–2281.
- Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31 (3), 307–327.
- Boubaker, H., Sghaier, N., 2012. Portfolio optimization in the presence of dependent financial returns with long memory: A copula based approach. *Journal of Banking & Finance*.
- Brailsford, T. J., Faff, R. W., 1996. An Evaluation of Volatility Forecasting Techniques. *Journal of Banking & Finance* 20 (3), 419–438.

- Brave, S., Butters, A., 2010. CFNAI Now Turns Ten - Analyzing Its First Decade of Performance. *Essays on Issues* (Federal Reserve Bank of Chicago) 10 (173).
- Breymann, W., Dias, A., Embrechts, P., 2003. Dependence Structures for Multivariate High-Frequency Data in Finance. *Quantitative Finance* 3 (1), 1–14.
- Briys, E., De Varenne, F., 1997. Valuing Risky Fixed Rate Debt: An Extension. *Journal of Financial and Quantitative Analysis* 32 (2).
- Brooks, C., 1996. Testing for Non-linearity in Daily Sterling Exchange Rates. *Applied Financial Economics* 6 (4), 307–317.
- Bystrom, H., 2006. CreditGrades and the iTraxx CDS Index Market. *Financial Analysts Journal* 62 (6), 65–76.
- Bystrom, H., 2008. Credit Default Swaps and Equity Prices: The iTraxx CDS Index Market. In: Wagner, N. (Ed.), *Credit Risk - Models, Derivatives, and Management*. Vol. 6. Chapman & Hall.
- Cantor, R., ap Gwilym, O., Thomas, S., 2007. The Use of Credit Ratings in Investment Management in the US and Europe. *Journal of Fixed Income* 17 (2), 13–26.
- Cao, C., Yu, F., Zhong, Z., 2010. The Information Content of Option-implied Volatility for Credit Default Swap Valuation. *J Financ Market* 13 (3), 321–343.
- Carling, K., Jacobson, T., Lindé, J., Roszbach, K., 2002. Capital Charges under Basel II: Corporate Credit Risk Modelling and the Macroeconomy. *Sveriges Riskbank Working Paper*.
URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=488542
- Carty, L. V., Fons, J. S., 1994. Measuring Changes in Corporate Credit Quality. *Journal of Fixed Income* 4 (1), 27–41.
- Chan, K., Gray, P., 2006. Using Extreme Value Theory to Measure Value-at-Risk for Daily Electricity Spot Prices. *International Journal of Forecasting* 22 (2), 283–300.

- Charette, R. N., 1990. Application Strategies for Risk Analysis. McGraw Hill.
- Cherubini, U., Luciano, E., Vecchiato, W., 2004. Copula Methods in Finance. John Wiley and Sons.
- Chollete, L., Heinen, A., Valdesogo, A., 2009. Modeling International Financial Returns with a Multivariate Regime-switching Copula. *Journal of Financial Econometrics* 7 (4), 437–480.
- Christensen, J. H., Hansen, E., Lando, D., 2004. Confidence Sets for Continuous-Time Rating Transition Probabilities. *Journal of Banking & Finance* 28 (11), 2575–2602.
- Christoffersen, P. F., 1998. Evaluating Interval Forecasts. *International Economic Review* 39 (4), 841–862.
- Christoffersen, P. F., Errunza, V., Jacobs, K., Langlois, H., 2012. Is the potential for international diversification disappearing? a dynamic copula approach. *Review of Financial Studies* 25, 3711–3751.
- Couderc, F., Renault, O., 2005. Times-to-Default: Life Cycle, Global and Industry Cycle Impacts. FAME Research Paper.
URL <http://ideas.repec.org/p/fam/rpseri/rp142.html>
- Credit Suisse First Boston, 1997. CreditRisk+: A Credit Risk Management Framework. Credit Suisse First Boston, Switzerland.
- Crook, J., F. M., 2011. Checking for Asymmetric Default Dependence in a Credit Card Portfolio: A Copula Approach. *Journal of Empirical Finance* 18, 728–742.
- Das, S., Geng, G., 2006. The Credit Market Handbook. Advanced Modeling Issues. John Wiley & Sons, Inc.
- Demarta, S., McNeil, A. J., 2005. The t Copula and Related Copulas. *International Statistical Review* 73 (1), 111–129.
- Di Clemente, A., Romano, C., 2004. Measuring and Optimizing Portfolio Credit Risk: A Copula-based Approach. *Economic Notes* 33 (3), 325–357.

- Diebold, F. X., Gunther, T. A., Tay, A. S., 1998. Evaluating Density Forecasts with Applications to Financial Risk Management. *International Economic Review* 39 (4), 863–883.
- Diebold, F. X., Schuermann, T., Stroughair, J. D., 2000. Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management. *Journal of Risk Finance* 1 (2), 30–35.
- Diks, C., V. P., van Dijk, D., 2010. Out-of-sample Comparison of Copula Specifications in Multivariate Density Forecasts. *Journal of Economic Dynamics and Control* 34, 1596–1609.
- Duffie, D., Saita, L., Wang, K., 2007. Multi-period Corporate Default Prediction with Stochastic Covariates. *Journal of Financial Economics* 83 (3), 635–665.
- Duffie, D., Singleton, K. J., 2003. *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press.
- Efron, B., Tibshirani, R. J., 1997. *An introduction to the bootstrap*. Chapman & Hall.
- Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. *Modelling Extreme Events for Insurance and Finance*. Springer Verlag.
- Embrechts, P., McNeil, A. J., Straumann, D., 2002. Correlation and Dependence in Risk Management: Properties and Pitfalls. In: Dempster, M. A. H. (Ed.), *Value at Risk and Beyond*. Cambridge University Press, Ch. 7, pp. 176–223.
- Embrechts, P., Resnick, S. I., Samorodnitsky, G., 1999. Extreme Value Theory as a Risk Management Tool. *North American Actuarial Journal* 3 (2), 30–41.
- Engle, R., 2002. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics* 20 (3), 339–350.

- Engle, R. F., Manganelli, S., 2004. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business and Economic Statistics* 22 (4), 367–381.
- Eom, Y., Helwege, J., Huang, J., 2004. Structural Models of Corporate Bond Pricing: An Empirical Analysis. *Review of Financial Studies* 17 (2), 499–544.
- Ericsson, J., Jacobs, K., Oviedo-Helfenberger, R., 2004. The Determinants of CDS Premia. *Journal of Fixed Income*, 1059–8596.
- Evans, C. L., Liu, C. T., Pham-Kanter, G., 2002. The 2001 Recession and the CFNAI: Identifying Business Cycle Turning Points. *Economic Perspectives* 26 (3), 26–43.
- Faff, R. F., Parwada, J. T., Poh, H.-L., 2007. The Information Content of Australian Managed Fund Ratings. *Journal of Business Finance & Accounting* 34 (9), 1528–1547.
- Fernandez, C., Steel, M. F. J., 1998. On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association* 93 (441), 359–371.
- Figlewski, S., Frydman, H., Liang, W., 2012. Modeling the Effect of Macroeconomic Factors on Corporate Default and Credit Rating Transitions. *International Review of Economics & Finance* 21 (1), 87–105.
- Fisher, R., Tippett, L., 1928. Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample. *Mathematical Proceedings of the Cambridge Philosophical Society* 24, 180–190.
- Francis, W. B., Osborne, M., 2012. Capital Requirements and Bank Behaviour in the UK: Are There Lessons for International Capital Standards? *Journal of Banking & Finance* 36 (3), 803–816.
- Frydman, H., Schuermann, T., 2008. Credit Rating Dynamics and Markov Mixture Models. *Journal of Banking & Finance* 32 (6), 1062–1075.

- Fuertes, A., Olmo, J., 2012. Optimally Harnessing Inter-Day and Intra-Day Information for Daily Value-at-Risk Prediction. *International Journal of Forecasting* 29 (1), 28–42.
- Fuertes, A.-M., Kalotychou, E., 2007. On Sovereign Credit Migration: A Study of Alternative Estimators and Rating Dynamics. *Computational Statistics & Data Analysis* 51 (7), 3448–3483.
- Fung, H.-G., Sierra, G. E., Yau, J., Zhang, G., 2008. Are the US Stock Market and Credit Default Swap Market Related? Evidence from the CDX Indices. *The Journal of Alternative Investments* 11 (1), 43–61.
- Gaivoronski, A. A., Pflug, G., 2005. Value at Risk in Portfolio Optimization: Properties and Computational Approach. *Journal of Risk* 7 (2), 1–31.
- Galbraith, J. W., Tkacz, G., 2007. Forecast Content and Content Horizons for some Important Macroeconomic Time Series. *Canadian Journal of Economics* 40 (3), 935–953.
- Garcia, R., Tsafack, G., 2011. Dependence Structure and Extreme Comovement In International Equity and Bond Markets. *Journal of Banking & Finance* 35 (8), 1954–1970.
- Gencay, R., Selcuk, F., 2004. Extreme Value Theory and Value-at-Risk: Relative Performance in Emerging Markets. *International Journal of Forecasting* 20 (2), 287–303.
- Giammarino, F., Barrieu, P., 2009. A Semiparametric Model for the Systematic Factors of Portfolio Credit Risk Premia. *Journal of Empirical Finance* 16 (4), 655–670.
- Goodhart, C. A., 2006. A Framework for Assessing Financial Stability. *Journal of Banking & Finance* 30 (12), 3415–3422.
- Gordy, M. B., Howells, B., 2006. Procyclicality in Basel II: Can We Treat the Disease Without Killing the Patient? *Journal of Financial Intermediation* 15 (3), 395–417.

- Hamilton, D. T., Cantor, R., 2004. Rating Rransitions and Defaults Conditional on Watchlist, Outlook and Rating History. Tech. rep., Moody's Investors Service.
- Hamilton, J. D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57 (2), 357–384.
- Hansen, B. E., 1994. Autoregressive Conditional Density Estimation. *International Economic Review* 35 (3), 705–730.
- Hanson, S., Schuermann, T., 2006. Confidence Intervals for Probabilities of Default. *Journal of Banking & Finance* 30 (8), 2281–2301.
- Haque, M., Varela, O., Hassan, M., 2007. Safety-first and Extreme Value Bilateral US–Mexican Portfolio Optimization around the Peso Crisis and NAFTA in 1994. *The Quarterly Review of Economics and Finance* 47 (3), 449–469.
- Hill, P., Faff, R., 2010. The Market Impact of Relative Agency Activity in the Sovereign Ratings Market. *Journal of Business Finance & Accounting* 37 (9-10), 1309–1347.
- Hsieh, D., 1989. Testing for Nonlinear Dependence in Daily Foreign Exchange Rates. *The Journal of Business* 62 (3), 339–368.
- Hu, W., 2009. Calibration of Multivariate Generalized Hyperbolic Distributions: Applications in Risk Management, Portfolio optimization, and Portfolio Credit Risk. VDM Verlag Dr. Müller.
- Hull, J., White, A., 2006. Valuing credit derivatives using an implied copula approach. *Journal of Derivatives* 14 (2), 8–28.
- Hull, J. C., Nelken, I., White, A. D., 2004a. Merton's Model, Credit Risk and Volatility Skews. *The Journal of Credit Risk* 1 (1), 1–27.
- Hull, J. C., Predescu, M., White, A. D., 2004b. The Relationship Between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements. *Journal of Banking & Finance* 28 (11), 2789–2811.

- Jafry, Y., Schuermann, T., 2004. Measurement, Estimation and Comparison of Credit Migration Matrices. *Journal of Banking & Finance* 28 (11), 2603–2639.
- Jarrow, R. A., Lando, D., Turnbull, S. M., 1997. A Markov Model for The Term Structure of Credit Risk Spreads. *Review of Financial Studies* 10 (2), 481–523.
- Jenkinson, A. F., 1955. The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements. *Quarterly Journal of the Royal Meteorological Society* 81 (348), 158–171.
- Jondeau, E., Rockinger, M., 2006. The Copula-GARCH Model of Conditional Dependencies: An International Stock Market Application. *Journal of International Money and Finance* 25 (5), 827–853.
- Jones, E., Mason, S., Rosenfeld, E., 1984. Contingent Claims Analysis of Corporate Capital Structures. *Journal of Finance* 39 (3), 611–625.
- Kavvathas, D., 2001. Estimating Credit Rating Transition Probabilities for Corporate Bonds. Working Paper - University of Chicago.
URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=252517
- Kim, C.-J., 1994. Dynamic Linear Models with Markov-switching. *Journal of Econometrics* 60 (1-2), 1–22.
- Kim, I. J., Ramaswamy, K., Sundaresan, S., 1993. Does Default Risk in Coupons Affect the Valuation of Corporate Bonds?: A Contingent Claims Model. *Financial Management*, 117–131.
- Kisgen, D. J., 2009. Do Firms Target Credit Ratings or Leverage Levels. *Journal of Financial & Quantitative Analysis* 44 (6), 1323–1344.
- Kloman, H. F., 1990. Risk Management Agonists. *Risk Analysis* 10 (2), 201–205.
- Koopman, S. J., Lucas, A., Monteiro, A., 2008. The Multi-State Latent Factor Intensity Model for Credit Rating Transitions. *Journal of Econometrics* 142 (1), 399–424.

- Krokhmal, P., Palmquist, J., Uryasev, S., 2002. Portfolio Optimization with Conditional Value-at-Risk Objective and Constraints. *Journal of Risk* 4, 43–68.
- Kuester, K., Mittnik, S., Paolella, M., 2006. Value-at-Risk Prediction: A Comparison of Alternative Strategies. *Journal of Financial Econometrics* 4 (1), 53–89.
- Kupiec, P. H., 1995. Techniques for Verifying the Accuracy of Risk Management Models. *Journal of Derivatives* 3 (2), 73–84.
- Lando, D., Skodeberg, T., 2002. Analyzing Rating Transitions and Rating Drift with Continuous Observations. *Journal of Banking & Finance* 26 (2-3), 423–444.
- Levin, C., Coburn, T., 2011. Wall Street and the Financial Crisis: Anatomy of a Financial Collapse. U.S. Senate Permanent Subcommittee on Investigations.
URL <http://hsgac.senate.gov/public/>
- Löffler, G., Posch, P. N., 2007. Credit Risk Modeling Using Excel and VBA. Wiley.
- Longerstaey, J., Spencer, M., 1994. RiskMetrics - Technical Document. Tech. rep., JP Morgen.
- Longin, F., Solnik, B., 1995. Is the Correlation in International Equity Returns Constant: 1960-1990. *Journal of International Money and Finance* 14 (1), 3–26.
- Longin, F., Solnik, B., 2001. Extreme Correlation of International Equity Markets. *Journal of Finance* 56 (2), 649–676.
- Longstaff, F. A., Mithal, S., Neis, E., 2003. The Credit-Default Swap Market: Is Credit Protection Priced Correctly? Working Paper - USC FBE Finance Seminar.
URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.199.8915>
- Longstaff, F. A., Mithal, S., Neis, E., 2005. Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. *Journal of Finance* 60 (5), 2213–2253.

- Longstaff, F. A., Schwartz, E. S., 1995. A Simple Approach to Valuing Risky Fixed and Floating Rate Debt. *Journal of Finance* 50, 789–819.
- Lopez, J. A., 1998. Methods for Evaluating Value-at-Risk Estimates. Federal Reserve Bank of New York research paper.
URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1029673
- Madan, D., Unal, H., 2000. A Two-factor Hazard Rate Model for Pricing Risky Debt and the Term Structure of Credit Spreads. *Journal of Financial and Quantitative Analysis* 35 (1), 43–65.
- Mählmann, T., 2009. Multiple Credit Ratings, Cost of Debt and Self-Selection. *Journal of Business Finance & Accounting* 36 (9-10), 1228–1251.
- Manner, H., Reznikova, O., 2010. A Survey on Time-varying Copulas: Specification, Simulations and Application. Working Paper - University of Cologne.
URL http://www.gloriamundi.org/Library_Journal_View.asp?Journal_id=10800
- Markit, 2010. Markit Credit Indices - A Primer. Tech. rep., Markit Group Limited.
- Markowitz, H., 1952. Portfolio Selection. *Journal of Finance* 7 (1), 77–91.
- Mashal, R., Zeevi, A., 2002. Beyond Correlation: Extreme Co-movements between Financial Assets. Working Paper - Columbia University.
URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=317122
- McNeil, A., Frey, R., 2001. Estimation of Tail Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach. *Journal of Empirical Finance* 7 (3), 271–300.
- McNeil, A. J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques, Tools. Princeton.
- Merton, R. C., 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* 2 (2), 449–470.

- Nelsen, R. B., 2006. An Introduction to Copulas. Springer Verlag.
- Nickell, P., Perraudin, W., Varotto, S., 2000. Stability of Ratings Transitions. *Journal of Banking & Finance* 24 (1), 203–227.
- Norden, L., Weber, M., 2009. The Co-movement of Credit Default Swap, Bond and Stock Markets: An Empirical Analysis. *European Financial Management* 15 (3), 529–562.
- Okimoto, T., 2008. New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis* 43 (3), 787–816.
- Patton, A. J., 2006. Modelling Asymmetric Exchange Rate Dependence. *International Economic Review* 47 (2), 527–556.
- Pickands, J., 1975. Statistical Inference Using Extreme Order Statistics. *Annals of Statistics* 119-131, 3.
- Rockafellar, R., Uryasev, S., 2000. Optimization of Conditional Value-at-Risk. *Journal of Risk* 2 (2), 21–42.
- Rockafellar, R. T., Uryasev, S., 2002. Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking & Finance* 26 (7), 1443–1471.
- Rodriguez, J. C., 2007. Measuring Financial Contagion: A Copula Approach. *Journal of Empirical Finance* 14 (3), 401–423.
- Scheinkman, J. A., Lebaron, B., 1989. Nonlinear Dynamics and Stock Returns. *The Journal of Business* 62 (3), 311–337.
- Shaw, W. T., Lee, K. A., 2007. Copula Methods vs Canonical Multivariate Distributions: The Multivariate Student t Distribution with General Degrees of Freedom. Working Paper - King's College London.
- URL http://www.defaultrisk.com/pp_corr_97.htm

- Shcherbakova, A. V., 2008. Credit Ratings Migration: Quantifying Obligor Risk. Working Paper - Emory University.
URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1214062
- Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231.
- Stefanescu, C., Tunaru, R., Turnbull, S., 2009. The Credit Rating Process and Estimation of Transition Probabilities: A Bayesian Approach. *Journal of Empirical Finance* 16 (2), 216–234.
- Stock, J., Watson, M., 1999. Forecasting Inflation. *Journal of Monetary Economics* 44 (2), 293–335.
- Trueck, S., 2008. Forecasting Credit Migration Matrices with Business Cycle Effects - A Model Comparison. *The European Journal of Finance* 14, 359–379.
- Wei, J. Z., 2003. A Multi-Factor, Credit Migration Model for Sovereign and Corporate Debts. *Journal of International Money & Finance* 22, 709–735.
- Xiong, J. X., Idzorek, T., 2011. Mean-Variance Versus Mean-Conditional Value-at-Risk Optimization - The Impact of Incorporating Fat Tails and Skewness into the Asset Allocation Decision. *Financial Analysts Journal* 67 (2), 23–35.
- Yu, F., 2006. How Profitable is Capital Structure Arbitrage? *Financial Analysts Journal* 62 (5), 47–62.
- Yu, J., Yang, X., Li, S., 2009. Portfolio Optimization with CVaR under VG Process. *Research in International Business and Finance* 23 (1), 107–116.
- Zhang, B. Y., Zhou, H., Zhu, H., 2009. Explaining Credit Default Swap Spreads with Equity Volatility and Jump Risks of Individual Firms. *The Review of Financial Studies* 22 (12), 5099–5131.
- Zhu, L., Coleman, T. F., Li, Y., 2009. Min-max Robust and CVaR Robust Mean-variance Portfolios. *Journal of Risk* 11 (3), 55–85.