Modeling Time Variation in Systemic Risk
ISBN 798 90 361 0343 5

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 551 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.
Modeling Time Variation in Systemic Risk

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Economische Wetenschappen en Bedrijfskunde
op maandag 11 februari 2013 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Xin Zhang

geboren te Anhui, China
promotoren: Prof.dr. A. Lucas
              Prof.dr. S.J. Koopman
# Contents

1 **Introduction**  
1.1 Financial risk modeling .................................. 3  
1.1.1 Systemic risk modeling ................................ 3  
1.1.2 Sovereign risk modeling ................................ 6  
1.1.3 Modeling challenges ................................... 7  
1.2 Multivariate dynamic non-normal risk models .............. 9  
1.3 Outline ..................................................... 12

2 **Modeling Dynamic Volatilities and Correlations under Skewness and Fat Tails**  
2.1 Introduction ............................................... 15  
2.2 The dynamic GH model .................................... 19  
2.3 Model parameterizations ................................... 23  
2.4 Time-varying scale matrix and an EM algorithm .......... 26  
2.5 Monte Carlo evidence ..................................... 29  
2.5.1 Simulation design ...................................... 29  
2.5.2 Simulation results ..................................... 30  
2.6 Empirical application ..................................... 31  
2.7 Conclusion ................................................ 38  
Appendices to Chapter 2 ........................................ 40

3 **Conditional Probabilities for Euro Area Sovereign Default Risk**  
3.1 Introduction ............................................... 45  
3.2 Conceptual framework ..................................... 48  
3.3 Statistical model .......................................... 51
## CONTENTS

3.3.1 Generalized Autoregressive Score dynamics .................................. 51
3.3.2 Parameter estimation ................................................................. 54
3.4 Empirical application: Euro area sovereign risk .................................. 55
   3.4.1 CDS data ................................................................................. 55
   3.4.2 Marginal and joint risk .............................................................. 56
   3.4.3 Joint probabilities of Eurozone financial stress ................................. 62
   3.4.4 Spillover measures: What if . . . failed? ........................................ 65
   3.4.5 Event study: the May 9, 2010 rescue package and risk dependence . 71
3.5 Conclusion ....................................................................................... 75
Appendix to Chapter 3 ............................................................................ 75

4 Measuring Credit Risk in a Large Banking System: Econometric Mod-
eling and Empirics ................................................................. 79
   4.1 Introduction ................................................................................... 79
   4.2 A framework for simultaneous financial firm failures ........................... 82
   4.3 The model .................................................................................... 84
      4.3.1 The Dynamic GH skewed-\( t \) model ........................................... 84
      4.3.2 Estimation and restrictions ....................................................... 86
      4.3.3 The Block GAS-Equirrelation model ........................................ 88
   4.4 The risk measures in a large system .................................................. 91
   4.5 Empirical application .................................................................... 94
      4.5.1 The system of major European banks ........................................ 94
      4.5.2 European large financial institutions ....................................... 100
   4.6 What factors drive the bank equity correlation? ................................. 104
   4.7 Conclusion .................................................................................... 111
Appendix to Chapter 4 ............................................................................ 113

5 Conclusion ....................................................................................... 117
   5.1 Summary of the main conclusions ................................................... 117
   5.2 Direction for future research .......................................................... 119

Bibliography ......................................................................................... 120

Samenvatting (Summary in Dutch) ................................................. 131
Chapter 1

Introduction

Over the past few decades, researchers have been developing models to study the risks of different financial institutions and financial instruments, see Christoffersen (2011). But interestingly, a large proportion of studies focuses on the understanding of individual risks rather than systemic risk. The difficulty of studying systemic risk arises from the interconnectedness of member institutions in a large financial system. For example, as default risks are correlated, systemic risk is not a simple addition of individual risks. We need to incorporate the possibly changing correlation and dependence structure among risk factors, which cannot be observed or quantified directly. Part of the changes in dependence are due to the changes in latent economic processes and financial market conditions, leading to the time variation in systemic linkages and systemic risk. This thesis is written to provide a unified framework to measure the dynamics in systemic risk of financial systems and sovereign systems.

Originally used to measure the risk of bank runs and currency crises (Billio et al. (2012)), systemic risk has been extended to describe the risk of any system-wide breakdown. Systemic risk is one of the biggest threats to the stability of modern financial systems, as financial institutions are more and more interconnected with visible links, for instance bank loans, and invisible links such as the common exposure to counterparty risks. The correlated and complex global financial system creates a network in which marginal credit events may drive up the risks in other institutions, and lead to further failures of multiple institutions. From a risk management perspective, the clustering of risks is of particular interest. What seems manageable in isolation may not be so if the rest of the system is also under stress, see Acharya et al. (2010). A good example is the
failure of Lehman Brothers, which occurred at a time when many other financial institutions were also under distress. As a result, their collapse was perceived to be much more severe.

The same intuition that applies to financial institutions, also applies to the system of Euro Area sovereigns. The dependence of sovereign risk is a natural consequence of a large degree of legal, economic, and financial interconnectedness between Euro Area member states. Such dependence is also expected to be time-varying, as market participants adapt to changes in policy and in the institutional settings. While adverse developments in one country’s public finances could perhaps still be handled with the support of a healthy remainder of the union, the situation may quickly become untenable if two, three, four, five, or more, of its members are also distressed at the same time. From a systemic perspective, it is necessary to treat the individual sovereign risks as a collection of dependent risks. Despite all the efforts for understanding corporate defaults, there is no commonly accepted framework and measurement system for sovereign credit risk. We will explore the systemic nature of Euro Area sovereign risks and use a general unified model for both financial institutions’ and sovereign credit risks.

In the interlinked system of multiple firms or sovereigns, the investigation of systemic risk involves a thorough study on the time variation of financial risks, and the dependence structure which may lead to systemic credit events. As the strong connection between the firms or members cannot be observed and quantified directly, we need to model the dynamic dependence of the default risks in the system. The risk assessing framework in this thesis is fitted to the financial markets, using equity returns and credit default swap spread changes. Both of these are usually asymmetrically distributed with fat-tails. Their variance and correlation are also changing over time, especially during crisis times. Our model is general enough to fit these non-Gaussian features and capture the time variation in conditional covariances as well. With the estimated correlations and variances, we are able to produce the time-varying risk assessment for individual member and for the whole financial system. This model also allows for the determination of the contagion effects of individual credit events on the computation of joint risk measure for the whole system.
1.1 Financial risk modeling

The studies on modeling financial risk date back to the start of modern financial research. Recognized by many researchers, the basic trade-off between expected return and risk is critical for economic decisions and financial activities.

Financial risk models can be categorized into aggregated and disaggregated models, see Andersen et al. (2007). Aggregated risk modeling focuses on portfolio-level studies. These models are univariate and study the risk and return on entire financial portfolios. Hence it is not suitable for active risk management that may require detailed understanding of financial risk associated with a single asset component. Disaggregated modeling is based on asset-level analysis and more suited for this purpose. Disaggregate models are multivariate models that capture the movements of all risk factors in an interdependent system. Within an asset-level study, we can identify and attribute the portfolio loss to the losses from individual counterparties in the portfolio. Both the aggregate and disaggregate approach focus on losses at the institutional level at most. Over the past few years, however, we have seen a revived interest in risk measurement at the system level. It is to this that we turn next to the systemic risk modeling.

1.1.1 Systemic risk modeling

Systemic risk is originally associated to currency crises and banks’ vulnerability to deposit runs. Later the scope of investigation has been extended to the contagion effect among financial intermediaries, with an emphasis on banks, see De Bandt and Hartmann (2000). The two most recent crises remind us that we might have to rethink the definition of systemic risk, as several non-bank sectors overlooked in the past appeared to be systemically important as well. Brunnermeier et al. (2009) argue that systemic risk should be considered as the risk of “negative risk spill-over” from individual credit distress to the highly interconnected financial system. In the thesis, we broadly define financial systemic risk as the risk of simultaneous failures of several financial institutions in the system, which leads to a higher likelihood of systemic breakdown.

The globalization and integration of financial markets link financial institutions stronger across countries and regions. Financial innovations in the industry like credit derivatives and asset backed securities increase the complexity of unobservable dependence among
the underlying risks. In the modern financial industry, banks, insurance companies, brokers/dealers and the shadow banking system all contribute to financial systemic risk, see Billio et al. (2012) and Gorton et al. (2010). Hence it is crucial to provide a unified econometric framework to measure systemic risk in these sectors.

The literature on financial systemic risk is already large and still growing. Much work has been devoted to theories that explain bank runs and bank contagion. The paper of Diamond and Dybvig (1983) paved the way for studying the instability of a single bank. Since then, many researchers extended their framework to study multiple bank systems and allow contagion across the financial sector. De Bandt and Hartmann (2000) review the developments in modeling systemic risk. The Asian financial crisis and the global financial crisis brought systemic risk back to the center of attention, especially in measuring systemic risk and systemic risk contributions.

The models on systemic risk measurement can be loosely classified into the following groups. Following the contingent claims analysis, Lehar (2005) estimates the dynamic correlation between banks’ assets using stock market data and a Merton model (Merton (1974)). He uses the default probabilities of financial institutions as the measure of systemic risk. Gray et al. (2007b) also use a contingent claims approach to provide an overall measure of systemic risk. Gray and Jobst (2009) apply a similar method to quantify the systemic risk contribution of large institutions. Some researchers take a game-theoretic perspective in explaining systemic risk and computing the risk contributions. Tarashev et al. (2010) propose an allocation of capital charges to financial institutions based on their systemic importance. Their calculation of the systemic risk contribution is based on Shapley values (Shapley (1953)). Farhi and Tirole (2009) presents a model for collective moral hazard and systemic bailouts. Another stream of literature considers the application of extreme value theory (EVT). The study on extreme dependence is first conducted by Mandelbrot (1963) and Jansen and De Vries (1991) studies stock return tail behaviors. Poon et al. (2004) and Hartmann et al. (2004) show the potential of modeling systemic risk with extreme value theory. Zhou (2010) estimates the systemic importance of banks under the extreme value theory framework.

One major alternative approach to construct systemic risk indicator is to view the financial system as a portfolio of equities or a basket of financial instruments. Assume the
1.1. **FINANCIAL RISK MODELING**

The financial system is a portfolio with return \( Y \) and consisting of \( N \) institutions' equities,

\[
Y = \sum_{i}^{N} w_{i} y_{i},
\]

where \( w_{i} \) is the weight on institution \( i \)'s equity return \( y_{i} \). Classic tools for risk management like Value-at-Risk (VaR) and Expected-Shortfall (ES) can be applied to this portfolio. The Value-at-Risk is defined as the maximum portfolio loss at a confidence level \( 1 - \alpha \). So the Value-at-Risk is equal to \(-Y_{\alpha}\), where

\[
Y_{\alpha} = \sup\{\tilde{Y} | P[Y < \tilde{Y}] \leq \alpha\}. \tag{1.2}
\]

The Expected-Shortfall measures the expected loss conditional on the return being less than a certain threshold \( Y_{\alpha} \),

\[
ES_{\alpha} = -E[Y | Y \leq Y_{\alpha}]. \tag{1.3}
\]

These two measures are adjusted and often used to measure systemic risk. For example, Adrian and Brunnermeier (2009) use the Value-at-Risk as the systemic risk indicator conditional on that a bank has had a VaR loss. Acharya et al. (2010) propose “Systemic Expected Shortfall” (SES) to measure the financial institution’s contribution to systemic risk. They conceptualize the contribution as the institution’s losses when the system is under-capitalized.

Some recent studies construct systemic risk measure by studying the multivariate density \( p(y, \theta) \) estimated from asset returns,

\[
y = (y_{1}, \cdots, y_{N})' \sim p(y; \theta), \tag{1.4}
\]

where \( \theta \) is a vector of unknown parameters in the model. Garcia Pascual et al. (2006) assess financial failure in a Gaussian factor model setting. Their determination of joint failure probabilities is in part based on the notion of an \( n \)th-to-default CDS basket, which can be set up and priced as suggested in Hull and White (2004). Huang et al. (2009) use credit default swap data and stock return correlations to compute systemic risk. They measure systemic risk as the theoretical price of insurance against a certain percentage loss on the portfolio. Alternatively, Segoviano and Goodhart (2009) propose a non-parametric
copula approach that reflects individual failure probabilities. They define Banking Stability Measures as probability statements about a certain number of banks falling into financial distress. In this thesis, we follow this disaggregated risk modeling approach and discuss credit risk in the system of financial institutions. The systemic risk indicators are defined as conditional and unconditional probabilities of joint defaults of the institutions, given the dynamic non-Gaussian multivariate distribution estimated from historical asset returns.

1.1.2 Sovereign risk modeling

The supranational bailouts of Greece, Ireland, Portugal and a series of more than 16 European Union summits ever since the start of the sovereign debt crisis eliminate any doubt about the relevance of sovereign risk modeling. Despite the long list of sovereign debt literature, little is known about the nature of sovereign systemic risk and the measurement of sovereign risk in advanced economies. The credit risk on Euro Area insolvent governments is not the same as, but quite comparable to financial systemic risk, especially if we notice the strong interconnectedness of the Euro Area due to the common currency.

Research on sovereign credit risk first gained attention in the 1980s due to the possible default of the developing countries at that time. After the 1990s, we observed the defaults on domestic/external debt in Russia, Ecuador, Argentina and semi-coercive restructuring of sovereign debts in Ukraine, Pakistan and Uruguay. There are also several cases that sovereign debt crisis were partly avoided via large amounts of support from international organizations and the private sector.

The current literature dealing with sovereign debt crises falls into four categories as in Manasse et al. (2003): theoretical models of sovereign default; empirical studies of the determinants of debt crises; empirical studies of the predictive power of credit ratings; and empirical studies of the determinants of sovereign credit spreads. Most theoretical research focuses on the incentive to repay the sovereign debt and the strategic default decision. A list of such references includes Eaton and Gersovitz (1987), Grossman and Van Huyck (1988), Bulow and Rogoff (1987), Atkeson (1991), Dooley and Svensson (1994), Cole and Kehoe (1996), Dooley (2000), Guembel and Sussman (2009) or Yue (2010) and many other recent papers. Research on the determinants of debt crises aim at finding the factors influencing the probability of a sovereign debt crisis. For example, see Detragiache
1.1. FINANCIAL RISK MODELING

and Spilimbergo (2001), Reinhart (2002), Catão and Sutton (2002), Hemming et al. (2003) and many others. Determinants of credit spreads have been another important element in the literature of sovereign credit risk. Dell’Ariccia et al. (2002) find a significant increase of spreads after the Russian Crisis of 1998. The recent paper by Gerlach et al. (2010) find that countries with large and highly leveraged banking sectors witness increasing yield spreads, reflecting the market’s expectations for prospective bailouts and increase of sovereign default risk. A line of literature that investigates the link between sovereign credit risk, country ratings, and macro fundamentals; see for example Haugh et al. (2009), Hilscher and Nosbusch (2010), or De Grauwe and Ji (2012). Unlike the private sector risk, it is difficult to adopt the contingent claims analysis (CCA) known as Merton model to sovereigns. We have to define the sovereign balance sheet and especially the assets of the country. Gray et al. (2007a) provide a similar approach in considering sovereign default events.

Recently some researchers attempt to disentangle the global, systemic and individual factors in determining credit spreads using asset pricing methodology. A few papers provide evidence that sovereign credit spreads are related to common global and financial market factors, see for example Pan and Singleton (2008) and Longstaff et al. (2011). With the comparison of Credit Default Swaps (CDS) structures in the U.S. and European Monetary Union (EMU), Ang and Longstaff (2011) conclude that European sovereigns have more systemic credit risk than the U.S. sovereigns, despite the stronger macroeconomic fundamentals in the U.S. Motivated by the empirical findings, we focus on modeling systemic sovereign risk in the Euro Area during the sovereign debt crisis. Credit derivatives written on sovereign bonds, such as Credit Default Swaps (CDS), contain implicit information on the sovereign default risk. So we propose a disaggregated model to assess the correlated credit risk in a portfolio of sovereign CDS contracts. We use Equation (1.4) again with a renewed definition of $y_i$ as the credit spread change on country $i$. The main goal of this analysis is to provide a quantitative analysis on the Euro Area sovereign credit risk, defined as the probability of sovereign failures, conditionally and unconditionally.

1.1.3 Modeling challenges

The investigation of systemic risk and sovereign credit risk requires a model that captures both the individual credit risk and the dependence structure of marginal risk factors. One
possible approach to analyze asset-specific and interrelated risks is to build a disaggregated model. However, there are a few difficulties researchers have to confront when dealing with disaggregated risk models.

The first difficulty arises from the dynamic nature of the financial risks. Financial risks are changing over time, as the economy fluctuates. The well-documented phenomenon, such as the business cycle and the credit cycle (Kiyotaki and Moore (1995), Koopman and Lucas (2005), Koopman et al. (2009) and Giannone et al. (2012)), influences the dynamics in financial risks. As a consequence of the evolving financial system, increasing institutional linkages and market integration have led to increased interdependencies among financial institutions and markets. Bekaert and Harvey (1995) show the time-varying integration in a few financial markets. Hardouvelis et al. (2006) examine the EMU stock market integration due to the launch of the single currency.

The second difficulty exists not only in the risk management literature, but also more widely in other areas of finance. Most finance models, like the CAPM and Black-Scholes-Merton model, have the underlying assumption of a Gaussian distribution. But the non-Gaussian properties of financial returns are widely-accepted findings in the literature. Financial asset returns exhibit non-zero skewness and excess kurtosis, see for example Eberlein and Keller (1995), Franses and van Dijk (2000) and Engle (2002). How to incorporate these salient features of financial data is an inevitable problem for financial risk modeling.

Last but not least, the disaggregated multivariate risk modeling is naturally prone to the “curse of dimensionality” (Bellman (1957)). As financial risks are usually correlated and interconnected, it is not trivial to model the dependence structure when the dimension is large. It is necessary to use a tractable model for the high-dimensional covariance matrices. We have to adopt dimensionality reduction techniques or a parsimonious covariance matrix structure.
1.2 Multivariate dynamic non-normal risk models

To tackle the challenges in modeling dynamic financial risks, many multivariate models have been proposed with a time-varying covariance matrix. The notion of volatility and correlation has been playing a central role in the risk management area. Sharpe (1975) points out that volatility is “a more relevant measure of risk”. The CAPM theory and the Arbitrage Pricing Theory use covariance as a measure of dependence between different financial instruments. The literature on portfolio risk management shows the necessity of accurate volatility and correlation modeling. Berkowitz and O’Brien (2002) show that existing bank risk models perform poorly and are easily outperformed by a simple univariate GARCH model. In the multivariate setting, Andersen et al. (2007) shows the usefulness of multivariate GARCH models in risk management modeling.

The literature on time-varying parameter models for volatilities and correlations is vast. A number of models for time-varying volatilities and correlations have been put forward in the literature. The simplest approach consists of using a rolling window or the RiskMatrics model. Other options include multivariate generalizations of the Stochastic Volatility model, see the overview in Shephard (2005), or the multivariate GARCH class of models, see for example the VEC model and the BEKK model by Engle and Kroner (1995), the Orthogonal GARCH by Alexander (1998) and Alexander (2001), and the Generalized Orthogonal GARCH by van der Weide (2002) and Boswijk and Van der Weide (2006). To reduce the number of parameters to be estimated, Bollerslev (1990) introduced the CCC model, which has GARCH volatility dynamics, but constant conditional correlations. Engle (2002) generalized the CCC model to a Dynamic Conditional Correlation (DCC) model, making the conditional correlation matrix time-varying while at the same time retaining parsimony. The \( N \)-dimensional financial portfolio follows a dynamic multivariate density \( y_t \sim p(y_t; \theta_t) \) with DCC covariance matrix \( \Sigma_t \in \theta_t \),

\[
\begin{align*}
\Sigma_t &= D_t R_t D_t, \\
D_t^2 &= \text{diag}\{\omega\} + \text{diag}\{\kappa\} \circ y_{t-1} y'_{t-1} + \text{diag}\{\lambda\} \circ D_{t-1}^2, \\
\epsilon_t &= D_t^{-1} y_t, \\
Q_t &= S \circ (\ell t' - A - B) + A \circ \epsilon_{t-1} \epsilon'_{t-1} + B \circ Q_{t-1}, \\
R_t &= \text{diag}\{Q_t\}^{-1} Q_t \text{diag}\{Q_t\}^{-1}.
\end{align*}
\]
Here $\ell$ is a vector of ones and $\circ$ is the Hadamard product of two identically sized matrices, $D_t$ is a diagonal matrix with volatilities in the diagonal and $Q_t$ is a decomposition of correlation matrix $R_t$ with non-identified parameters on the diagonal. The matrix $S$ is the unconditional correlation matrix of standardized returns $\epsilon_t$. $\text{diag}\{\cdot\}$ operation produces a diagonal matrix by keeping the diagonal of a matrix $Q_t$, or place a vector $\omega$ in the diagonal. $\omega$, $\kappa$ and $\lambda$ are $N \times 1$ parameter vectors of the GARCH volatility models. $A$ and $B$ are $N \times N$ parameter matrices in the correlation processes. But usually $A$ and $B$ are a scalar or diagonal rather than a whole matrix in practice.

Several researchers have made contributions in the area of DCC correlation modeling. Engle and Sheppard (2001) developed the theoretical and empirical properties of the DCC model. Aielli (2011) found out the inconsistency of the DCC estimator and provided a tractable alternative model (cDCC model). Boudt et al. (2011) proposed a robust multivariate volatility forecasting model from an extension of the DCC model.

The (M)GARCH model works well if the parametric assumption of $y_t$ is Gaussian. The GARCH equation indicates that we should increase the variance if we observe large absolute excess returns. It is intuitive because the thin tail of the Gaussian distribution implies that large excess returns should be fully attributed to a bigger volatility. The logic is less reasonable when we use non-Gaussian distributions for the GARCH model. If we assume a Student’s $t$ distribution as the GARCH-$t$ model of Bollerslev (1987), the large value of $|y_t|$ may simply be a result of the fat-tailed nature in the distribution, and need not be a consequence of an increased volatility. It is necessary to consider the impact of distributional assumption on the volatility dynamics we use. Recently Creal et al. (2011) proposed to adopt generalized autoregressive score (GAS) dynamics to obtain a time-varying covariance matrix of the multivariate Student’s $t$ distribution. They provide a correlation model in which the precise form of the non-Gaussian distribution that is used to model the data also governs how time-varying correlations and volatilities depend on realized data. The $t$-GAS model appears to perform better in-sample and out-of-sample, compared to a variety of competing models. The generalized autoregressive score framework of Creal et al. (2012) is a general family of observation driven models. The dynamic parameters we are interested in, for instance the volatilities and correlations,
1.2. MULTIVARIATE DYNAMIC NON-NORMAL RISK MODELS

depend on a dynamic factor $f_t$,

$$
\Sigma_t = h(f_t),
$$

$$
f_{t+1} = \omega + \sum_{i=0}^{p-1} A_i s_{t-i} + \sum_{j=0}^{q-1} B_j f_{t-j},
$$

$$
s_t = S_t \nabla_t,
$$

$$
\nabla_t = \frac{\partial \ln p(y_t; \theta_t)}{\partial f_t},
$$

where $\omega$, $A_i$ and $B_j$ are estimated model parameters, $S_t$ is a scaling matrix for the score $\nabla_t$, usually chosen as the local curvature of the score measured by the inverse information matrix. The GAS model creates an explicit link between the distribution of $y_t$ and the dynamic behavior of $\Sigma_t$. In particular, if $y_t$ is fat-tailed, observations that lie far outside the center automatically have less impact on future values of the time-varying parameters in $\Sigma_t$. Blasques et al. (2012) derive conditions characterizing the stationarity and ergodicity (SE) regions for this type of dynamic recursive models.

As shown earlier, the literature reveals that financial asset returns exhibit not only fat-tailness, but also non-zero skewness. To better accommodate these non-Gaussian features, we extend the GAS framework to our current setting where we have both skewness and kurtosis in a general multivariate distribution. We show that the asymmetric leptokurtic distribution influences the time variation in volatilities and correlations. The new model performs better in simulations and empirical studies. The dynamic mechanism in this model is driven by the GAS process. We can implement the model for a large dimensional dataset by imposing a common factor structure. Based on the restricted and unrestricted model, we propose several credit risk measures for systemic risk and sovereign risk. Hence the new risk model tackles the main challenges in disaggregate risk modeling: the time variation in the dependence structure, the non-Gaussian features of financial data, and the high dimension of the model.
1.3 Outline

In this thesis, we propose a multivariate econometric model for systemic risk and sovereign risk measurement. It is a disaggregated risk model with flexible time-varying dependence structure. As the financial returns are non-normally distributed, we rely on a general non-Gaussian distribution that captures different salient features of the data. Further, we extend the study to financial risk modeling in a large system. Each chapter is self-contained and can be read independently. At the end of each chapter, we will discuss the contribution.

In Chapter 2 (Modeling Dynamic Volatilities and Correlations under Skewness and Fat Tails), we propose a new model for dynamic volatilities and correlations of skewed and heavy-tailed data. Our model endows the Generalized Hyperbolic distribution with time-varying parameters driven by the score of the observation density function. The key novelty in our approach is the fact that the skewed and fat-tailed shape of the distribution directly affects the dynamic behavior of the time-varying parameters. It distinguishes our approach from familiar alternatives such as the generalized autoregressive conditional heteroskedasticity model and the dynamic conditional correlation model where distributional assumptions affect the likelihood but not the parameter dynamics. We present a modified expectation-maximization algorithm to estimate the model. Simulated and empirical evidence shows that the model outperforms its close competitors if skewness and kurtosis are relevant features of the data.

In Chapter 3 (Conditional Probabilities for Euro Area Sovereign Default Risk), we propose a novel empirical framework to assess the likelihood of joint and conditional failure for Euro area sovereigns. Our model is based on a dynamic skewed-$t$ copula which captures all the salient features of the data, including skewed and heavy-tailed changes in the price of CDS protection against sovereign default, as well as dynamic volatilities and correlations to ensure that failure dependence can increase in times of stress. We apply the framework to Euro area sovereign CDS spread changes from 2008 to mid-2011. Our results reveal significant time-variation in risk dependence and considerable spill-over effects in the likelihood of sovereign failures. We also investigate distress dependence around a key policy announcement by Euro area heads of state on May 9, 2010, and demonstrate the importance of capturing higher-order time-varying moments during times of crisis for the correct assessment of interacting risks.
1.3. **OUTLINE**

In Chapter 4 (Measuring credit risk in a large banking system: econometric modeling and empirics), two new measures for financial systemic risk are computed based on the time-varying conditional and unconditional probability of simultaneous failures of several financial institutions. These risk measures are derived from a multivariate model that allows for skewed and heavy-tailed changes in the market value of financial firms' equity. Our model can be interpreted as a Merton model with correlated Lévy drivers. This model incorporates dynamic volatilities and dependence measures and uses the overall information on the shape of the multivariate distribution. Our correlation estimates are robust against possible outliers and influential observations. For very large cross-sectional dimensions, we propose an approximation based on a conditional Law of Large Numbers to compute extreme joint default probabilities. We apply the model to assess the risk of joint financial firm failure in the European Union during the financial crisis. By augmenting the dynamic parameter model with Euribor-EONIA interest rate spreads and other variables that capture situations of systemic stress, we find that including extra economic variables helps to explain systemic correlation dynamics.

Chapter 5 (Conclusion) summarizes the main results and points out possible directions for future research.
Chapter 2

Modeling Dynamic Volatilities and Correlations under Skewness and Fat Tails

2.1 Introduction

We propose a new dynamic observation driven model for correlations and volatilities based on the class of multivariate Generalized Hyperbolic (GH) distributions. The GH distribution was introduced by Barndorff-Nielsen (1977) and further explored in Barndorff-Nielsen (1978) and Blæsild (1981). The distribution’s flexible form accommodates many of the relevant features in financial time series data, such as excess kurtosis, skewness, and time-varying volatilities and correlations; see McNeil et al. (2005), Eberlein and Keller (1995), Franses and van Dijk (2000), Engle (2002) and others.

The dynamics of the time-varying parameters in our GH distributions are driven by the scaled score of the local observation density. This is a distinguishing feature of our approach. By using the density scores, the skewed and fat-tailed nature of the observation distribution not only affects the likelihood, but also the dynamics of the volatilities and correlations. This differentiates our approach from other well-known models where the distributional assumptions affect the likelihood only, but not the parameter dynamics, e.g., the Dynamic Conditional Correlation (DCC) model of Engle (2002) with normal or Student’s $t$ distributed innovations.
The literature on time-varying parameter models for volatilities and correlations is vast. Our model follows the literature on observation driven rather than parameter driven models. For surveys of the latter in the current context, see for example Shephard (2005) and Asai and McAleer (2009). Within the observation driven class of volatility and correlation models, many multivariate extensions of the seminal generalized autoregressive conditional heteroskedasticity (GARCH) model have been proposed. Bollerslev (1990) introduced the idea of having dynamically evolving variance matrices with the individual variances specified as GARCH processes but with the corresponding conditional correlations treated as unknown constants. This specification has become known as the CCC model. Engle (2002) generalized the CCC model by introducing a simple and parsimonious observation driven mechanism for the conditional correlations. The parsimony of the DCC model combined with its time-varying full conditional correlation matrix makes the DCC model attractive to empirical researchers. This feature is retained in the new model proposed in this paper as well. Other multivariate extensions of the GARCH model include the VEC model of Engle and Kroner (1995), the BEKK model of Engle and Kroner (1995), the Orthogonal GARCH model of Alexander (1998) and Alexander (2001), and the Generalized Orthogonal GARCH model of van der Weide (2002) and Boswijk and Van der Weide (2006).

Most of the above models were originally derived under the assumption of a (conditionally) normal distribution for the underlying data. Since then, generalizations have been proposed to accommodate alternative distributions, including the Student's $t$, the skewed $t$, and the GH distribution; see, for example, Bauwens and Laurent (2005), Fiorentini et al. (2003), Hu (2005), Mencia and Sentana (2004), Peters (2001), and Prause (1999). In all of these models, the likelihood changes but the dynamic specifications for volatilities and correlations are unaffected.

It is rather surprising that the form of the distribution should have no impact on the specification of volatility and correlation dynamics. If, for example, the distribution is leptokurtic, we expect to see large (absolute) observations from time to time. The occurrence of a large observation should not automatically be attributed to a recent increase in volatility, as is done in a standard GARCH specification. Similarly, if the data are drawn from a skewed distribution, we would expect large negative or positive observations to convey different signals about current volatility levels. Again, this would
2.1. INTRODUCTION

imply a link between the shape of the observation distribution and the specification of the volatility and correlation dynamics. No such direct link is present in the standard GARCH and DCC models.

Our main contribution in this paper is to provide a general model for time-varying variances and correlations in which the form of the error distribution governs the specification of volatility and correlation dynamics. For this purpose we extend the framework of Creal et al. (2012) and Creal et al. (2011) to a multivariate setting with skewed and heavy-tailed data. Creal et al. (2011) treat the special case where time series are drawn from a multivariate Student’s $t$ distribution. Nelson and Foster (1994) and Harvey and Chakravarttry (2008) give treatments of the univariate version of this model. Similar to these approaches, our model provides an automatic mechanism that limits the impact of outlying or aberrant observations on future correlations and volatilities. At an intuitive level, the new model attributes part of the sign and magnitude of each observation to the skewed and fat-tailed nature of the data generating process rather than to direct changes in volatilities or correlations.

Our results provide a full treatment of skewness and kurtosis effects on volatilities and correlations in a multivariate setting. We show that the volatility and correlation updating mechanism includes a natural asymmetry effect to allow for a different impact of negative versus positive realizations. For example, if the distribution is left-skewed, large negative realizations are more likely and should not automatically be attributed to local volatility increases. A large positive realization for a left-skewed distribution, however, is extremely unlikely unless volatilities or correlations have increased recently. Via the density score, our dynamic specification for volatilities and correlations includes an interaction between the skewness coefficient and past observations. In this way, the possibly asymmetric impact of past observations on future volatilities and correlations enters the dynamic specification in a natural way.

Parameter estimation is straightforward for our model, since the model is defined in conditional terms similar to the standard GARCH model and its multivariate counterparts. This implies that the likelihood function can be specified in closed analytical form and computed using a prediction error decomposition. In the literature, maximum likelihood estimation of the parameters or the GH distribution is often carried out using the Expectation-Maximization (EM) algorithm of Dempster et al. (1977), see Mencia and
Sentana (2004). EM estimation, however, is not straightforward for our new model due to the highly non-linear functions of the data that are used to drive the volatility and correlation dynamics. We show how to modify the standard EM algorithm to our specific setting to make estimation by EM feasible again. The key step is to replace the density score as a driving mechanism by a conditional density score that runs in parallel to the conditional expectations taken in the expectations step of the EM algorithm.

In a simulation experiment, we compare the performance of our new model to its direct competitors, including versions of the DCC model. We carry out simulations with different correlation dynamics and a variety of error distributions. We also consider the DCC model with GH distributed observations. Although it is not our primary focus, the DCC model with GH errors can also be regarded as a contribution of our paper to the current literature. If the true error distribution is normal, differences in performance between the different statistical models are limited. For fat-tailed error distributions, our model with the GH distribution has superior performance. If in addition the error distributions are skewed, our model performs best.

We provide an empirical illustration of the new model to investigate the volatilities and correlations between four blue-chip stocks from different industries. The sample period includes the recent financial crisis. We find that the estimated correlation dynamics differ substantially between our new approach and a traditional DCC models. The new approach seems much less influenced by incidental influential observations. Accounting for the skewness and fat-tailed nature of the data, we show that volatilities for all series are relatively smaller and that the overall persistence of volatilities and correlations is generally higher.

The remainder of the paper is organized as follows. Section 2.2 introduces the model. Section 2.3 discusses some alternative model parameterizations. Section 2.4 extends the model for the scale rather than the covariance matrix and proposes a modified EM algorithm for parameter estimation. Section 2.5 provides Monte Carlo evidence on the performance of the model compared to some of its competitors. Section 2.6 presents the empirical illustration. Section 2.7 concludes.
2.2 The dynamic GH model

We assume our data generating process is given by

$$y_t = L_t \varepsilon_t, \quad \Sigma_t = L_t L'_t,$$

(2.1)

where \(y_t, \varepsilon_t \in \mathbb{R}^k\) for \(t = 1, \ldots, n\), \(L_t\) is a \(k \times k\) lower triangular matrix giving rise to a time-varying \(k \times k\) covariance matrix \(\Sigma_t\), and \(\varepsilon_t\) follows a Generalized Hyperbolic (GH) distribution with zero mean and unit covariance matrix. The specification (2.1) can easily be extended to include a conditional or unconditional non-zero and possibly time-varying mean for \(y_t\). In line with Engle (2002), we further decompose the covariance matrix \(\Sigma_t\) as

$$\Sigma_t = L_t L'_t = D_t R_t D_t,$$

(2.2)

with \(D_t\) a diagonal matrix containing the standard deviations of the elements in \(y_t\), and \(R_t\) the correlation matrix of \(y_t\).

The Generalized Hyperbolic (GH) distribution introduced by Barndorff-Nielsen (1977) is a flexible distribution that accommodates both thin and fat-tailed as well as positively and negatively-skewed distributions. We present the GH class as the normal mean-variance mixture model

$$\varepsilon_t = \mu_\varepsilon + \zeta_t T \gamma + \sqrt{\zeta_t} T z_t, \quad z_t \sim N(0, I_k),$$

(2.3)

where \(\zeta_t \in \mathbb{R}^+\) is a positively valued random scalar that is independent of \(z_t\), \(\mu_\varepsilon \in \mathbb{R}^k\) is the location parameter, \(k \times k\) matrix \(TT'\) is the scaling matrix and \(\gamma \in \mathbb{R}^k\) is the skewness parameter. The GH class includes distributions such as the normal (\(\gamma = 0\) and \(\zeta_t = 1\)), the (skewed) multivariate Student’s \(t\) (for which \(\zeta_t\) has an inverse Gamma distribution with \(\gamma = 0\) for the symmetric case and \(\gamma \neq 0\) for the asymmetric case), the (skewed) variance-gamma distribution (for which \(\zeta_t\) has a Gamma distribution) and the Generalized Hyperbolic distribution (for which \(\zeta_t\) has a Generalized Inverse Gaussian (GIG) distribution with parameters \(\lambda, \chi, \) and \(\psi\)).

Since we assume that \(\varepsilon_t\) has zero mean and unit covariance matrix, we obtain from (2.3) that

$$0 = \mathbb{E}[\varepsilon_t] = \mu_\varepsilon + \mu_\zeta T \gamma \quad \Leftrightarrow \quad \mu_\varepsilon = -\mu_\zeta T \gamma,$$

(2.4)
and

\[ I_k = \mathbb{E}[\varepsilon_t' \varepsilon_t'] = T \left( \mu_\zeta I + \sigma_\zeta^2 \gamma' \right) T' \ \iff \ \left( T'T \right)^{-1} = \mu_\zeta I + \sigma_\zeta^2 \gamma', \tag{2.5} \]

where \( \mu_\zeta \) and \( \sigma_\zeta^2 \) denote the mean and variance of \( \zeta_t \), respectively. The mean and variance of \( \varepsilon_t \) exist if the mean and variance of \( \zeta_t \) exist, respectively. The density of \( y_t \) for our specification of \( \gamma_t \) in (2.3) is given in the appendix.

We let the variances and correlations for \( y_t \) be time-varying by assuming that both \( D_t \) and \( R_t \) in (2.2) depend on a time-varying parameter \( f_t \), such that \( D_t = D(f_t) \) and \( R_t = R(f_t) \). This accommodates a setting where correlations and volatilities have their own dynamics, as well as a setting where correlations and volatilities are driven by a smaller set of time-varying common factors such as in the factor GARCH literature.

We model the dynamics of \( f_t \) using the framework of Creal, Koopman, and Lucas (2011, 2012). Their updating equation for the time-varying factor \( f_t \) is given by

\[ f_{t+1} = \sum_{i=0}^{p-1} A_i s_{t-i} + \sum_{j=0}^{q-1} B_j f_{t-j}, \tag{2.6} \]

where matrices \( A_i \) and \( B_j \), with appropriate dimensions, depend on a static parameter vector \( \theta \), that is \( A_i = A_i(\theta) \) and \( B_j = B_j(\theta) \). The innovation variable in (2.6) is \( s_t \) and is specified as a function of current and past values of \( y_t \) and \( f_t \). For example, in the univariate case with \( f_t = D_t^2 \) and normally distributed \( y_t \), the model in (2.6) embeds the standard GARCH model by setting \( s_t = y_t^2 \). The simplicity of this choice for \( s_t \) is appealing, but it is generally hard to extend it to a natural candidate \( s_t \) in other, more complicated cases. For example, in our current setting we want (2.6) to account for the possibly fat-tailed and skewed nature of the GH distribution, as well as for the adopted parameterization \( D(f_t) \) and \( R(f_t) \).

Creal et al. (2011, 2012) demonstrate that a good choice for \( s_t \) in a general non-linear time series context is the scaled density score, as given by

\[ s_t = S_t \nabla_t, \tag{2.7} \]

\[ \nabla_t = \partial \ln p(y_t | \mathcal{F}_{t-1}; f_t, \theta) / \partial f_t, \tag{2.8} \]

where \( S_t \) is an \( \mathcal{F}_{t-1} \)-adapted scaling matrix, and \( \mathcal{F}_t = \{ y_t, \ldots, y_1 \} \). By using the density score, the time-varying parameters are changed in the direction that increases the model’s
local fit as measured by the log-density. For our standard GH distributed $\varepsilon_t$ in (2.3), we rely on the following result.

**Result 1** Let $\otimes$ denote the Kronecker product, and let vec$(\cdot)$ denote the operator that stacks the columns of a matrix into a column vector. If $\varepsilon_t$ is modeled as in (2.3) with zero mean and unit covariance matrix, we have

$$
\nabla_t = \Psi'_t H'_t \left( w_t (y_t \otimes y_t) - \text{vec}(\tilde{\Sigma}_t) - (1 - w_t \mu_t)(y_t \otimes \tilde{L}_t \gamma) \right),
$$

(2.9)

where $\Psi_t = \partial \text{vech}(\Sigma_t)/\partial f'_t$, $\tilde{L}_t = L_t T$, $\tilde{\Sigma}_t = \tilde{L}_t \tilde{L}'_t$, $w_t$ is a scalar weight, and $H_t$ is a $k^2 \times k^2$ matrix. We define $w_t$ in (2.28) and $H_t$ in (2.31) in the appendix.

Our current model generalizes some of the well-known univariate and multivariate GARCH models. If $\varepsilon_t$ is normally distributed, i.e., $\gamma = 0$ and $T = I_k$, the weight $w_t$ reduces to $w_t = 1$ and equation (2.9) reduces to $\nabla_t = \Psi'_t H'_t \text{vec}(y_t y'_t - \Sigma_t)$. This is the usual expression for a multivariate GARCH model for time-varying volatilities and correlations. The matrix $H_t$ captures the relation between $\tilde{L}_t$ and $\Sigma_t$. The matrix $\Psi_t$ is determined by the parameterization of $D(f_t)$ and $R(f_t)$ in terms of the time-varying parameter vector $f_t$.

There are two interesting differences between a standard multivariate GARCH model, that is driven by squares and cross-products of vector $y_{t-1}$, and our model, that is driven by the score function (2.9) of the GH distribution. The first difference is the presence of the weighting factor $w_t$, which is fully defined in (2.28) in the Appendix. The second difference is the presence of the asymmetry term $y_t \otimes \gamma$. These differences are the result of the fat-tailedness and skewness of the distribution of $y_t$, respectively. For the case of a symmetric Student’s $t$ distribution, Creal et al. (2011) also obtain a weight effect but not the asymmetry term.

As shown in Appendix 2.7, the weight $w_t$ is generally a decreasing function of $d^*_t$, where

$$
d^*_t = \chi + x'_t x_t, \quad x_t = \tilde{L}^{-1}_t y_t + \mu \gamma,
$$

for fat-tailed distributions in the GH class, where $x_t$ is the standardized version of the original observation $y_t$. As a result, the impact of lagged (cross)-products in $y_t \otimes y_t$ on future values of $f_t$ (and thus on volatilities and correlations) is mitigated by $w_t$, if $y_t$ is large.
in the sense that $d_{x_t}$ is large. The intuition is as follows. If $y_t$ is drawn from a fat-tailed distribution, large values of $y_t$ are not necessarily due to local volatility or correlation increases. Instead, large $y_t$’s may be due to the fat-tailed nature of the distribution. The dynamics of $f_t$ (volatilities and correlations) should therefore only partly reflect the large value of $y_t$. The remainder is then attributed to the fat-tailed nature of the distribution and should not affect the volatility and correlation dynamics.

The second difference in (2.9) is the asymmetry term. The term takes a different role than the usual leverage effect in volatility models, which captures increases in volatilities if recent returns have been negative. Such a leverage effect can still be included in (2.9) in the usual way. Our asymmetry term $y_t \otimes \gamma$ is due to the skewness of the distribution. If, for example, $y_t$ is univariate and right-skewed ($\gamma > 0$), a large positive value of $y_t$ is more likely and is not necessarily attributable to a local volatility increase. However, a large negative value of $y_t$ should be taken as a very strong signal of a volatility increase, because large negative observations are extremely unlikely for a right-skewed distribution (unless the volatility has increased). This is precisely the effect of the asymmetry term $y_t \otimes \gamma$ in (2.9): for a right skew ($\gamma > 0$), the term mitigates the volatility increase if $y_t > 0$, and reinforces the volatility increase if $y_t < 0$.

Since both the shape and the parameterization of the distribution affect the dynamics of $f_t$, our current model is clearly different from the GARCH class of models with non-Gaussian observations. For the GARCH class of models, the non-Gaussian assumption only affects the likelihood function; it does not affect the dynamic behavior of $f_t$. In our framework, the distributional properties of $y_t$ affect both the likelihood and the dynamic evolution of $f_t$ at the same time.

Our model retains many of the convenient properties of the GARCH and DCC type models. For example, $s_t$ is $\mathcal{F}_t$-adapted and therefore parameter estimation in model (2.7)–(2.9) is carried out in the same convenient way as in GARCH models. Indeed, our likelihood function can be expressed in closed-form via the prediction error decomposition and the basic recursion (2.6). It leads to fast likelihood evaluation. The interpretation of the model is also intuitive. Depending on the choice of the scaling matrix $S_t$, the driver $s_t$ can be interpreted as a local Gauss-Newton or Steepest-Ascent improvement to the likelihood at time $t$. The score of the observation density at time $t$, evaluated at the current estimate $f_t$ of the time-varying parameter, determines in what direction $f_t$ is best
updated to improve the fit of the model. The additional lags and dynamics in (2.6) add further flexibility to the size and speed of these adjustments as time progresses.

We collect all static parameters of the model, such as $\gamma$, $\mu_\xi$, $\sigma_\xi^2$, $A_1, \ldots, A_p$, $B_1, \ldots, B_q$, into the vector $\theta$. The parameter vector $\theta$ is estimated by the method of Maximum Likelihood (ML). Inference on $\theta$ is carried out in the usual way by taking the negative inverse Hessian of the log likelihood function at the optimum as the covariance matrix of the ML estimator.

2.3 Model parameterizations

The GH distribution has a considerable number of parameters from which a selection cannot be identified simultaneously. In particular, $\chi$ and $\psi$ are not separately identified; only their product $\chi\psi$ is identified. Identification can be achieved in several ways. For example, we can set $|\Sigma_t|$ to a fixed constant, say unity, such that $\Sigma_t$ is normalized. Alternatively, we can simply fix $\chi$ or $\psi$ and estimate the other parameter in an unrestricted way. In our implementation, we estimate $\kappa = (\chi\psi)^{1/2}$ and extract $\chi$ and $\psi$ separately through the identifying assumption $\mu_\xi = 1$. This normalization turns out to be particularly useful when estimating the GH model using the Expectation Maximization (EM) algorithm of Section 2.4. Given the identifying restriction $\mu_\xi = 1$, we can obtain $\chi$ and $\psi$ for a fixed value of $\kappa$ by the equality

$$1 = \mu_\xi = \frac{\sqrt{\chi\psi}K_{\lambda+1}(\sqrt{\chi\psi})}{\psi K_{\lambda}(\sqrt{\chi\psi})} \Leftrightarrow \psi = \frac{\kappa \cdot K_{\lambda+1}(\kappa)}{K_{\lambda}(\kappa)}, \quad (2.10)$$

with $\chi = \kappa^2/\psi$ and $K_\lambda(\cdot)$ is the modified Bessel function of the second kind.

Following Creal et al. (2011), we can consider two obvious choices for the parameterization of both the diagonal matrix of variances $D_t^2$ and the correlation matrix $R_t$ in (2.2). We can take the variances $\text{diag}(D_t^2)$ themselves or the log-variances $\ln(\text{diag}(D_t^2))$ as parameters. The advantage of taking log-variances as parameters is that the resulting variances are always positive. When the variances themselves are taken as parameters, we need to impose restrictions on the coefficient matrices $A_i$ and $B_j$ in (2.6) to ensure positive variances at all times. In higher dimensional models with more lags in the updating equation (2.6), such restrictions become rather complicated. We therefore take
log-variances as parameters.

The specification of the correlation matrix \( R_t \) is subject to the constraints that \( R_t \) is a positive definite matrix with diagonal elements equal to one, for all \( t \). A possible parameterization of \( R_t \) is similar to the DCC model of Engle (2002). Let \( Q_t = Q(f_t) \) be an auxiliary time-varying parameter matrix, and set

\[
R_t = \Delta_t^{-1} Q_t \Delta_t^{-1},
\]

(2.11)

where \( \Delta_t^2 \) is a diagonal matrix holding the diagonal elements of \( Q_t \). The matrix \( Q_t \) has \( k \) redundant elements compared to the correlation matrix \( R_t \). As a result, only \( k(k-1)/2 \) independent signals in \( \nabla_t \) are distributed over the \( k^2 \) elements of \( Q_t \). The details of this specification and its implication for \( \Psi_t \) in (2.9) are presented in Creal et al. (2011).

An alternative specification for the correlation matrix is given by the hypersphere transformation as adopted by, for example, Jäckel and Rebonato (1999), van der Weide (2002), and Creal et al. (2011). The correlation matrix is obtained from the Choleski decomposition \( R_t = X_t'X_t \) where \( X_t \) is a upper triangular \( k \times k \) matrix that is constructed from a set of \( k(k-1)/2 \) time-varying angles \( \phi_{ijt} \) in \([0, \pi]\) and is given by

\[
X_t = \begin{pmatrix}
1 & c_{12t} & c_{13t} & \cdots & c_{1kt} \\
0 & s_{12t} & c_{23t}s_{13t} & \cdots & c_{2kt}s_{1kt} \\
0 & 0 & s_{23t}s_{13t} & \cdots & c_{3kt}s_{2kt}s_{1kt} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & c_{kkt}s_{kkt}s_{kkt} \\
0 & 0 & 0 & \cdots & \prod_{\ell=1}^{k-1} s_{\ell kt}
\end{pmatrix},
\]

(2.12)

with \( c_{ijt} = \cos(\phi_{ijt}) \) and \( s_{ijt} = \sin(\phi_{ijt}) \). For the 2-dimensional case, we have the Choleski and correlation matrices given by

\[
X_t = \begin{pmatrix}
1 & \cos(\phi_{12,t}) \\
0 & \sin(\phi_{12,t})
\end{pmatrix}, \quad R_t = X_t'X_t = \begin{pmatrix}
1 & \cos(\phi_{12,t}) \\
\cos(\phi_{12,t}) & 1
\end{pmatrix},
\]

(2.13)

with the correlation given by \( \cos(\phi_{12,t}) \). The second column of \( X_t \) in (2.13) expresses a two-dimensional unit-length vector in terms of its polar rather than its Cartesian coordinates.
The generalization to the $k$-dimensional setting is given by the $k$th column of $X_t$ in (3.15).

The number of unknown coefficients in $X_t$ equals the number of correlations in the matrix $R_t$ such that there are no redundancies as in the specification (2.11). We collect all angles $\phi_{ijt}$ in the vector $\phi_t$ which is specified as a function of $f_t$. For any value of $\phi_t$, the matrix $R_t = X_t'X_t$ satisfies the properties of a correlation matrix. The specification of $\Psi_t$ in (2.9) when using the hypersphere parameterization of $R_t$ is provided in Creal et al. (2011).

The definition of $s_t$ in (2.6) is completed by the choice of a scaling matrix $S_t$. Creal et al. (2012) discuss a number of possible choices, all of which are based on the local curvature of the model density at time $t$ via the (local) information matrix. Computing the information matrix for the general GH distribution, however, is analytically intractable. Therefore, we consider the computationally feasible alternative by setting the scaling matrix equal to the inverse information matrix for the symmetric Student’s $t$ distribution as a special member of the GH class. This information matrix is known analytically and is derived in Creal et al. (2011). This choice accommodates both the possible fat-tailed nature of the distribution and the time-variation in the volatilities and correlations. The form of scaling can be implemented efficiently and has shown to work well for both simulated and empirical data, see also the results in the following sections. Finally, this choice also makes our current model directly comparable to the familiar multivariate GARCH models if the distribution is Gaussian.

An interesting final feature of our model is that one can easily impose a factor structure on the volatilities and correlations. This can be done by picking the dimension of $f_t$ to be lower than the number of elements in $\text{diag}(D_t)$ and $Q_t$ or $X_t$. This approach can be used if the same factors drive more correlations, or if correlations and volatilities are driven by the same factors. The model allows the dynamic factors $f_t$ to adapt automatically via the specification of $\Psi_t$ in (2.9). Through the score of the density function, our framework naturally weights and combines the different sources of information in $y_t$ to improve the current estimates of volatilities and correlations.
2.4 Time-varying scale matrix and an EM algorithm

The time-varying covariance matrix $\Sigma_t = L_t L_t'$ is specified by means of the factor $f_t$ which is modeled by (2.6) – (2.9). We assume that the variance of the multivariate GH distribution exists and therefore we must constrain the fat-tailedness of the mixing variable $\zeta_t$. For example, in the case of a skewed Student’s $t$ distribution, we require the degrees of freedom parameter to be higher than 4, rather than the usual 2 for the symmetric case. This constraint may not be realistic for financial data, especially returns on individual equities that have many jumps and outliers. As an alternative, we can specify the time-varying scaling matrix $\hat{\Sigma}_t$ in (2.9) rather than the time-varying covariance matrix $\Sigma_t$. Moment restrictions are then no longer needed since the scaling matrix $\hat{\Sigma}_t$ always exists.

The GH distribution relies on many parameters. This can complicate parameter estimation, particularly when the dimension of $y_t$ is high; see the discussion in Hu (2005). This is one of the reasons why maximum likelihood estimation for the GH distribution is usually carried out by the Expectation-Maximization (EM) algorithm of Dempster et al. (1977). A basic introduction of the EM algorithm for the GH distribution with a time-invariant covariance matrix is provided in McNeil et al. (2005). Parameter estimation for a multivariate GARCH model with a GH distribution is considered by Hu (2005). A key simplification in the EM algorithm is that the parameters for the mixing distribution can be separated from the location, skewness, and scale parameters. This convenient property does not hold for the model specification with the covariance matrix $\Sigma_t$. However, if we consider the model specification in terms of the scale matrix $\hat{\Sigma}_t$, we are able to develop a newly modified EM algorithm for estimation. The usual advantages of EM estimation then again apply to our setting of a GH distribution with time-varying parameters.

First, we reformulate the model in terms of the scaling matrix $\hat{\Sigma}_t = \hat{L}_t \hat{L}_t'$. Second, we develop the modified EM algorithm for estimating the static parameter vector $\theta$. The mean-variance normal mixture model for the observations $y_t$ using the square root scaling matrix $\hat{L}_t$ is given by

$$y_t = \mu_{ty} + \zeta_t \hat{L}_t' \gamma + \sqrt{\hat{\Sigma}_t} z_t,$$

with $\mu_{ty} = -\mu_z \hat{L}_t' \gamma$. This specification follows from (2.3). Since $\hat{\Sigma}_t = \hat{L}_t \hat{L}_t'$ is a covariance matrix for the normal variable in the mixture specification (2.3), we can use similar matrices as developed in the previous section, that is $\hat{\Sigma}_t = \hat{D}_t \hat{R}_t \hat{D}_t$ and $\hat{\Psi}_t = \partial \text{vech}(\hat{\Sigma}_t)/\partial f_t'$. 
2.4. TIME-VARYING SCALE MATRIX AND AN EM ALGORITHM

In the implementation of the EM algorithm for the GH distribution, as proposed by McNeil et al. (2005), estimation of parameters governing the mixing variable specification (2.3) can be separated from estimation of the other parameters. The main difficulty in our current context is the dynamic process for \( f_t \) that is driven by the scaled score \( s_t \) of the GH distribution and depends on the parameters of the mixing variable. It appears difficult to split the parameter vector and to reduce a high-dimensional likelihood optimization into two lower dimensional optimization problems. Our modification of the EM algorithm, however, circumvents this problem on the basis of Result 2.

**Result 2** We can express the score function of the conditional observation density by

\[
\nabla_t = \frac{\partial \ln p(y_t | f_t, \mathcal{F}_{t-1}; \theta)}{\partial f_t} = E \left[ \frac{\partial \ln p(y_t | \zeta_t, f_t, \mathcal{F}_{t-1}; \theta)}{\partial f_t} \bigg| \mathcal{F}_t \right]. 
\]

(2.15)

The result enables us to partition the parameter vector as \( \theta' = (\theta_1', \theta_2') \) where \( \theta_2 \) contains the parameters associated with the distribution of the mixing variable \( \zeta_t \), in particular \( \lambda, \chi, \) and \( \psi \). The remaining parameters are collected in \( \theta_1 \). We define the joint log likelihood of the observation \( y_t \) and the unobserved mixing variable \( \zeta_t \) as

\[
\sum_{t=1}^{n} \ln p(y_t, \zeta_t | f_t, \mathcal{F}_{t-1}; \theta) = \mathcal{L}_{1n}(\theta) + \mathcal{L}_{2n}(\theta_2),
\]

(2.16)

with

\[
\mathcal{L}_{1n}(\theta) = \sum_{t=1}^{n} \ln p(y_t | \zeta_t, f_t, \mathcal{F}_{t-1}; \theta), \quad \mathcal{L}_{2n}(\theta_2) = \sum_{t=1}^{n} \ln p(\zeta_t; \theta_2),
\]

(2.17)

where the conditional density \( p(y_t | \zeta_t, f_t, \mathcal{F}_{t-1}; \theta) \) is Gaussian and the marginal density \( p(\zeta_t; \theta_2) \) is Generalized Inverse Gaussian (GIG) denoted by \( N^{-}(\lambda, \chi, \psi) \). For the implementation of the E-step in the EM algorithm, we define

\[
Q_1(\theta, \hat{\theta}) = \int \ldots \int \mathcal{L}_{1n}(\theta) \left( \prod_{t=1}^{n} p(\zeta_t | y_t, \mathcal{F}_{t-1}; \hat{\theta}) \right) \, d\zeta_n \ldots d\zeta_1 = E_{\hat{\theta}} [ \mathcal{L}_{1n}(\theta) | \mathcal{F}_n ],
\]

(2.18)

and, similarly,

\[
Q_2(\theta_2, \hat{\theta}) = E_{\hat{\theta}} [ \mathcal{L}_{2n}(\theta_2) | \mathcal{F}_n ].
\]

(2.19)

In Appendix 2.7 we show that under the normalization constraint \( \mu_\zeta = 1 \), \( Q_1(\theta, \hat{\theta}) \) depends on \( \theta_1 \) only. Consequently, we write \( Q_i(\theta, \hat{\theta}) \) for \( i = 1, 2 \), with a slight abuse of notation.
The EM algorithm for parameter estimation is given follows.

**Modified EM algorithm for a the dynamic GH model for the scale matrix**

1. Start with an initial guess of the parameters, $\hat{\theta}^{(0)}$, and set $\ell = 0$.

2. Given a trial value of the parameters $\hat{\theta}^{(\ell)}$, define the modified transition equation for the scaling matrix as

$$f_{t+1} = \sum_{i=0}^{p-1} A_i \tilde{s}_{t-i} + \sum_{j=0}^{q-1} B_j f_{t-j}, \quad (2.20)$$

where $\tilde{s}_t = S_t \tilde{\nabla}_t$, and

$$\tilde{\nabla}_t^{(\ell)} = E_{\hat{\theta}^{(\ell)}} \left[ \partial p(y_t | \zeta_t, f_t, F_{t-1}; \theta) / \partial f_t \mid F_t \right], \quad (2.21)$$

with $\tilde{\nabla}_t^{(\ell)}$ fully specified in the appendix.

3. Given the modified dynamics, compute $Q_1(\theta_1, \hat{\theta}^{(\ell)})$ and maximize it numerically with respect to $\theta_1$. The maximum is obtained at $\tilde{\theta}_1$.

4. Update $\hat{\theta}^{(\ell)}$ to $\tilde{\theta}^{(\ell)} = (\tilde{\theta}_1^{(\ell)}, (\tilde{\theta}_2^{(\ell)})')$, compute $Q_2(\theta_2, \tilde{\theta}^{(\ell)})$ and maximize it numerically with respect to $\theta_2$. The maximum is obtained at $\tilde{\theta}_2$.

5. Update $\hat{\theta}^{(\ell)}$ to $\hat{\theta}^{(\ell+1)} = (\tilde{\theta}_1^{(\ell)}, \tilde{\theta}_2^{(\ell)})'$, increase $\ell$ by one, and iterate steps 2–5 until convergence.

Steps 3–5 are standard for the GH-EM algorithm; see, for example, McNeil et al. (2005). The E-step is developed in Appendix 2.7. An important feature of our modified EM algorithm is that the optimization can still be split into two lower dimensional problems in steps 3 and 4, even though we have a GH model with time-varying parameters governed by complex dynamics. The key to this result is that step 3 of the algorithm is effectively based on fitting a standard multivariate Gaussian GARCH model with the updating equation (2.20). The crucial part that enables this is our modification to the standard EM algorithm in step 2. In this step, the updating equation that depends on $\theta_2$ only via the score $\nabla_t$ is replaced by a simple equation that does not depend on $\theta_2$. The intuition follows from Result 2. In the same way as in the E-step of the EM algorithm, the
score function is replaced by a conditional expectation of a score function that depends on parameter values from the previous iteration, that is \( \hat{\theta}^{(t)} \). As this score function is conditional on \( \zeta_t \), it is the score of a Gaussian density and therefore takes a very simple form. It follows that as the EM iterations converge to the ML estimates, the score \( \tilde{\nabla}_t^{(t)} \) in the EM algorithm converges to the score \( \nabla_t \) of the full GH distribution via (2.15). This is confirmed by numerical experiments, where the ML parameter estimates are obtained by the modified EM algorithm and by directly maximizing the likelihood.

2.5 Monte Carlo evidence

To study the behavior of the new model, we carry out a Monte Carlo study. In the next section, we investigate the model’s performance in an empirical study. In both settings, we benchmark the model’s performance to the well-known DCC model. The simulations test the accuracy of the different models in estimating correlation patterns, similar to the experiments in Engle (2002). We describe the set-up in Subsection 2.5.1 and present the results in Subsection 2.5.2.

2.5.1 Simulation design

The design of our Monte Carlo experiments are similar to the original experiments for correlations as described in Engle (2002). We take the same deterministic functions as in Engle’s paper, namely

1. Constant: \( f_t = 0.9 \),
2. Sine: \( f_t = 0.5 + 0.4 \cos(2\pi t/200) \),
3. Fast Sine: \( f_t = 0.5 + 0.4 \cos(2\pi t/20) \),
4. Step: \( f_t = 0.9 - 0.5(t > 500) \),
5. Ramp: \( f_t = \text{mod} \ (t/200) \).

This allows us to study the properties of competing statistical models under a range of correlation dynamics, such as slow and fast oscillations, and structural breaks.

The simulation experiment concentrates on recovering dynamic correlation patterns. We consider a bivariate series \( y_t \) with zero mean and unit variances, such that we can fully concentrate on the correlations. To limit the number of parameters, we concentrate on a
particular subclass of the GH family, namely the GH Skewed $t$ (GHST) distribution. Using the five deterministic patterns for correlations described above, we generate bivariate time series $y_t$ as

$$y_t \sim GHST(0, D_t R_t D_t, \gamma, \nu), \quad D_t = I_2, \quad R_t = \begin{pmatrix} 1 & f_t \\ f_t & 1 \end{pmatrix}. \quad (2.22)$$

Given the five different correlation patterns, we consider three different GHST distributions in our experiments. The GHST distribution contains the symmetric Student’s $t$ and the normal distribution as special cases. In particular, the GHST collapses to the symmetric Student’s $t$ distribution if the skewness parameter $\gamma$ goes to zero. It further reduces to the normal distribution if the degrees of freedom parameter $\nu$ goes to infinity. As a benchmark, we start with the normal distribution. Then we introduce moderate kurtosis by setting $\nu = 5$. Finally, we introduce mild skewness by setting $\gamma = (-0.03, -0.03)'$.

In the experiment, we take the DCC model of Engle (2002) as our benchmark. Again, for each simulated DGP we use the correct class of distributions when computing the likelihood. We like to emphasize that we are not aware of an earlier application that considers a DCC model with GHST or GH distributed error terms. The DCC models are compared to our new model with a diagonal structure for the $3 \times 3$ matrices $A_1$ and $B_1$. To model the correlation, we use the hypersphere parameterization. The performance of the different statistical models is measured using the Mean Absolute Error (MAE) based on the difference between the estimated correlation and its true value. The MAEs are averaged across time and across simulations. We generate samples of size $T = 1,100$, discarding the first 100 observations to avoid dependence on initial conditions, and use 100 Monte Carlo replications.

### 2.5.2 Simulation results

Table 2.1 contains the results for our experiment. For the normal distribution, the performance of both models is roughly the same. There appears to be no noticeable loss in efficiency in this case of using the over-parameterized GHST distribution in the new model. Again we note that as the error distribution becomes more complex, the MAEs of the DCC model increases, whereas the MAEs of the new model remain rather stable.
2.6. EMPIRICAL APPLICATION

We also see that the new model outperforms the DCC in four out of five cases for the fat-tailed and skewed DGPs. The improved performance is mainly due to the weighting function and asymmetry effect in the updating equations (2.6) and (2.9) for the factor \( f_t \). Due to this weighting incidentally large observations result in less distortions for the estimated correlation dynamics.

Table 2.1: Mean Absolute Errors for Correlation Estimates
The table presents the average Mean Absolute Error (MAE) over 100 Monte Carlo replications and 1,000 time series observations for the correlation estimates of three different distributions (in pairs of columns) and five different correlation patterns. The distributions used are the normal, Student’s \( t(5) \), and GHST(0, \( \Sigma_t \), 0.03, \( \nu \)). The boldface numbers show the models with the smallest MAE for a given DGP.

<table>
<thead>
<tr>
<th>Dynamic Correlations</th>
<th>normal model (2.6)–(2.9)</th>
<th>DCC model (2.6)–(2.9)</th>
<th>GHST model (2.6)–(2.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.004</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>Sine</td>
<td>0.135</td>
<td>0.139</td>
<td>0.151</td>
</tr>
<tr>
<td>Fast Sine</td>
<td>0.225</td>
<td>0.255</td>
<td>0.254</td>
</tr>
<tr>
<td>Step</td>
<td>0.066</td>
<td>0.068</td>
<td>0.094</td>
</tr>
<tr>
<td>Ramp</td>
<td>0.159</td>
<td>0.165</td>
<td>0.168</td>
</tr>
</tbody>
</table>

2.6 Empirical application

In an empirical study, we examine the correlations in a multivariate dataset with four blue-chip stocks from different industries: Coca-Cola, IBM, Merck and J.P. Morgan. All four stocks are part of the Dow Jones 30 index. We use daily log returns from January 1989 to December 2009 from CRSP. The final dataset contains 5295 daily observations. Descriptive statistics are provided in Table 2.2. It is clear that the series exhibit significant excess kurtosis and skewness, warranting the use of the GH distribution.
Table 2.2: Data descriptive statistics.
The descriptive statistics for the CRSP stock returns between January 1989 and December 2009. All observations are daily log returns. All four stocks are part of the Dow Jones 30 composite index. All skewness and excess kurtosis statistics have p-values below $10^{-4}$.

<table>
<thead>
<tr>
<th></th>
<th>Coca-Cola</th>
<th>IBM</th>
<th>Merck</th>
<th>JP Morgan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean $\times 10^4$</td>
<td>6.33</td>
<td>5.27</td>
<td>5.69</td>
<td>7.56</td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Standard Deviation $\times 100$</td>
<td>1.56</td>
<td>1.89</td>
<td>1.82</td>
<td>2.62</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.23</td>
<td>0.29</td>
<td>-0.12</td>
<td>0.74</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>2.05</td>
<td>3.89</td>
<td>3.00</td>
<td>8.33</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.10</td>
<td>-0.16</td>
<td>-0.15</td>
<td>-0.21</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.14</td>
<td>0.13</td>
<td>0.13</td>
<td>0.25</td>
</tr>
</tbody>
</table>

To estimate the volatilities and correlations, we use the DCC model with a Gaussian and GHST error distribution as our benchmark. We also implement our own model using the GHST, the GH Variance-Gamma (GHVG, with $\psi = 2\lambda$ and $\chi = 0$), and the GH error distribution. Our model has ten factors: four volatilities, and six correlations. We estimated the model using both the DCC and hypersphere parameterization for the correlation matrix. The estimation results for the dynamic parameters were similar, so we only report the results obtained under the DCC parameterization to maximize comparability with the DCC model. We use $p = q = 1$ in (2.6) and impose the same parsimony as in the DCC model. This means that we use diagonal matrices $A_1$ and $B_1$ in (2.6), and that the diagonal elements corresponding to the correlation equations have the same value.

The estimation results are presented in Table 2.3. The parameters governing the dynamics are statistically significant for all models. For the DCC model with a normal distribution, the persistence parameters for the volatilities ($A + B$) are high. All the standard stationarity conditions are satisfied. Changing the specification to a DCC model with a GHST distribution has several effects. First, the likelihood increases by more than 1,200 points by adding only five parameters. The GHST distribution, therefore, provides a much better fit to the data. Second, the volatilities of the first two stocks (Coca Cola, IBM) are less affected by lagged squared errors. This can be seen from the reduced values for the $A$ coefficients. By contrast, the persistence of the volatility dynamics of Merck ($B_{d3}$) goes up substantially. This is due to some highly influential observations during the sample period for this stock.
Table 2.3: Empirical Estimation Results

Empirical results based on stock return data between January 1989 and December 2009 for Coca Cola, IBM, Merck, and JP Morgan. The DCC model is defined as in Engle (2002) and uses a normal or GH skewed $t$ (GHST) likelihood. The DGH model uses the GH, GHST, and GH Variance Gamma (GHVG) distribution for the likelihood and the parameter dynamics. Intercepts are not reported to save space. $A_{d1}$ to $A_{d4}$ and $B_{d1}$ to $B_{d4}$ contain the diagonal elements of $A_1$ and $B_1$ from (2.6) corresponding to the volatilities, and $A_\rho$ and $B_\rho$ the parameter corresponding to the correlations. $\gamma_i$ is the skewness parameter for series $i$ (1: Coca Cola, 2: IBM, 3: Merck, 4: JP Morgan), $\kappa = (\chi \psi)^{1/2}$, with $\chi$, $\psi$ and $\lambda$ the GH parameters. For the GHST, we report $\gamma = -2\lambda$.

<table>
<thead>
<tr>
<th></th>
<th>DCC</th>
<th></th>
<th>DGH</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gaussian</td>
<td>GHST</td>
<td>GHST</td>
<td>GHVG</td>
</tr>
<tr>
<td></td>
<td>$A_{d1}$</td>
<td>0.037$a$</td>
<td>0.029$a$</td>
<td>0.032$a$</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.017)</td>
</tr>
<tr>
<td></td>
<td>$A_{d2}$</td>
<td>0.035$a$</td>
<td>0.026$a$</td>
<td>0.034$a$</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.016)</td>
</tr>
<tr>
<td></td>
<td>$A_{d3}$</td>
<td>0.038$a$</td>
<td>0.030$a$</td>
<td>0.038$a$</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.020)</td>
</tr>
<tr>
<td></td>
<td>$A_{d4}$</td>
<td>0.057$a$</td>
<td>0.053$a$</td>
<td>0.053$a$</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.007)</td>
<td>(0.024)</td>
</tr>
<tr>
<td></td>
<td>$A_\rho$</td>
<td>0.010$a$</td>
<td>0.010$a$</td>
<td>0.010$a$</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.005)</td>
</tr>
<tr>
<td></td>
<td>$B_{d1}$</td>
<td>0.959$a$</td>
<td>0.969$a$</td>
<td>0.996$a$</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td>$B_{d2}$</td>
<td>0.959$a$</td>
<td>0.969$a$</td>
<td>0.994$a$</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>$B_{d3}$</td>
<td>0.913$a$</td>
<td>0.956$a$</td>
<td>0.989$a$</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td>$B_{d4}$</td>
<td>0.939$a$</td>
<td>0.944$a$</td>
<td>0.994$a$</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>$B_\rho$</td>
<td>0.986$a$</td>
<td>0.985$a$</td>
<td>0.996$a$</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.045</td>
<td>0.089$a$</td>
<td>0.131$a$</td>
<td>0.089$a$</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.029)</td>
<td>(0.038)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>0.026</td>
<td>0.074$a$</td>
<td>0.099$a$</td>
<td>0.074$a$</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.028)</td>
<td>(0.037)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>-0.038</td>
<td>-0.026</td>
<td>-0.025</td>
<td>-0.026</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.028)</td>
<td>(0.037)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>$\gamma_4$</td>
<td>0.056$a$</td>
<td>0.083$a$</td>
<td>0.126$a$</td>
<td>0.083$a$</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.028)</td>
<td>(0.038)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>6.434$a$</td>
<td>6.318$a$</td>
<td>3.738$a$</td>
<td>3.160$a$</td>
</tr>
<tr>
<td></td>
<td>(0.235)</td>
<td>(0.245)</td>
<td>(0.128)</td>
<td>(0.123)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td></td>
<td></td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.467)</td>
<td></td>
</tr>
<tr>
<td>$\kappa$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-lik</td>
<td>-39991</td>
<td>-38787</td>
<td>-38684</td>
<td>-38994</td>
</tr>
</tbody>
</table>
The skewness parameters in the DCC model with the GHST distribution are mostly insignificant. The only exception is the positive skewness for J.P. Morgan (stock 4). The signs of the skewness coefficients $\gamma_i$ in Table 2.3 are compatible with the descriptive statistics in Table 2.2. The degrees of freedom is estimated at 6.4 with a relatively small standard error.

For the new model with a GHST distribution, we see a further increase in the likelihood of more than 100 points. This increase in the likelihood is obtained without adding any parameters relative to the DCC model with a GHST distribution. The persistence parameters $B$ and the degrees of freedom $\nu$ are estimated at similar values as for the DCC-GHST model. Note that the $B$ parameters for the new model must be compared to the $A + B$ parameter of the DCC model. The primary reason for the increase in the likelihood is the effect of the fat-tailed and skewed GHST distribution on the volatility and correlation dynamics.

The effect of the altered dynamic specification under fat tails on the correlation dynamics can be clearly illustrated in Figure 2.1. The figure shows the estimated correlations of the different statistical models for the sub-sample 2002–2005 for the two pairs Coca Cola-IBM and Merck-J.P. Morgan. During this period, we note that several influential observations caused abrupt shifts in the estimated correlation levels based on the DCC model. For Coca Cola-IBM, clear examples of this behavior are seen at the end of the first quarter of 2002, mid 2003, mid 2004 and September 2004, and April 2005. For Merck-JP Morgan, similar patterns are observed around October 2003, November 2003, October 2004 (very clear), and February, March, and October 2005. During all these episodes, the evolution of the correlations for the new model is much more stable and in line with expectations that correlations should behave rather smoothly. We also note that the estimated dynamics of the correlations for the DCC models repeatedly take a long time to revert to their old pattern. For example, for Coca Cola-IBM, it takes roughly three months starting from the big drop in mid-2003 before the DCC model and the new model exhibit similar correlation levels. The same is true for mid-2004. This holds even more strongly for Merck-JP Morgan after October 2004, when Merck experienced a large incidental drop in its stock price after it announced a major worldwide withdrawal of its products. As seen from the DCC-GHST model, the use of the GHST model alone does not remedy the distortive impact of such an influential observation. To adequately cope
with such outliers, changes in the dynamic equation of the correlations are needed in addition to a fat-tailed observation density. Our current score driven modeling framework provides such corrections in a natural way.

The correlation differences between the DCC and score driven models over the entire sample and for all six pairs of stocks are presented in Figure 2.2. A positive value in the graph indicates that the DCC estimate of the correlation is lower than that of the score driven model DGH-GHST. We see that the differences in the correlations can be substantial at times. When the underlying observations causing the differences are extreme, the differences can persist for months and in some cases even years. In particular, we note that during the dotcom crash in 2000 the IBM correlations in the new model are estimated at a higher level than the estimates in the DCC models. Also, for the pairs involving J.P. Morgan, the correlations from the score driven models during the Financial Crisis are larger than for the DCC models. Such differences can have important implications for diversification, risk management, and asset allocation.

The skewness parameters for the new model with a GHST distribution are significant, except for Merck. The signs are in line with the descriptive statistics from Table 2.2. We note again that the skewness parameters also contribute to the different correlation dynamics via (2.6).

Figure 2.3 shows the volatility estimates of both models. The volatility patterns are at first sight much more in line between the two different volatility specifications. However, closer inspection shows that the same effect of large innovations affects the volatility dynamics. This is most clearly seen for Merck. For example, in October 2004 the stock price drops significantly for reasons explained earlier. This causes a large spike in the volatility estimate of the GARCH-DCC model, despite the use of a GHST error distribution. The spike in volatility only decreases very slowly to normal levels over a period of almost a year. The score driven model, by contrast, also shows an increase in volatility since October 2004, but on a much more modest and realistic scale. Though this is one of the most striking differences between the two models, there are many more. Particularly the IBM stock shows over the entire sample period various cases where the volatility as estimated by the DCC model first jumps and then gradually recedes to normal levels. This results in small reverse saw-tooth like patterns in the graph. The corresponding volatility dynamics for the score driven models do not exhibit such peculiar
Figure 2.1: Estimated correlation subsample for the DCC-GHST, and DGH-GHST models.

To show the difference of estimates for the DCC(1,1) and DGH(1,1) models, we plot the estimated correlations of these two models for the sub-sample 2002-2005. The correlation pairs shown in the figure are Coca Cola-IBM and Merck-J.P. Morgan.
To look into the difference of the DCC(1,1) and DGH(1,1) models, we plot the difference of correlation estimates under GHST distributions. It appears that the DCC differs from DGH model even under the same parametric assumption.
To conclude the empirical analysis, we also estimate a specification based on the GH Variance Gamma (GHVG) distribution, and on the general GH distribution. The GHVG has $\chi = 0$ and $\psi = 2\lambda$ and has a clear link to Lévy driven stochastic processes. The likelihood of the GHVG model is lower than that of the GHST model, and even lower than the DCC-GHST specification. This is confirmed by the model using the unrestricted GH distribution. Interestingly, the unconstrained GH estimates reveal that the GHST model is a good model for the data set at hand. We see that the parameter $\kappa = (\psi\chi)^{1/2}$ is very close to zero, and that the value of $\lambda$ is negative. For the GHST, we have $\lambda = -\nu/2$, and this is precisely the value that is estimated under the GH specification. We conclude that the GHST distribution provides sufficient flexibility to accommodate the current levels of fat-tailedness and skewness combined with correlation and volatility dynamics.

2.7 Conclusion

We have proposed a new time-varying conditional correlation model that accounts for skewness and fat tails through the use of the Generalized Hyperbolic (GH) distribution with time-varying parameters. The distinguishing feature of the model is that the shape of the observation distribution directly affects the mechanism by which volatilities and correlations are updated. The key mechanism for this is the use of the local density score to update volatilities and correlations. As a result, large observations are reweighted before they enter the updating equation. Because of this, the model is much less sensitive to outliers and incidental influential observations. The new model also includes a natural asymmetry term if the GH distribution is skewed.

We showed that the model is easy to estimate by standard maximum likelihood and Expectation-Maximization procedures. In a simulation experiment, we demonstrated that the model does a better job at estimating the unknown correlation dynamics than competing models if the error distribution is fat-tailed and skewed. When applied to real data, we showed that the model yields a more robust assessment of local volatility and correlation dynamics. Because the new model accounts for fat tails and skewness in the volatility and correlation dynamics, it is less affected by aberrant observations and therefore produces a clearer picture of actual volatilities and correlations.
2.7. CONCLUSION

The volatility estimates from the DCC under GHST distributions and DGH(1,1)-GHST with stock return data. From the graphs, we can see that the volatility from DGH-GHST is smoother than the DCC estimates.
CHAPTER 2. THE DYNAMIC VOLATILITIES & CORRELATION MODEL

Appendices to Chapter 2

Skewness of the GH distribution

Define \( \tilde{y}_t = \tilde{L}_t^{-1} y_t \) and \( m_{2\zeta} = E[(\zeta_t - \mu_\zeta)^2] \) for integer \( i \). Let \( e_i \) denote the \( i \)th column of \( I_k \). We obtain

\[
E[\tilde{y}_t] = 0, \quad (2.23)
\]

\[
E[\tilde{y}_t \tilde{y}_t'] = \mu_\zeta I + m_{2\zeta} \gamma' = (T' T)^{-1}, \quad (2.24)
\]

\[
E[\tilde{y}_t \otimes \tilde{y}_t \tilde{y}_t'] = m_{3\zeta} \gamma \otimes \gamma' + m_{2\zeta} \begin{pmatrix} \gamma_1 I_k + \gamma e_1' + e_1 \gamma' \\ \vdots \\ \gamma_k I_k + \gamma e_k' + e_k \gamma' \end{pmatrix}, \quad (2.25)
\]

such that the skewness of \( \tilde{y}_t \) only depends on \( \gamma \) and on the variance and skewness of the mixing variable \( \zeta_t \).

The Score of the GH distribution

Define the matrix \( \text{vec}(L) = D^0_k \text{vech}(L) \) for a lower triangular matrix \( L \). Note that \( D^0_k \) is different from the standard duplication matrix \( D_k \) for a symmetric matrix \( S \), i.e., \( \text{vec}(S) = D_k \text{vech}(S) \) with \( B_k = (D_k' D_k)^{-1} D_k' \). Also note that \( D^0_k D^0_k = I_k \), such that \( B^0_k = D^0_k \). Finally, let \( C_k \) be the commutation matrix, \( \text{vec}(S') = C_k \text{vec}(S) \) for an arbitrary matrix \( S \). For completeness, we mention that \( \tilde{L}_t = L_t T \), and \( \tilde{\Sigma}_t = \tilde{L}_t \tilde{L}_t' \).

An intermediate result is

\[
d\Sigma_t = d(L_t L_t') \quad \Leftrightarrow \quad \text{vec}(d\Sigma_t) = (I_k + C_k) (L_t \otimes I_k) \text{vec}(dL_t) \quad \Leftrightarrow \quad D_k \text{vech}(d\Sigma_t) = (I_k + C_k) (L_t \otimes I_k) D^0_k \text{vech}(dL_t) \quad \Leftrightarrow \quad \text{vech}(dL_t) = (B_k (I_k + C_k) (L_t \otimes I_k) D^0_k)^{-1} \text{vech}(d\Sigma_t). \quad (2.26)
\]

First define the standardized \( y_t \) as \( x_t = \tilde{L}_t^{-1} y_t + \mu_\zeta \gamma \). The random variable \( x_t \) has a GH distribution with location 0 and scaling matrix \( I_k \). Let \( d'z = \nu + z'z \) for a scalar \( \nu \) and a
vector $z$. With this notation, the density of the GH distribution of $y_t$ is given by

$$p_{GH}(y_t|f_t; \lambda, \chi, \psi, \mu, \Sigma_t) = \frac{e^{\gamma z_t} \left( \frac{d^\chi}{d^\gamma} \right)^{\lambda-k/2}}{2\pi L_t L_t^t} \cdot \frac{K_{\lambda-k/2}}{(\sqrt{\chi/\psi})^\lambda \cdot K_{\lambda}} \cdot \frac{d^\gamma}{d^\psi},$$

(2.27)

Let $k(\cdot) = \ln K(\cdot)$ with first derivative $k'(\cdot)$. Define the scalar weight

$$w_t = -\frac{\lambda - k/2}{d^x} - \frac{k'_{\lambda-k/2}}{d^\gamma} \cdot \frac{d^\gamma}{d^\psi}.$$

(2.28)

We obtain

$$\nabla_t = \frac{\partial \text{vech}(\Sigma_t)^t}{\partial f_t} \cdot \frac{\partial \text{vech}(\Sigma_t)}{\partial f_t} \cdot \frac{\partial \text{vech}(\tilde{\Sigma}_t)^t}{\partial f_t} \cdot \frac{\partial \ln p_{GH}(y_t|f_t)}{\partial f_t} = \Psi_t \bar{H}_t \frac{\partial \ln p_{GH}(y_t|f_t)}{\partial \text{vech}(f_t)},$$

with $\Psi_t = \partial \text{vech}(\Sigma_t)/\partial f_t$ and

$$\bar{H}_t = (T' \otimes I_k) D_k (B_k (I_k^2 + C_k) (L_t \otimes I_k) D_k^0)^{-1}$$

(2.29)

using the intermediate result (2.26).

Taking the derivative of the log-density with respect to vec$(\tilde{L}_t)$ and then via the chain rule with respect to $f_t$, we get

$$\frac{\partial \ln p_{GH}(y_t|f_t)}{\partial \text{vech}(f_t)} = \frac{\partial x_t'}{\partial \text{vech}(f_t)} \left( -0.5 \omega_t \frac{\partial d^x}{\partial x_t} + \gamma \right) - \text{vec}((\tilde{L}_t)^{-1})$$

$$= (\tilde{L}_t^{-1} y_t \otimes (\tilde{L}_t)^{-1})(w_t x_t - \gamma) - \text{vec}((\tilde{L}_t)^{-1})$$

$$= (\tilde{L}_t^{-1} \otimes (\tilde{L}_t)^{-1})(y_t \otimes 1)(w_t \tilde{L}_t^{-1} y_t + w_t \mu_t \gamma - \gamma) - \text{vec}((\tilde{L}_t)^{-1})$$

$$= (\tilde{L}_t' \otimes 1)(\tilde{x}_t^{-1} \otimes \tilde{x}_t^{-1}) \left( w_t y_t \otimes y_t - \text{vec}(\tilde{X}_t) - (1 - w_t \mu_t)(y_t \otimes \tilde{L}_t \gamma) \right).$$

(2.30)

The main result is now obtained by defining

$$H_t' = \bar{H}_t'(\tilde{L}_t' \otimes 1)(\tilde{x}_t^{-1} \otimes \tilde{x}_t^{-1}).$$

(2.31)
CHAPTER 2. THE DYNAMIC VOLATILITIES & CORRELATION MODEL

EM algorithm for time-varying scale matrix $\tilde{\Sigma}_t$

We first prove Result 2. It is easy to check that

$$
\nabla_t = \frac{\partial \ln p(y_t|f_t, \mathcal{F}_{t-1}; \theta)}{\partial f_t} = \frac{1}{p(y_t|f_t, \mathcal{F}_{t-1}; \theta)} \int \frac{\partial p(y_t, \zeta_t|f_t, \mathcal{F}_{t-1}; \theta)}{\partial f_t} \frac{p(\zeta_t; \theta_2)}{p(y_t|f_t, \mathcal{F}_{t-1}; \theta)} d\zeta_t
$$

$$
= \int \frac{\partial \ln p(y_t|\zeta_t, f_t, \mathcal{F}_{t-1}; \theta)}{\partial f_t} \frac{p(\zeta_t; \theta_2)}{p(y_t|f_t, \mathcal{F}_{t-1}; \theta)} d\zeta_t
$$

$$
= \mathbb{E} \left[ \frac{\partial \ln p(y_t|\zeta_t, f_t, \mathcal{F}_{t-1}; \theta)}{\partial f_t} \bigg| \mathcal{F}_t \right] = \nabla_t. \tag{2.32}
$$

Throughout, we impose the normalization constraint $\mu_\zeta = 1$. We note that

$$
\ln p(y_t|\zeta_t, f_t, \mathcal{F}_{t-1}; \theta) = -\frac{1}{2} \ln |\tilde{\Sigma}_t| - \frac{k}{2} \ln (\zeta_t) - \frac{k}{2} \ln (2\pi)
$$

$$
- \frac{1}{2\zeta_t} (y_t - (\zeta_t - \mu_\zeta) \tilde{L}_t) \tilde{\Sigma}_t^{-1} (y_t - (\zeta_t - \mu_\zeta) \tilde{L}_t), \tag{2.33}
$$

and

$$
\ln p(\zeta_t; \theta_2) = -\frac{\lambda}{2} \ln (\chi/\psi) - \ln (2) - \ln K_{\lambda} \left( \sqrt{\chi \psi} \right) - (\lambda - 1) \ln (\zeta_t) - \frac{1}{2} (\chi \zeta_t^{-1} + \psi \zeta_t), \tag{2.34}
$$

where $\tilde{L}_t = \tilde{L}(f_t)$ and $\tilde{\Sigma}_t = \tilde{\Sigma}(f_t)$, and where the mapping from $f_t$ to $\tilde{\Sigma}_t$ does not depend on $\theta_2$.

We define $\tilde{x}_t = \tilde{L}_t^{-1} y_t + \gamma$. From (2.38) and (2.34) and the properties of the Generalized Inverse Gaussian distribution (see Appendix A.2 of McNeil et al. (2005)), we get

$$
\delta_{1t}^{(f)} = \mathbb{E}_{\tilde{q}(t)} \left[ \zeta^{-1} \big| \mathcal{F}_n \right] = \left( \frac{d\tilde{x}_t}{d\gamma} \right)^{-1/2} K_{\lambda_1-k/2} \left( \sqrt{d\tilde{x}_t d\tilde{x}_t} \right) K_{\lambda_{1-k/2}} \left( \sqrt{d\tilde{x}_t d\tilde{x}_t} \right), \tag{2.35}
$$

$$
\delta_{2t}^{(f)} = \mathbb{E}_{\tilde{q}(t)} \left[ \zeta \big| \mathcal{F}_n \right] = \left( \frac{d\tilde{x}_t}{d\gamma} \right) K_{\lambda_1-k/2} \left( \sqrt{d\tilde{x}_t d\tilde{x}_t} \right) K_{\lambda_{1-k/2}} \left( \sqrt{d\tilde{x}_t d\tilde{x}_t} \right), \tag{2.36}
$$
2.7. APPENDICES TO CHAPTER 2

\[
\delta_{3t}^{(t)} = E_{\hat{\gamma}(t)} \left[ \ln(\zeta_t) \mid \mathcal{F}_n \right] = \frac{\partial}{\partial \xi} \left( \frac{d_{\xi t}^2}{d_{\gamma t}^2} \right)^{\xi/2} \left. \frac{K_{\lambda+\xi-k/2} \left( \sqrt{d_{\xi t}^2 d_{\gamma t}^2} \right)}{K_{\lambda-k/2} \left( \sqrt{d_{\xi t}^2 d_{\gamma t}^2} \right)} \right|_{\xi=0}, \tag{2.37}
\]

where \( d_t \) is defined below (2.26).

From (2.38) and using \( \mu_c = 1 \), we obtain

\[
\frac{\partial \ln p(y_t | \zeta_t, f_t, \mathcal{F}_{t-1}; \theta)}{\partial f_t} = \tilde{\Psi}_t' \tilde{H}_t' \text{vec} \left( \zeta_t^{-1} y_t (y_t + \tilde{L}_t \gamma)' - \tilde{\Sigma}_t \right) \tag{2.38}
\]

with \( \tilde{\Psi}_t = \partial \text{vech}(\tilde{\Sigma}_t) / \partial f_t' \) and

\[
\tilde{H}_t = \left( \tilde{\Sigma}_t^{-1} \otimes \tilde{\Sigma}_t^{-1} \right) (\tilde{L}_t \otimes I) \mathcal{D}^0_k (B_k (I_k^2 + C_k) (\tilde{L}_t \otimes I_k) \mathcal{D}^0_k)^{-1},
\]

and with \( \tilde{L}_t \) a lower triangular matrix. Taking conditional expectations, we obtain

\[
\tilde{\nabla}_t^{(t)} = \tilde{\Psi}_t' \tilde{H}_t' \left( \delta_{1t}^{(t)} y_t \otimes y_t - \text{vec}(\tilde{\Sigma}_t) - (1 - \delta_{1t}^{(t)}) (y_t \otimes \tilde{L}_t \gamma) \right), \tag{2.39}
\]

which only depends on \( \hat{\theta}^{(t)}, \gamma \), and \( \tilde{\Sigma}_t \), and therefore not on \( \theta_2 \). As a result, the modified model for \( y_t \) conditional on \( \zeta_t \) depends on \( \theta_1 \) only.

Using these results, it is clear that \( Q_1(\cdot) \) only depends on \( \theta_1 \) and \( \hat{\theta}^{(t)} \). We have

\[
Q_1(\theta_1, \hat{\theta}^{(t)}) = -\frac{1}{2} \ln |\tilde{\Sigma}_t| - \frac{k}{2} \delta_{3t}^{(t)} - \frac{k}{2} \ln(2\pi) - \frac{1}{2} \delta_{1t}^{(t)} \tilde{x}_t' \tilde{x}_t + \tilde{x}_t' \gamma - \frac{1}{2} \delta_{2t}^{(t)} \gamma' \gamma. \tag{2.40}
\]

For expositional purposes, we restrict our attention to the model with order \((1,1)\) dynamics,

\[
f_{t+1} = A_1 S_t \tilde{\nabla}_t^{(t)} + B_1 f_t. \tag{2.41}
\]

Optimizing (2.40) using the dynamics in (2.41) now becomes similar to estimating a Gaussian multivariate GARCH in Mean model. The transition equation uses weighted (by \( \delta_{1t}^{(t)} \)) rather than unweighted squared observations to drive volatilities and correlations, see (2.39). Similarly, there is weighting by \( \delta_{1t}^{(t)} \) in the objective (2.40). When optimizing over \( \theta_1 \), however, these weights are fixed. Numerical optimization should therefore be faster than direct ML estimation of the full \( \theta \) vector due to the less complicated likelihood and the lower dimensional parameter space.
Using the new estimate of $\theta_1$ obtained by maximizing (2.40), we update the parameter estimate to $\tilde{\theta}^{(l)}$, and use this new estimate to update the weights $\delta_{\ell t}^{(l)}$. The second part of the EM maximization step then follows from

\[ Q_2(\theta_2, \tilde{\theta}^{(l)}) = -\frac{\lambda}{2} \ln(\chi/\psi) - \ln(2) - \ln K_\lambda \left( \sqrt{\chi/\psi} \right) + (\lambda - 1)\delta_{\ell t}^{(l)} - \frac{1}{2}(\chi\delta_{\ell t}^{(l)} + \psi\delta_{2t}^{(l)}) \],

(2.42)

which can be optimized numerically with respect to $\theta_2$ under the constraint $\mu_\zeta = 1$. This can be achieved by optimizing over $\kappa = \chi/\psi > 0$ and $\lambda$, and using (2.10).

The similarity of (2.39) and (2.30) can be taken a step further by noting that $w_t = \delta_{1t}^{(\infty)}$, where $\delta_{1t}^{(\infty)}$ is evaluated using the true parameters. This follows from the fact that

\[ w_t - \delta_{1t}^{(\infty)} = -\frac{\lambda - k/2}{d_{2t}^\kappa} + 0.5 \frac{K_{\lambda-k/2+1} \left( \sqrt{d_{2t}^\kappa \sigma} \right) - K_{\lambda-k/2-1} \left( \sqrt{d_{2t}^\kappa \sigma} \right)}{K_{\lambda-k/2} \left( \sqrt{d_{2t}^\kappa \sigma} \right) \sqrt{d_{2t}^\kappa \sigma}} \]

(2.43)

and the properties of the modified Bessel function of the second kind,

\[ K_{\lambda+1}(\kappa) = 2\lambda \cdot \kappa^{-1} \cdot K_{\lambda}(\kappa) + K_{\lambda-1}(\kappa), \]

and

\[ \frac{\partial \ln K_\lambda(\kappa)}{\partial \kappa} = \frac{K_{\lambda+1}(\kappa) + K_{\lambda-1}(\kappa)}{2 \cdot K_\lambda(\kappa)}, \]

such that from (2.43) it follows that $w_t - \delta_{1t}^{(\infty)} = 0$. 
Chapter 3

Conditional Probabilities for Euro Area Sovereign Default Risk

3.1 Introduction

The Eurozone debt crisis raises the issue of measuring and monitoring interconnected sovereign credit risk. In this paper we construct a novel empirical framework to assess the likelihood of joint and conditional failure for Euro area sovereigns. This new framework allows us to estimate marginal, joint, and conditional probabilities of sovereign default from observed prices for credit default swaps (CDS) on sovereign debt. We define failure as any credit event that would trigger a sovereign CDS contract. Examples of such failures are the non-payment of principal or interest when it is due, a forced exchange of debt into claims of lower value, or a moratorium or official repudiation of the debt. Unlike marginal probabilities, conditional probabilities of sovereign default cannot be obtained from raw market data alone, but instead require a proper joint modeling framework. Our methodology is novel in that our probability assessments are derived from a multivariate framework based on a dynamic Generalized Hyperbolic (GH) skewed-t density that naturally accommodates all relevant empirical features of the data, such as skewed and heavy-tailed changes in individual country CDS spreads, as well as time variation in their volatilities and dependence. Moreover, the model can easily be calibrated to match current market expectations regarding the marginal probabilities of default, similar to for example Segoviano and Goodhart (2009) and Huang et al. (2009).
We make four main contributions. First, we provide estimates of the time variation in Euro area joint and conditional sovereign default risk using a new model and a 10-dimensional data set of sovereign CDS spreads from January 2008 to June 2011. For example, we estimate the conditional probability of a default on Portuguese debt given a Greek failure to be around 30% at the end of our sample. We report similar conditional probabilities for other countries. At the same time, we infer which countries are more exposed than others to certain credit events.

Second, we analyze the extent to which parametric modeling assumptions matter for such joint and conditional risk assessments. Perhaps surprisingly, and despite the widespread use of joint risk measures to guide policy decisions, we are not aware of a detailed investigation of how different parametric assumptions matter for joint and conditional risk assessments. We therefore report results based on a dynamic multivariate Gaussian, symmetric-$t$, and GH skewed-$t$ (GHST) specification. The distributional assumptions turn out to be most important for our conditional assessments, whereas simpler joint failure probability estimates are less sensitive to the assumed dependence structure. In particular, and much in line with Forbes and Rigobon (2002), we show that it is important to account for the different salient features of the data, such as non-zero tail dependence and skewness when interpreting time-varying volatilities and increases in correlations in times of stress.

Third, our modeling framework allows us to investigate the presence and severity of market implied spill-overs in the likelihood of sovereign failure. Specifically, we document spill-overs from the possibility of a Greek failure to the perceived riskiness of other Euro area countries. For example, at the end of our sample we find a difference of about 25% between the one-year conditional probability of a Portuguese default given that Greece does versus that Greece does not default. This suggests that the cost of debt refinancing in some European countries depends to a considerable extent on developments in other countries.

Fourth, we provide an in-depth analysis of the impact on sovereign joint and conditional risks of a key policy announcement on May 9, 2010. On this day, Euro area heads of state announced a comprehensive rescue package to mitigate sovereign risk conditions and perceived risk contagion in the Eurozone. The rescue package contained the European Financial Stability Facility (EFSF), a rescue fund, and the ECB’s Securities
Markets Program (SMP), under which the central bank can purchase government bonds in secondary markets. This event study shows how our model can be used to disentangle market assessments of joint and conditional probabilities. In particular, for May 9, 2010 we find that market perceptions of joint sovereign default risk have decreased, while market perceptions of conditional sovereign default risk have increased at the same time. From a risk perspective, our joint approach is in line with for example Acharya et al. (2010) who focus on financial institutions: bad outcomes are much worse if they occur in clusters. What seems manageable in isolation may not be so if the rest of the system is also under stress. While adverse developments in one country’s public finances could perhaps still be handled with the support of the remaining healthy countries in the Eurozone, the situation may quickly become untenable if one, two, or more countries are already in distress. Relevant questions regarding joint and conditional sovereign default risks would be hard if not impossible to answer without an empirical model such as the one proposed in this paper.

The literature on sovereign credit risk has expanded rapidly and branched off into different fields. Part of the literature focuses on the theoretical development of sovereign default risk and strategic default decisions; see for example Guembel and Sussman (2009) or Yue (2010). Another part of the literature tries to disentangle the different priced components of sovereign credit risk using asset pricing methodology, including the determination of common risk factors across countries; see for example Pan and Singleton (2008), Longstaff et al. (2011), or Ang and Longstaff (2011). Finally, there is a line of literature that investigates the link between sovereign credit risk, country ratings, and macro fundamentals; see for example Haugh et al. (2009), Hilscher and Nosbusch (2010), or De Grauwe and Ji (2012).

Our paper primarily relates to the empirical literature on sovereign credit risk as proxied by sovereign CDS spreads and focuses on spill-over risk as perceived by financial markets. We take a pure time-series perspective instead of assuming a specific pricing model as in Longstaff et al. (2011) or Ang and Longstaff (2011). The advantage of such an approach is that we are much more flexible in accommodating all the relevant empirical features of CDS changes given that we are not bound by the analytical (in)tractability of a particular pricing model. This appears particularly important for the data at hand. In particular, our paper relates closely to the statistical literature for multiple defaults,
such as for example Li (2001), Hull and White (2004) or Garcia Pascual et al. (2006). These papers, however, typically build on a Gaussian or sometimes symmetric Student $t$ dependence structure, whereas we impose a dependence structure that allows for non-zero tail dependence, skewness, and time variation in both volatilities and correlations. Our approach therefore also relates to an important strand of literature on modeling dependence in high dimensions, see for example Demarta and McNeil (2005), Christoffersen et al. (2011), Oh and Patton (2012), and Engle and Kelly (2012), as well as to a growing literature on observation-driven time-varying parameter models, such as for example Patton (2006), Harvey (2010), and Creal, Koopman and Lucas (2011, 2012). Finally, we relate to the CIMDO framework of Segoviano and Goodhart (2009). This is based on a multivariate prior distribution, usually Gaussian or symmetric-$t$, that can be calibrated to match marginal risks as implied by the CDS market. Their multivariate density becomes discontinuous at so-called threshold levels: some parts of the density are shifted up, others are shifted down, while the parametric tails and extreme dependence implied by the prior remain intact at all times. Our model does not have similar discontinuities, while it allows for a similar calibration of default probabilities to current CDS spread levels as Segoviano and Goodhart (2009).

The remainder of the paper is set up as follows. Section 3.2 introduces the conceptual framework for joint and conditional risk measures. Section 3.3 introduces the multivariate statistical model for failure dependence. The empirical results are discussed in Section 4.5. Section 4.7 concludes.

3.2 Conceptual framework

In a corporate credit risk setting, the probability of failure is often modeled as the probability that the value of a firm’s assets falls below the value of its debt at (or before) the time when the debt matures, see Merton (1974) and Black and Cox (1976). To allow for default clustering, the default processes of individual firms can be linked together using a copula function, see for example McNeil et al. (2005). In a sovereign credit risk setting, a similar approach can be adopted, though the interpretation has to be slightly altered given the different nature of a sovereign compared to a corporate default. Rather than to consider asset levels falling below debt values, it is more convenient for sovereign credit
risk to compare costs and benefits of default, see for example Calvo (1988). Default costs may arise from losing credit market access for some time, obstacles to conducting international trade, difficulties in borrowing in the domestic market, etc., while default benefits include immediate debt relief.

To accommodate this interpretation, we introduce a variable \( v_{it} \) that triggers default if \( v_{it} \) exceeds a threshold value \( c_{it} \). The variable \( v_{it} \) captures the time-varying changes in the difference between the perceived benefits and cost of default for sovereign \( i \) at time \( t \). Since a cost, or penalty, can always be recast in terms of a benefit, we incur no loss of generality if we focus on a model with time-varying benefits of default and fixed costs, or vice versa, see Calvo (1988). The \( v_{it}, i = 1, \ldots, n \), are linked together via a Generalized Hyperbolic Skewed Student’s \( t \) (GHST) copula,

\[
v_{it} = (s_t - \mu_s)\tilde{L}_t\gamma + \sqrt{s_t}\hat{L}_t\epsilon_t, \quad i = 1, \ldots, n,
\]

where \( \epsilon_t \in \mathbb{R}^n \) is a vector of standard normally distributed risk factors, \( \tilde{L}_t \) is an \( n \times n \) matrix of risk factor sensitivities, and \( \gamma \in \mathbb{R}^n \) is a vector controlling the skewness of the copula. The random scalar \( \zeta_t \in \mathbb{R}^+ \) is assumed to be an inverse-Gamma distributed risk factor that affects all sovereigns simultaneously, where \( \zeta_t \) and \( \epsilon_t \) are independent, and \( \mu_s = \mathbb{E}[\zeta_t] \). The GHST model can be further generalized to the GH model by assuming a generalized inverse Gaussian distribution for \( \zeta_t \), see McNeil et al. (2005). The current simpler GHST model, however, already accounts for all the empirical features in the CDS data at hand, including skewness and fat tails.

Default dependence in model (3.1) stems from two sources: common exposures to the normally distributed risk factors \( \epsilon_t \) as captured by the time-varying matrix \( \tilde{L}_t \); and a common exposure to the scalar risk factor \( \zeta_t \). The former captures spillover effects through the correlations, while the latter captures such effects through the tail-dependence of the copula. To see this, note that if \( \zeta_t \) is non-random, the first term in (3.1) drops out of the equation and there is zero tail dependence. Conversely, if \( \zeta_t \) is large, all sovereigns are affected at the same time, making joint defaults of two or more sovereigns more likely.

The probability of default \( p_{it} \) of sovereign \( i \) at time \( t \) is given by

\[
p_{it} = \Pr[v_{it} > c_{it}] = 1 - F_i(c_{it}) \quad \Leftrightarrow \quad c_{it} = F_i^{-1}(1 - p_{it}),
\]

(3.2)
where $F_i(\cdot)$ is the cumulative distribution function of $v_{it}$. In our case, $F_i(\cdot)$ is the univariate GHST distribution, which follows directly from the mean-variance mixture construction in equation (3.1). Our main interest, however, is not in the marginal default probability $p_{it}$, but rather in the joint default probability $\Pr[v_{it} > c_{it}, v_{jt} > c_{jt}]$ or the conditional default probability $\Pr[v_{it} > c_{it} | v_{jt} > c_{jt}]$, for $i \neq j$. The (market implied) marginal default probabilities are typically estimated directly from CDS market data under a number of simplifying assumptions. We follow this practice. First, we fix the recovery rate at a stressed level of $rec_i = 25\%$ for all countries and use the 6 months LIBOR rate as the discount rate $r_t$. We assume that the premium payments occur continuously, such that the standard CDS pricing formula as in for example Hull and White (2000) simplifies and can be inverted to extract the market-implied probability of default $p_{it}$. The relation is given by

$$p_{it} = \frac{s_{it} \times (1 + r_t)}{1 - rec_i}, \quad (3.3)$$

where $s_{it}$ is the CDS spread for sovereign $i$ at time $t$, and $r_t$ is our discount rate; see also Brigo and Mercurio (2006, Chapter 21) and Segoviano and Goodhart (2009).

Given our market implied estimates of the default probabilities, we can make use of our multivariate model in (3.1) to infer the magnitude and time-variation in joint and conditional default probabilities. To do this, we proceed in two simple steps. In the first step, we estimate the dependence structure in (3.1) from observed CDS data as explained in Section 3.3, and we infer the threshold values $c_{it}$ by inverting the univariate GHST distributions using our market implied estimates of the default probabilities. In the second step, we then use the calibrated thresholds $c_{it}$ and the estimated dependence structure of the $v_{it}$s to simulate joint and conditional default probabilities. We show in Section 4.5 how the combination of marginal default probabilities calibrated to current CDS spread levels with the time-varying copula structure in (3.1) can lead to new insights into sovereign credit spread spillovers.
3.3 Statistical model

3.3.1 Generalized Autoregressive Score dynamics

As mentioned in Section 3.2, we use sovereign CDS spreads to estimate the time-varying dependence structure in (3.1) and to calibrate the model’s marginal default probabilities through equation (3.3). The statistical model, therefore, closely follows the set-up of the previous section while allowing for time variation in the parameters using the Generalized Autoregressive Score dynamics of Creal et al. (2012).

We assume that we observe a vector \( y_t \in \mathbb{R}^n \), \( t = 1, \ldots, \tau \), of changes in sovereign CDS spreads for sovereign \( i = 1, \ldots, n \), where

\[
y_t = \mu + L_t \varepsilon_t,
\]

with \( \mu \in \mathbb{R}^n \) a vector of fixed unknown means, and \( \varepsilon_t \) a GHST distributed random variable with zero mean, \( \nu \) degrees of freedom, skewness parameter \( \gamma \), and covariance matrix \( I \). To ease the notation, we set \( \mu = 0 \) in the remaining exposition. For \( \mu \neq 0 \), all derivations go through if \( y_t \) is replaced by \( y_t - \mu \). The density of \( y_t \) is denoted by

\[
p(y_t; \tilde{\Sigma}_t, \gamma, \nu) = \frac{\nu^\nu 2^{1-\nu+n}}{\Gamma(\frac{\nu}{2}) \pi^{\nu/2} |\tilde{\Sigma}_t|^\frac{\nu}{2}} \frac{K_{\nu+n}(\sqrt{d(y_t) \cdot (\gamma' \gamma)})}{d(y_t)^{\nu+n} \cdot (\gamma' \gamma)^{-\nu+n}} e^{\gamma' \tilde{\Sigma}_t^{-1}(y_t - \tilde{\mu}_t)},
\]

\[
d(y_t) = \nu + (y_t - \tilde{\mu}_t)' \tilde{\Sigma}_t^{-1} (y_t - \tilde{\mu}_t),
\]

\[
\tilde{\mu}_t = -\frac{\nu}{\nu - 2} \tilde{L}_t \gamma,
\]

where \( \nu > 4 \) is the degrees of freedom parameter, \( \tilde{\mu}_t \) is the location vector, and \( \tilde{\Sigma}_t = \tilde{L}_t \tilde{L}_t' \) is the scale matrix,

\[
\tilde{L}_t = L_t \mathcal{T},
\]

\[
(\mathcal{T}' \mathcal{T})^{-1} = \frac{\nu}{\nu - 2} + \frac{2\nu^2}{(\nu - 2)(\nu - 4)} \gamma \gamma',
\]

and \( K_a(b) \) is the modified Bessel function of the second kind. The matrix \( L_t \) characterizes the time-varying covariance matrix \( \Sigma_t = L_t L_t' \). We consider the standard decomposition

\[
\Sigma_t = L_t L_t' = D_t R_t D_t,
\]
where $D_t$ is a diagonal matrix containing the time-varying volatilities of $y_t$, and $R_t$ is the time-varying correlation matrix.

The fat-tailedness and skewness of the CDS data $y_t$ creates challenges for standard dynamic specifications of volatilities and correlations, such as standard GARCH or DCC type dynamics, see Engle (2002). In the presence of fat tails, large absolute observations $y_{it}$ occur regularly even if volatility is not changing rapidly. If not properly accounted for, such observations lead to biased estimates of the dynamic behavior of volatilities and correlations. The Generalized Autoregressive Score (GAS) framework of Creal et al. (2012) as applied in Chapter 2 to the case of GHST distributions provides a coherent approach to deal with such settings. The GAS model creates an explicit link between the distribution of $y_t$ and the dynamic behavior of $\Sigma_t$, $L_t$, $D_t$, and $R_t$. In particular, if $y_t$ is fat-tailed, observations that lie far outside the center automatically have less impact on future values of the time-varying parameters in $\Sigma_t$. The same holds for observations in the left-hand tail if $y_t$ is left-skewed. The intuition for this is that the score dynamics attribute the effect of a large observation $y_t$ partly to the distributional properties of $y_t$ and partly to a local increase of volatilities and/or correlations. The estimates of dynamic volatilities and correlations thus become more robust to incidental influential observations, which are prevalent in the CDS data used in our empirical analysis. We refer to Creal et al. (2011) and Chapter 2 for more details.

We assume that the time-varying covariance matrix $\Sigma_t$ is driven by a number of unobserved dynamic factors $f_t$, or $\Sigma_t = \Sigma(f_t) = L(f_t)L(f_t)'$. The number of factors coincides with the number of free elements in $\Sigma_t$ in our empirical application later on, but may also be smaller. The dynamics of $f_t$ are specified using the GAS framework for GHST distributed random variables and are given by

$$f_{t+1} = \omega + \sum_{i=0}^{p-1} A_i s_{t-i} + \sum_{j=0}^{q-1} B_j f_{t-j}; \quad \text{(3.11)}$$

$$s_t = S_t \nabla_t, \quad \text{(3.12)}$$

$$\nabla_t = \partial \ln p(y_t; \tilde{\Sigma}(f_t), \gamma, \nu)/\partial f_t, \quad \text{(3.13)}$$

where $\nabla_t$ is the score of the GHST density with respect to $f_t$, $\tilde{\Sigma}(f_t) = L(f_t)T^TL(f_t)', \omega$ is a vector of fixed intercepts, $A_i$ and $B_j$ are appropriately sized fixed parameter matrices,
3.3. STATISTICAL MODEL

\( S_t \) is a scaling matrix for the score \( \nabla_t \), and \( \omega = \omega(\theta), A_i = A_i(\theta), \) and \( B_j = B_j(\theta) \) all depend on a static parameter vector \( \theta \). Typical choices for the scaling matrix \( S_t \) are the unit matrix or inverse (powers) of the Fisher information matrix \( \mathcal{I}_{t-1} \), where

\[
\mathcal{I}_{t-1} = \text{E}[\nabla_t \nabla_t' | y_{t-1}, y_{t-2}, \ldots].
\]

For example, \( S_t = \mathcal{I}_{t-1}^{-1} \) accounts for the curvature in the score \( \nabla_t \).

For appropriate choices of the distribution, the parameterization, and the scaling matrix, the GAS model (3.11)–(3.13) encompasses a wide range of familiar models such as the (multivariate) GARCH model, the autoregressive conditional duration (ACD) model, and the multiplicative error model (MEM); see Creal et al. (2012) for more examples. Details on the parameterization \( \Sigma_t = \Sigma(f_t), D_t = D(f_t), \) and \( R_t = R(f_t) \), and the scaling matrix \( S_t \) used in our empirical application can be found in the appendix.

Using the GHST specification in equation (3.5), the appendix shows that

\[
\nabla_t = \Psi_t' H_t' \text{vec} \left( w_t \cdot y_t y_t' - \hat{\Sigma}_t - \left( 1 - \frac{\nu}{\nu-2} w_t \right) \tilde{L}_t \gamma y_t' \right),
\]

where \( w_t \) is a scalar weight function that decreases in the Mahalanobis distance of \( y_t \) from its center \( \hat{\mu}_t \) as defined in (4.9). The matrices \( \Psi_t \) and \( H_t \) are time-varying, parameterization specific and depend on \( f_t \), but not on the data. Due to the presence of \( w_t \) in (2.9), observations that are far out in the tails receive a smaller weight and therefore have a smaller impact on future values of \( f_t \). This robustness feature is directly linked to the fat-tailed nature of the GHST distribution and allows for smoother correlation and volatility dynamics in the presence of heavy-tailed observations (i.e., \( \nu < \infty \)).

For skewed distributions (\( \gamma \neq 0 \)), the score in (2.9) shows that positive CDS changes have a different impact on correlation and volatility dynamics than negative ones. As explained earlier, this aligns with the intuition that CDS changes from for example the left tail are less informative about changes in volatilities and correlations if the (conditional) observation density is itself left-skewed. For the symmetric Student’s t case, we have \( \gamma = 0 \) and the asymmetry term in (2.9) drops out. If furthermore the fat-tailedness is ruled out by considering \( \nu \to \infty \), one can show that the weights \( w_t \) tend to 1 and that \( \nabla_t \) collapses to the intuitive form for a multivariate GARCH model, \( \nabla_t = \Psi_t' H_t' \text{vec}(y_t y_t' - \Sigma_t) \).
3.3.2 Parameter estimation

The parameters of the dynamic GHST model can be estimated by standard maximum likelihood procedures as the likelihood function is known in closed form using a standard prediction error decomposition. The joint estimation of all parameters in the model, however, is rather cumbersome. Therefore, we split the estimation in two steps relating to (i) the marginal behavior of the coordinates \( y_{it} \) and (ii) the joint dependence structure of the vector of standardized residuals \( D^{-1}_t y_t \). Similar two-step procedures can be found in Engle (2002), Hu (2005), and other studies that are based on a multivariate GARCH framework.

In the first step, we estimate a dynamic GHST model for each series \( y_{it} \) separately using a GAS(1,1) dynamic specification with \( p = q = 1 \) and taking our time-varying parameter \( f_t \) as the log-volatility \( \log(\sigma_{it}) \). The skewness parameter \( \gamma_i \) is also estimated for each series separately, while the degrees of freedom parameter \( \nu \) is fixed at a pre-determined value. This restriction ensures that the univariate GHST distributions are the marginal distributions from the multivariate GHST distribution and that the model is therefore internally consistent.

In the second step, we consider the standardized data \( z_{it} = y_{it} / \hat{\sigma}_{it} \), where \( \hat{\sigma}_{it} \) are obtained from the first step. Using \( z_t = (z_{1t}, \ldots, z_{nt})' \), we estimate a multivariate dynamic GHST model using again a GAS(1,1) dynamic specification. The GHST distribution in this second step has mean zero, skewness parameters \( \hat{\gamma}_i, i = 1, \ldots, n \), as estimated in the first step, the same pre-determined value for \( \nu \), and covariance matrix \( \text{cov}(z_t) = R_t = R(f_t) \), where \( f_t \) contains the spherical coordinates of the choleski decomposition of the correlation matrix \( R_t \); see the appendix for further details.

The advantages of the two-step procedure for computational efficiency are substantial, particularly if the number \( n \) of time series considered in \( y_t \) is large. The univariate models of the first step can be estimated at low computational cost. Using these estimates, the univariate dynamic GHST models are used as a filter to standardize the individual CDS spread changes. In the second step, only the parameters that determine the dynamic correlations remain to be estimated.
3.4 Empirical application: Euro area sovereign risk

3.4.1 CDS data

We compute joint and conditional probabilities of failure for a set of ten countries in the Euro area. We focus on sovereigns that have a CDS contract traded on their reference bonds since the beginning of our sample in January 2008. We select ten countries: Austria (AT), Belgium (BE), Germany (DE), Spain (ES), France (FR), Greece (GR), Ireland (IE), Italy (IT), the Netherlands (NL) and Portugal (PT). CDS spreads are available for these countries at a daily frequency from January 1, 2008 to June 30, 2011, yielding $\tau = 913$ observations. The CDS contracts have a five year maturity. They are denominated in U.S. dollars and therefore do not depend on foreign exchange risk concerns should a European credit event materialize. Such contracts are also far more liquidly traded than their Euro denominated counterparts. All time series data are obtained from Bloomberg. We prefer CDS spreads to bond yield spreads as a measure of sovereign default risk since the former are less affected by liquidity and flight-to-safety issues, see for example Pan and Singleton (2008) and Ang and Longstaff (2011). In addition, our CDS series are likely to be less affected than bond yields by the outright government bond purchases that might have taken place under the Securities Markets Program during the second half of our sample, see Section 3.4.5 below.

The use of CDS data to estimate market implied failure probabilities means that our probability estimates combine physical failure probabilities with the price of sovereign default risk. As a result, our risk measures constitute an upper bound for an investor worried about losing money due to a joint sovereign failure. This has to be kept in mind when interpreting the empirical results later on. Estimating failure probabilities directly from observed defaults, however, is impossible in our context, as OECD defaults are not observed over our sample period. Even if such defaults would have been observed, they would not have allowed us to perform the detailed empirical analysis in the current section on the dynamics of joint and conditional failure probabilities.

Table 3.1 provides summary statistics for daily de-meaned changes in these ten CDS spreads. All time series have significant non-Gaussian features under standard tests and significance levels. In particular, we note the non-zero skewness and large values of kurtosis for almost all time series in the sample. All series are covariance stationary according to
standard unit root (ADF) tests.

Table 3.1: CDS descriptive statistics

The summary statistics correspond to daily changes in observed sovereign CDS spreads for ten Euro area countries from January 2008 to June 2011. Mean, Median, Standard Deviation, Minimum and Maximum are multiplied by 100. Almost all skewness and excess kurtosis statistics have p-values below $10^{-4}$, except the skewness parameters of France and Ireland.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Std.Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>1.07</td>
<td>18.74</td>
<td>-0.27</td>
<td>0.42</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.33</td>
<td>8.29</td>
<td>-0.21</td>
<td>0.27</td>
</tr>
<tr>
<td>Germany</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.41</td>
<td>7.98</td>
<td>-0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>Spain</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>-0.71</td>
<td>18.47</td>
<td>-0.79</td>
<td>0.50</td>
</tr>
<tr>
<td>France</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.14</td>
<td>6.38</td>
<td>-0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>Greece</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.30</td>
<td>-0.31</td>
<td>46.81</td>
<td>-3.64</td>
<td>2.91</td>
</tr>
<tr>
<td>Ireland</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.12</td>
<td>0.02</td>
<td>9.13</td>
<td>-0.79</td>
<td>0.55</td>
</tr>
<tr>
<td>Italy</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>-0.82</td>
<td>25.54</td>
<td>-0.77</td>
<td>0.45</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>1.62</td>
<td>19.59</td>
<td>-0.10</td>
<td>0.24</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.13</td>
<td>-2.60</td>
<td>51.49</td>
<td>-1.85</td>
<td>0.74</td>
</tr>
</tbody>
</table>

3.4.2 Marginal and joint risk

We model the CDS spread changes with the framework explained in Section 3.3 based on the dynamic GHST specification (3.11). We consider three different choices for the parameters, corresponding to a Gaussian, a Student-$t$, and a GHST distribution, respectively. We treat the degrees of freedom parameter $\nu$ as a robustness parameter; compare Franses and Lucas (1998). This implies we fix the degrees of freedom at $\nu = 5$ rather than estimating it. The advantage of such an approach is that it further simplifies the estimation process, while retaining many of the robustness features of model (3.11). In particular, fixing $\nu$ at $\nu = 5$ may seem high at first sight given some of the high kurtosis values in Table 3.1. The value is small enough, however, to result in a substantial robustification of the results via the weights $w_t$ in (2.9), both in terms of likelihood evaluation as well as in terms of the volatility and correlation dynamics.

Figure 3.1 plots estimated volatility levels for the three different models along with the squared CDS changes. Figure 3.2 plots the volatility differences from these models for two countries: Greece and Portugal. The assumed statistical model (Gaussian, Student-$t$, GHST) directly influences the volatility estimates. The volatilities from the univariate Gaussian models repeatedly seem to be too high. The thin tails of the Gaussian distri-
bution imply that volatility increases sharply in response to a jump in the CDS spread, see for example the Spanish CDS spread around April 2008, and many countries around Spring 2010. In particular, the magnitude of the increase in volatility appears too large when compared to the subsequent squared CDS spread changes. The volatility estimates based on the Student-\(t\) and GHST distribution change less abruptly after incidental large changes than the Gaussian ones due to the weighting mechanism in (2.9). The results for the Student-\(t\) and GHST are very similar and in line with the subsequent squared changes in CDS spreads. Some differences are visible for the series that exhibit significant skewness, such as the time series for Greece and Portugal.

Table 3.2 reports the parameter estimates for the ten univariate country-specific models. In all cases, volatility is highly persistent, i.e., \(B\) is close to one. Note that the parameterization of our score driven model is different than that of a standard GARCH model. In particular, the persistence is completely captured by \(B\) rather than by \(A + B\) as in the GARCH case. Also note that \(\omega\) sometimes takes on negative values. This is natural as we define \(f_t\) to be the log-volatility rather than the volatility itself.

Next, we estimate the dynamic correlation coefficients for the standardized CDS spread changes. Given \(n = 10\), there are 45 different elements in the correlation matrix. Figure 3.3 plots the average correlation, averaged across 45 time-varying bivariate pairs, for each model specification. As a robustness check, we benchmark each multivariate model-based estimate to the average over 45 correlation pairs obtained from a 60 business days rolling window. Over each window we use the same pre-filtered marginal data as for the multivariate model estimates.

If we compare the correlation estimates across the different specifications, the GHST model matches the rolling window estimates most closely. Rolling window and GHST correlations are low in the beginning of the sample at around 0.3 and increase to around 0.75 during 2010 and 2011. In the beginning of the sample the GHST-based average correlation is lower than that implied by the two alternative specifications. The pattern reverses in the second half of the sample. This result is in line with correlations that tend to increase during times of stress.

The correlation estimates vary considerably over time across all model specifications considered. Estimated dependence across Euro area sovereign risk increases sharply for the first time around September 15, 2008, on the day of the Lehman failure, and around
Figure 3.1: Estimated time-varying volatilities for CDS spread changes of EA countries

We report three different estimates of time-varying volatility that pertain to changes in CDS spreads on sovereign debt for 10 countries. The volatility estimates are based on different parametric assumptions regarding the univariate distribution of sovereign CDS spread changes: Gaussian, symmetric $t$, and GHST. As a direct benchmark, the squared CDS spread changes are plotted as well.
Figure 3.2: Estimated volatility differences for CDS spread changes: Greece and Portugal.

We report the difference of the volatility estimates for CDS spreads on Greece and Portugal. Top two panels contain the time series plots of the CDS spread change data. We plot the differences of volatility estimates from Gaussian/symmetric $t$ and GHST, as percentages of the GHST volatilities.
**Table 3.2: Model parameter estimates**

The table reports parameter estimates that pertain to three different model specifications. The sample consists of daily changes from January 2008 to June 2011. The degree of freedom parameter $\nu$ is set to five for the $t$ distributions. Parameters in $\gamma$ are estimated in the marginal distributions.

<table>
<thead>
<tr>
<th></th>
<th>AT</th>
<th>BE</th>
<th>DE</th>
<th>ES</th>
<th>FR</th>
<th>GR</th>
<th>IE</th>
<th>TT</th>
<th>NL</th>
<th>PT</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gaussian</td>
</tr>
<tr>
<td>$A$</td>
<td>0.06</td>
<td>0.10</td>
<td>0.08</td>
<td>0.15</td>
<td>0.11</td>
<td>0.12</td>
<td>0.08</td>
<td>0.11</td>
<td>0.08</td>
<td>0.16</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>$B$</td>
<td>0.99</td>
<td>0.98</td>
<td>0.97</td>
<td>0.94</td>
<td>0.97</td>
<td>0.99</td>
<td>0.96</td>
<td>0.99</td>
<td>0.97</td>
<td>0.99</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.14</td>
<td>-0.18</td>
<td>-0.12</td>
<td>0.00</td>
<td>-0.09</td>
<td>0.00</td>
<td>-0.11</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.03)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>$t$</td>
<td>0.28</td>
<td>0.30</td>
<td>0.35</td>
<td>0.39</td>
<td>0.40</td>
<td>0.42</td>
<td>0.30</td>
<td>0.34</td>
<td>0.26</td>
<td>0.36</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.31)</td>
<td>(0.18)</td>
<td>(0.68)</td>
<td>(0.00)</td>
<td>(0.22)</td>
<td>(0.17)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.07</td>
<td>0.05</td>
<td>-0.07</td>
<td>0.09</td>
<td>0.00</td>
<td>0.14</td>
<td>0.09</td>
<td>0.08</td>
<td>-0.02</td>
<td>0.11</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(1.62)</td>
<td>(0.79)</td>
<td>(4.14)</td>
<td>(0.00)</td>
<td>(0.84)</td>
<td>(0.82)</td>
<td>(0.28)</td>
<td>(0.15)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>GHST</td>
<td>0.13</td>
<td>0.15</td>
<td>0.21</td>
<td>0.16</td>
<td>0.22</td>
<td>0.17</td>
<td>0.14</td>
<td>0.16</td>
<td>0.16</td>
<td>0.15</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>$B$</td>
<td>0.99</td>
<td>0.98</td>
<td>0.93</td>
<td>0.98</td>
<td>0.95</td>
<td>0.97</td>
<td>0.98</td>
<td>0.98</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>-0.04</td>
<td>-0.08</td>
<td>-0.29</td>
<td>-0.05</td>
<td>-0.18</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.18</td>
<td>-0.05</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.05)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.05)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.11</td>
<td>0.17</td>
<td>0.04</td>
<td>0.12</td>
<td>0.12</td>
<td>0.35</td>
<td>0.22</td>
<td>0.10</td>
<td>0.06</td>
<td>0.29</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 3.3: Average correlation over time
Plots of the estimated average correlation over time, where averaging takes place over 45 estimated correlation coefficients. The correlations are estimated based on different parametric assumptions: Gaussian, symmetric $t$, and GH Skewed-$t$ (GHST). The time axis runs from March 2008 to June 2011. The corresponding rolling window correlations are each estimated using a window of sixty business days of pre-filtered CDS changes. The bottom-right panel collects four series for comparison.
September 30, 2008, when the Irish government issued a blanket guarantee for all deposits and borrowings of six large financial institutions. Average GHST correlations remain high afterwards, around 0.75, until around May 10, 2010. At this time, Euro area heads of state introduced a rescue package that contained government bond purchases by the ECB under the so-called Securities Markets Program, and the European Financial Stability Facility, a fund designed to provide financial assistance to Euro area states in economic difficulties. After an eventual decline to around 0.6 towards the end of 2010, average correlations increase again towards the end of the sample.

The parameter estimates for volatility and correlations are shown in Table 3.2. Unlike the raw sample skewness, the estimated skewness parameters are all positive, indicating a fatter right tail of the distribution of CDS changes. The negative raw skewness may be the result of several influential outliers. These are accommodated in a model specification with fat-tails.

3.4.3 Joint probabilities of Eurozone financial stress

This section reports marginal and joint risk estimates that pertain to Euro area sovereign default. First, Figure 3.4 plots estimates of CDS-implied probabilities of default (pd) over a one year horizon based on (3.3). These are directly inferred from CDS spreads, and do not depend on parametric assumptions regarding their joint distribution. Market-implied pd’s range from around 1% for Germany and the Netherlands to above 10% for Greece, Portugal, and Ireland at the end of our sample.

The top panel of Figure 3.5 tracks the market-implied probability of two or more failures among the ten Euro area sovereigns in the portfolio over a one year horizon. The joint failure probability is calculated by simulation, using 50,000 draws at each time $t$. This simple estimate combines all marginal and joint failure information into a single time series plot and reflects the deterioration of debt conditions since the beginning of the Eurozone crisis. The overall dynamics are roughly similar across the different distributional specifications.

The probability of two or more failures over a one year horizon, as reported in Figure 3.5, starts to pick up in the weeks after the Lehman failure and the Irish blanket guarantee in September 2008. The joint probability estimate peaks in the first quarter of 2009, at the height of the Irish debt crisis, then decreases until the third quarter of 2009.
Figure 3.4: Implied marginal failure probabilities from CDS markets
The risk neutral marginal probabilities of failure for ten Euro area countries extracted from CDS markets. The time axis is from January 2008 to June 2011.
Figure 3.5: Probability of two or more failures

The top panel plots the time-varying probability of two or more failures (out of ten) over a one-year horizon. Estimates are based on different distributional assumptions regarding marginal risks and multivariate dependence: Gaussian, symmetric-$t$, and GH skewed-$t$ (GHST). The bottom panel plots model-implied probabilities for $n^*$ sovereign failures over a one year horizon, for $n^* = 0, 1, 2, 3$. 
It is increasing since then until the end of the sample. The joint probability decreases sharply, but only temporarily, around the May 10, 2010 announcement of the the European Financial Stability Facility and the European Central Bank’s intervention in government debt markets starting at around the same time. We come back to this later.

In the beginning of our sample, the joint failure probability from the GHST model is higher than that from the Gaussian and symmetric-$t$ model. This pattern reverses towards the end of the sample, when the Gaussian and symmetric-$t$ estimates are slightly higher than the GHST estimate. Towards the end of the sample, the joint probability measure is heavily influenced by the possibility of a credit event in Greece and Portugal. The CDS changes for each of these countries are positively skewed, i.e., have a longer right tail. As the crisis worsens, we observe more frequent positive and extreme changes, which increase the volatility in the symmetric models more than in the skewed setting. Higher volatility translates into higher marginal risk, or lower estimated default thresholds. This explains the (slightly) different patterns in the estimated probabilities of joint failures.

The bottom panel in Figure 3.5 plots the probability of a pre-specified number of failures. The lower level of our GHST joint failure probability in the top panel of Figure 3.5 towards the end of the sample is due to the higher probability of no defaults in that case. Altogether, the level and dynamics in the estimated measures of joint failure from this section do not appear to be very sensitive to the precise model specification.

### 3.4.4 Spillover measures: What if ...failed?

This section investigates conditional probabilities of failure. Such conditional probabilities relate to questions of the “what if?” type and reveal which countries may be most vulnerable to the failure of a given other country. We condition on a credit event in Greece to illustrate our general methodology. We pick this case since it has by far the highest market-implied probability of failure at the end of our sample period. To our knowledge, this is the first attempt in the literature on evaluating the spill-over effects and conditional probability of sovereign failures. Clearly, conditioning on a credit event is different from conditioning on incremental changes in other countries’ risks, see Caceres et al. (2010) and Caporin et al. (2012).

Figure 3.6 plots the conditional probability of default for nine Euro area countries if Greece defaults. We distinguish four cases, i.e., Gaussian dependence, symmetric-$t$,
GHST, and GHST with zero correlations. The last experiment is included to disentangle the effect of correlations and tail dependence, see our discussion below equation (3.1). Regardless of the parametric specification, Ireland and Portugal seem to be most affected by a Greek failure, with conditional probabilities of failure of around 30%. Other countries may be perceived as more ‘ring-fenced’ as of June 2011, with conditional failure probabilities below 20%. The level and dynamics of the conditional estimates are sensitive to the parametric assumptions. The conditional default probability estimates are highest in the GHST case. The symmetric-\(t\) estimates in turn are higher than those obtained under the Gaussian assumption. The bottom right panel of Figure 3.6 demonstrates that even if the correlations are put to zero, the GHST still shows extreme dependence due to the mixing variable \(\xi_t\) in (3.1). The correlations and mixing construction thus operate together to capture the dependence in the data.

Figure 3.7 plots the pairwise correlation estimates for Greece with each of the remaining nine Euro area countries. The estimated correlations for the GHST model are higher than for the other two models in the second half of the sample. This is consistent with the higher level of conditional probabilities of default in the GHST case compared to the other distributional assumptions, as discussed above for Figure 3.5. Interestingly, the dynamic correlation estimates of Euro area countries with Greece increased most sharply in the first half of 2009. These are the months before the media attention focused on the Greek debt crisis, which was more towards the end of 2009 up to Spring 2010.

Figure 3.8 and 3.9 plot the difference between the conditional probability of failure of a given country given that Greece fails and the respective conditional probability of failure given that Greece does not fail. We refer to this difference as a spillover component or contagion effect as the differences relate to the question whether CDS markets perceive any spillovers from a potential Greek default to the likelihood of other Euro area countries failing. The level of estimated spillovers are substantial. For example, the difference in the conditional probability of a Portuguese failure given that Greece does or does not fail, is about 25%. The spillover estimates do not appear to be very sensitive to the different parametric assumptions. In all cases, Portugal and Ireland appear the most vulnerable to a Greek default since around mid-2010.

The conditional probabilities can be scaled by the time-varying marginal probability of a Greek failure to obtain pairwise joint failure risks. These joint risks are increasing
3.4. **EMPIRICAL APPLICATION: EURO AREA SOVEREIGN RISK**

![Figure 3.6: Conditional probabilities of failure given that Greece fails](image)

Plots of annual conditional failure probabilities for nine Euro area countries given a Greek failure. We distinguish estimates based on a Gaussian dependence structure, symmetric-$t$, GH skewed-$t$ (GHST), and a GHST with zero correlations.
Figure 3.7: Dynamic correlation of Euro area countries with Greece
The time-varying bivariate correlation pairs for nine Euro area countries and Greece. The correlation estimates are obtained from the ten-dimensional multivariate model with a Gaussian, symmetric-\(t\), and GH skewed-\(t\) (GHST) dependence structure, respectively.
3.4. **EMPIRICAL APPLICATION: EURO AREA SOVEREIGN RISK**

Figure 3.8: Risk spillover components

The difference between the (simulated) probability of failure of \( i \) given that Greece fails and the probability of failure of \( i \) given that Greece does not fail. The underlying distributions are multivariate Gaussian, symmetric-\( t \), and GH skewed-\( t \) (GHST), respectively.
Figure 3.9: Risk spillover components for each country

Figures for every country of the difference between the (simulated) probability of failure of $i$ given that Greece fails and the probability of failure of $i$ given that Greece does not fail. The underlying distributions are multivariate Gaussian, symmetric-$t$, and GH skewed-$t$ (GHST), respectively.
towards the end of the sample and are higher in 2011 than in the second half of 2009. Annual joint probabilities for nine countries are plotted in Figure 3.10. For example, the risk of a joint failure over a one year horizon of both Portugal and Greece, as implied by CDS markets, is about 10% at the end of our sample.

3.4.5 Event study: the May 9, 2010 rescue package and risk dependence

During a weekend meeting on May 8–9, 2010, Euro area heads of state ratified a comprehensive rescue package to mitigate sovereign risk conditions and perceived risk contagion in the Eurozone. This section analyses the impact of the resulting simultaneous announcement of the European Financial Stability Facility (EFSF) and the ECB’s Securities Markets Program (SMP) on Euro area joint risk and conditional risk as implied by our empirical model. We do so by comparing CDS-implied risk conditions closely before and after the announcement of May 9, 2010.

The agreed upon rescue fund, the European Financial Stability Facility (EFSF), is a limited liability company with an objective to preserve financial stability of the Euro area by providing temporary financial assistance to Euro area member states in economic difficulties. Initially committed funds were 440bn Euro. The announcement made clear that EFSF funds can be combined with funds raised by the European Commission of up to 60bn Euro, and funds from the International Monetary Fund of up to 250bn Euro, for a total safety net up to 750bn Euro.

A second key component of the May 9, 2010 package consisted of the ECB’s government bond buying program, the SMP. Specifically, the ECB announced that it would start to intervene in secondary government bond markets to ensure depth and liquidity in those market segments that are qualified as being dysfunctional. These purchases were meant to restore an appropriate transmission of monetary policy actions targeted towards price stability in the medium term. The SMP interventions were almost always sterilized through additional liquidity-absorbing operations.

The joint impact of the May 9, 2010 announcement of the EFSF and SMP as well as of the initial bond purchases on joint risk estimates can be seen in the top panel of Figure 3.5. The figure suggests that the probability of two or more credit events in our sample of ten countries decreases from about 7% to approximately 3% before and after the May 9, 2010
Figure 3.10: Joint default risk with Greece

The time-varying probability of two simultaneous credit events in Greece and a given other Euro area country. The estimates are obtained from a multivariate model based on a Gaussian, symmetric-$t$, and GH skewed-$t$ (GHST) density, respectively.
3.4. EMPIRICAL APPLICATION: EURO AREA SOVEREIGN RISK

announcement. Figure 3.4 indicates that marginal risks decreased considerably as well. The graphs also suggest that these decreases were temporary. The average correlation plots in Figure 3.3 do not suggest a wide-spread and prolonged decrease in dependence. Instead, there seems to be an up-tick in average correlations. Overall, the evidence so far suggest that the announcement of the policy measures and initial bond purchases may have substantially lowered joint risks, but not necessarily through a decrease in joint dependence.

To further investigate the impact on joint and conditional sovereign risk from actions communicated on May 9, 2010 and implemented shortly afterwards, Table 3.3 reports model-based estimates of joint and conditional risk. We report our risk estimates for two dates, Thursday May 6, 2010 and Tuesday May 11, 2010, i.e., two days before and after the announced change in policy. The top panel of Table 3.3 confirms that the joint probability of a credit event in, say, both Portugal and Greece, or Ireland and Greece, declines from 4.8% to 2.1% and 3.0% to 1.7%, respectively. These are large decreases in joint risk. For any country in the sample, the probability of that country failing simultaneously with Greece or Portugal over a one year horizon is substantially lower after the May 9, 2010 policy announcement than before.

The bottom panel of Table 3.3, however, indicates that the decrease in joint failure probabilities is generally not due to a decline in failure dependence, ‘interconnectedness’, or ‘contagion’. Instead, the conditional probabilities of a credit event in for example Greece or Ireland given a credit event in Portugal increases from 77% to 81% and from 45% to 56%, respectively. Similarly, the conditional probability of a credit event in Belgium or Ireland given a credit event in Greece increases from 10% to 13% and from 24% to 26%, respectively.

As a bottom line, based on the initial impact of the two policy measures on CDS prices, our analysis suggests that the two policies may have been perceived to be less of a ‘firewall’ or ‘ringfence’ measure, i.e., intended to lower the impact and spread of an adverse development should it actually occur. Markets perceived the measures much more as a means to affect the probability of individual adverse outcomes downwards, but without decreasing dependence. These findings are robust to, for example, alternative choices for the degrees of freedom parameter $\nu$ in the copula, and different choices for the expected recovery rate in case of defaults.
Table 3.3: Joint and conditional failure probabilities

The top and bottom panels report model-implied joint and conditional probabilities of a credit event for a subset of countries, respectively. The probabilities are based on the model estimated over the whole sample. For the conditional probabilities $\Pr(i \text{ failing} \mid j \text{ failed})$, the conditioning events $j$ are in the columns (PT, GR, DE), while the events $i$ are in the rows (AT, BE, ..., PT). Avg contains the averages for each column.

<table>
<thead>
<tr>
<th></th>
<th>Thu May 6, 2010</th>
<th>Tue May 11, 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PT</td>
<td>GR</td>
</tr>
<tr>
<td>AT</td>
<td>1.1%</td>
<td>1.1%</td>
</tr>
<tr>
<td>BE</td>
<td>1.4%</td>
<td>1.3%</td>
</tr>
<tr>
<td>DE</td>
<td>1.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td>ES</td>
<td>2.9%</td>
<td>2.9%</td>
</tr>
<tr>
<td>FR</td>
<td>1.1%</td>
<td>1.0%</td>
</tr>
<tr>
<td>GR</td>
<td>4.8%</td>
<td></td>
</tr>
<tr>
<td>IR</td>
<td>2.8%</td>
<td>3.0%</td>
</tr>
<tr>
<td>IT</td>
<td>2.7%</td>
<td>2.7%</td>
</tr>
<tr>
<td>NL</td>
<td>1.0%</td>
<td>0.9%</td>
</tr>
<tr>
<td>PT</td>
<td>4.4%</td>
<td>4.4%</td>
</tr>
<tr>
<td>Avg</td>
<td>2.1%</td>
<td>2.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Thu May 6, 2010</th>
<th>Tue May 11, 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PT</td>
<td>GR</td>
</tr>
<tr>
<td>AT</td>
<td>18%</td>
<td>8%</td>
</tr>
<tr>
<td>BE</td>
<td>22%</td>
<td>10%</td>
</tr>
<tr>
<td>DE</td>
<td>16%</td>
<td>8%</td>
</tr>
<tr>
<td>ES</td>
<td>47%</td>
<td>23%</td>
</tr>
<tr>
<td>FR</td>
<td>18%</td>
<td>8%</td>
</tr>
<tr>
<td>GR</td>
<td>77%</td>
<td>8%</td>
</tr>
<tr>
<td>IR</td>
<td>45%</td>
<td>24%</td>
</tr>
<tr>
<td>IT</td>
<td>43%</td>
<td>21%</td>
</tr>
<tr>
<td>NL</td>
<td>17%</td>
<td>7%</td>
</tr>
<tr>
<td>PT</td>
<td>35%</td>
<td>35%</td>
</tr>
<tr>
<td>Avg</td>
<td>34%</td>
<td>16%</td>
</tr>
</tbody>
</table>
3.5 Conclusion

We have proposed a novel empirical framework to assess the likelihood of joint and conditional failure for Euro area sovereigns. Our methodology is novel in that our joint risk measures are derived from a multivariate framework based on a dynamic Generalized Hyperbolic skewed-t (GHST) density that naturally accommodates skewed and heavy-tailed changes in marginal risks as well as time variation in volatility and multivariate dependence. When applying the model to Euro area sovereign CDS data from January 2008 to June 2011, we find significant time variation in risk dependence, as well as considerable spillover effects in the likelihood of sovereign failures. We also documented how parametric assumptions, including assumptions about higher order moments, matter for joint and conditional risk assessments. Using the May 9, 2010 new policy measures of the European heads of state, we illustrated how the model contributes to our understanding of market perceptions about specific policy measures.

Appendix to Chapter 3

The Generalized Autoregressive Score model of Creal et al. (2011, 2012) for the GH skewed-t (GHST) density (3.5) adjusts the time-varying parameter $f_t$ at every step using the scaled score of the density at time $t$. This can be regarded as a steepest ascent improvement of the parameter using the local (at time $t$) likelihood fit of the model. Under the correct specification of the model, the scores form a martingale difference sequence.

We partition $f_t$ as $f_t = (f^v_t, f^c_t)$ for the (diagonal) matrix $D_t^2 = D(f^v_t)^2$ of variances and correlation matrix $R_t = R(f^c_t)$, respectively, where $\Sigma_t = D_t R_t D_t = \Sigma(f_t)$. We set $f^v_t = \ln(\text{diag}(D^2_t))$, which ensures that variances are always positive, irrespective of the value of $f^v_t$. For the correlation matrix, we use the hypersphere transformation also used in Creal et al. (2011) and Chapter 2. This ensures that $R_t$ is always a correlation matrix, i.e., positive semi-definite with ones on the diagonal. We set $R_t = R(f^c_t) = X_t'X_t$, with $f^c_t$
as a vector containing \( n(n - 1)/2 \) time-varying angles \( \phi_{ijt} \in [0, \pi] \) for \( i > j \), and

\[
X_t = \begin{pmatrix}
1 & c_12 & c_13 & \cdots & c_1kt \\
0 & s_12 & c_23s_{13t} & \cdots & c_2ks_{1kt} \\
0 & 0 & s_{23} & \cdots & c_3ks_{2kt}s_{1kt} \\
0 & 0 & 0 & \cdots & c_4ks_{3kt}s_{2kt}s_{1kt} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & c_{k-1}ks_{kkt} \prod_{\ell=1}^{k-2}s_{\ellkt} \\
0 & 0 & 0 & \cdots & \prod_{\ell=1}^{k-1}s_{\ellkt}
\end{pmatrix}, \tag{3.15}
\]

where \( c_{ijt} = \cos(\phi_{ijt}) \) and \( s_{ijt} = \sin(\phi_{ijt}) \). The dimension of \( f_t^c \) thus equals the number of correlation pairs.

As implied by equation (3.13), we take the derivative of the log-density with respect to \( f_t \), and obtain

\[
\nabla_t = \frac{\partial \text{vech}(\Sigma_t)'}{\partial f_t} \frac{\partial \text{vech}(L_t)'}{\partial \Sigma_t} \frac{\partial \text{vech}(\tilde{L}_t)'}{\partial L_t} \frac{\partial \ln p^{GH}(y_t|f_t)}{\partial f_t}
\]

\[
= \Psi_t H_t' \left( w_t(y_t \otimes y_t) - \text{vec}(\tilde{\Sigma}_t) - \left(1 - \frac{\nu}{\nu - 2}w_t\right)(y_t \otimes \tilde{L}_t)\gamma \right)
\]

\[
= \Psi_t H_t' \text{vec} \left( w_t y_t y_t' - \tilde{\Sigma}_t - \left(1 - \frac{\nu}{\nu - 2}w_t\right)\tilde{L}_t \gamma y_t' \right), \tag{3.16}
\]

\[
\Psi_t = \frac{\partial \text{vech}(\Sigma_t)}{\partial f_t}, \tag{3.17}
\]

\[
H_t = (\tilde{\Sigma}_t^{-1} \otimes \tilde{\Sigma}_t^{-1})(\tilde{L}_t \otimes I) \left( (T' \otimes I_k) D_k^0 \right) \left( B_k (I_{k2} + C_k) (L_t \otimes I_k) D_k^0 \right)^{-1}, \tag{3.18}
\]

\[
w_t = \frac{\nu + n}{2 \cdot d(y_t)} - k'_{\nu+n}/2 \left( \frac{\sqrt{d(y_t)} \cdot \gamma' \gamma}{d(y_t) \cdot \gamma' \gamma} \right), \tag{3.19}
\]

where \( k'_{\nu}(b) = \partial \ln K_n(b)/\partial b \) is the derivative of the log modified Bessel function of the second kind, \( D_k^0 \) is the the duplication matrix \( \text{vec}(L) = D_k^0 \text{vech}(L) \) for a lower triangular matrix \( L \), \( D_k \) is the standard duplication matrix for a symmetric matrix \( S \) \( \text{vec}(S) = D_k \text{vech}(S) \), \( B_k = (D_k^s D_k)^{-1} D_k^s \), and \( C_k \) is the commutation matrix, \( \text{vec}(S') = C_k \text{vec}(S) \) for an arbitrary matrix \( S \). For completeness, we mention that \( \tilde{L}_t = L_t T, \tilde{\Sigma}_t = \tilde{L}_t \tilde{L}_t', \) and

\[
(T'T)^{-1} = \frac{\nu}{\nu - 2} I + \frac{2\nu^2}{(\nu - 2)^2(\nu - 4)} \gamma' \gamma.
\]

To scale the score \( \nabla_t \), Creal et al. (2012) propose the use of powers of the inverse information matrix. The information matrix for the GHST distribution, however, does not have a tractable form. Therefore, we scale by the information matrix of the symmetric
Student’s $t$ distribution,

$$S_t = \left\{ \Psi'(I \otimes \tilde{L}_t^{-1})'[gG - \text{vec}(I)\text{vec}(I)'](I \otimes \tilde{L}_t^{-1})\Psi \right\}^{-1},$$  \hfill (3.20)

where $g = (\nu + n)/(\nu + 2 + n)$, and $G = E[x_t x_t' \otimes x_t x_t']$ for $x_t \sim N(0, I_n)$. Chapter 2 demonstrate that this results in a stable model that outperforms alternatives such as the DCC if the data are fat-tailed and skewed.

Using the dynamic GH model for the individual CDS series, we first estimate the parameters for the $f_t^v$ process. Applying equations (3.16) to (3.19) in the univariate setting, we compute the $f_t^v$s and use them to filter the data. The time-varying factor for country $i$’s volatility follows as

$$f_{i,t+1}^v = \omega_i^v + a_i^v s_{i,t}^v + b_i^v f_{i,t}^v,$$  \hfill (3.21)

with $a_i^v$ and $b_i^v$ scalar parameters corresponding to the $i$th series.

Next, we estimate the parameters for the $f_t^c$ process using the filtered data $y_{it}/\exp(f_{it}^v/2)$. Assuming the variances are constant ($D_t = I_n$), the covariance matrix $\Sigma_t$ is equivalent to $R_t$. The matrix $\Psi_t$ should only contain the derivative with respect to $R_t$. The dynamic model can be estimated directly as explained above. For parsimony, we follow a similar parameterization of the dynamic evolution of $f_t^c$ as in the DCC model and assume

$$f_{i,t+1}^c = \omega^c + A^c s_{i,t}^c + B^c f_{i,t}^c,$$  \hfill (3.22)

where $A^c, B^c \in \mathbb{R}$ are scalars, and $\omega^c$ is an $n(n - 1)/2$ vector. To reduce the number of parameters in the maximization, we obtain $\omega^c$ from the hypersphere transformation of the unconditional correlation matrix of the transformed data. All remaining parameters are estimated by maximum likelihood. Inference is carried out by taking the negative inverse Hessian of the log likelihood at the optimum as the covariance matrix for the estimator.
Chapter 4

Measuring Credit Risk in a Large Banking System: Econometric Modeling and Empirics

4.1 Introduction

We propose a new approach to measure the credit risk in a large system of European financial institutions, based on the time-varying probability of simultaneous failure of multiple financial institutions. Such joint failures are akin to financial crises when the banking sector is in distress. Our measures for joint financial firm failure are based on a dynamic multivariate Generalized Hyperbolic skewed-$t$ (GHST) density that allows for skewed and heavy-tailed changes in the market value of financial firms’ equity. The model incorporates dynamic volatilities and failure dependence, while being consistent with expectations about firms’ marginal probabilities of failure at each point in time. By applying the new model to the data of European large financial institutions, we show that the model works well even when the cross-sectional dimension is large. Since the model can be treated as a statistical factor model, it can also be used to explore the possible economic variables driving the variation in the default dependence structure.

The systemic risk or the joint default probability of financial institutions has drawn considerable attention since the recent global financial crisis. How to measure the systemic risk and safeguard the financial system during periods of stress has become the key interest of policy makers. There are several commonly used approaches to measure the systemic risk. The Macro stress tests, such as the 2009 SCAP exercise in the U.S. and the 2010 and 2011 CEBS/EBA stress tests in the E.U., are widely used to assess financial
stability conditions. However, they are expensive to conduct (both in terms of manpower at supervisory agencies as well as at the involved financial institutions), subject to a wide range of political sensitivities, and as a result not suitable for regular financial sector surveillance at monthly frequency. Model-based Banking Stability Measures (BSM) are considered a valuable alternative to more accurate financial stability assessments. The model proposed in this paper yields two financial stability measures related to the conditional and unconditional default probability of a certain percentage of banks in the system at one point in time. Such banking stability measures are currently widely used in central banks and international policy institutions, see for example ECB (2010).

The construction of useful systemic risk or banking stability measures, however, is not straightforward. First, the risk of a systemic event, such as the simultaneous failure of multiple financial firms, usually involves a high cross-sectional dimension, even if only large and possibly systemically important financial institutions are considered. Extending a copula or multivariate density model beyond, say, five time series is difficult. Second, the failure dependence among financial institutions is time-varying. In particular, the interconnectedness across financial firms appears to be stronger during times of turmoil. For example, fire-sale externalities may connect financial firms through market prices in bad times even in the absence of direct business links, see for example Lorenzoni (2008), Brunnermeier and Pedersen (2009), and Korinek (2011). As a result, taking into account higher correlations during times of stress, in addition to higher marginal risks, is an important feature of financial systemic risk. We overcome the two problems of a high dimension and time-varying parameter values by proceeding in two steps. First, we separate the univariate from the multivariate analysis, as in Engle (2002). Second we impose a parsimonious equicorrelation structure into our dynamic density, similar to the approach taken by Engle and Kelly (2012). The parsimonious structure then ensures that the computations remain tractable. The time variation in volatility and correlation parameters is modeled following the Generalized Autoregressive Score (GAS) framework of Creal et al. (2011), and Chapter 2. In our setting, the scaled score of the local log-likelihood drives the dynamic behavior of the time-varying parameters. As a result, the log-likelihood is available in closed form and can be easily maximized.

Two studies in particular relate to our construction of financial stability measures. In each case, the banking system is seen as a portfolio of financial intermediaries whose
4.1. INTRODUCTION

Multivariate dependence structure is inferred from equity returns. Garcia Pascual et al. (2006) assess financial failure in a Gaussian factor model setting. The determination of joint failure probabilities is in part based on the notion of an \( n \)th-to-default CDS basket, which can be set up and priced as suggested in Hull and White (2004). Alternatively, Segoviano and Goodhart (2009) propose a non-parametric copula approach. Here, the banking system’s multivariate density is recovered by minimizing the distance between the so-called banking system multivariate density and a parametric prior density subject to tail constraints that reflect individual failure probabilities. We regard each of these approaches as polar cases, and attempt to strike a middle ground. The proposed GAS framework in our current paper retains the ability to describe the salient equity data features in terms of skewness, fat tails, and time-varying correlations (which the Gaussian copula fails to do), and in addition retains the ability to fit a cross-sectional dimension larger than fifteen (which the non-parametric approach fails to do due to computational problems). In addition, and for the first time, parameter non-constancy is addressed explicitly in our new modeling setup. The two above approaches are inherently static, and rely on a rolling window approach to capture time variation in parameters. By contrast, we model the parameter dynamics explicitly in a parsimonious way.

The remainder of the paper is structured as follows. Section 4.2 introduces a framework for simultaneous failures of financial sector firms. The econometric framework is introduced in Section 4.3 and two new risk measures are proposed in Section 4.4. Section 4.5 presents empirical results on the likelihood of joint failures of large financial institutions in the European Union. In Section 4.6, we study the explanatory power of a few economic variables in understanding the equity correlation dynamics. Section 4.7 concludes.
4.2 A framework for simultaneous financial firm failures

The structural approach due to Merton (1974) and Black and Cox (1976) is probably the most widely used approach for the estimation of individual firms’ credit risk. In this firm value framework, a firm’s underlying asset value evolves stochastically over time, and default is triggered if the firm’s asset value falls below a certain default threshold. This threshold is in general determined by a firm’s current liability structure. It is straightforward to extend the basic premise of the Merton model to a portfolio credit risk setting. In the case of multiple firms, however, the assumptions regarding the correlation (more generally, dependence) structure between the firm value processes are important for overall risk.

First, consider the simple case of two firms $i = 1, 2$, whose asset values $S_{i,t}$ follow

$$dS_{i,t} = S_{i,t}(\mu_i dt + \sigma_i dW_{i,t}),$$

(4.1)

where $W_{i,t}$ is a standard Brownian Motion, $\mu_i$ and $\sigma_i^2$ are drift and variance parameters, respectively, and $dW_{1,t}W_{2,t} = \rho dt$. The solution to Equation (4.1) is

$$S_{i,t} = S_{i,0} \exp \left[ (\mu_i - \sigma_i^2/2) t + \sigma_i W_{i,t} \right].$$

(4.2)

If $\log S_{i,0} = 0$, the log asset values are normally distributed as

$$y_{i,t} = \log S_{i,t} \sim N \left[ (\mu_i - \sigma_i^2/2) t, \sigma_i^2 t \right].$$

(4.3)

The use of Brownian Motions and Gaussian distributions has been popular in the literature for modeling asset returns. However, the conditions of Brownian Motions and the log-normal distribution are too restrictive for financial datasets. The asset returns are usually skewed and heavy-tailed, with time-varying (co)variances. The price process does not have a continuous path as the Brownian Motion, but is identified as a semi-martingale with jumps (Cont and Tankov (2004)). To incorporate these empirical features, the Generalized Hyperbolic (GH) Lévy process has gained more attention as a replacement for the Gaussian assumption. The GH distributions are infinitely divisible (Barndorff-Nielsen and Halgreen (1977)) and every member of this family can generate a Lévy process that is a semimartingale. We focus on the GH skewed-$t$ distribution in this paper, which is an
4.2. A FRAMEWORK FOR SIMULTANEOUS FINANCIAL FIRM FAILURES

asymmetric version of the Student’s $t$ distribution. Our analysis can be easily extended to several other GH distributions. Eberlein (2001) provides a useful survey on asset pricing models under the GH Lévy process assumption.

We write the firm values in a Lévy process framework as in Bibby and Sørensen (2001),

$$dS_{i,t} = \frac{1}{2}v(S_{i,t})[\log(f(S_{i,t})v(S_{i,t}))]'dt + \sqrt{v(S_{i,t})}dW_{i,t},$$

with $v(S_{i,t})$ and $f(S_{i,t})$ two continuously differentiable strictly positive real functions defined on $\mathbb{R}$. Following the arguments in Bibby and Sørensen (2003), we can find suitable functions for a prescribed marginal distribution, for instance a GH skewed-$t$ distribution. The asset value becomes

$$S_{i,t} = S_{0,t}\exp(L_{i,t}),$$

where $L_i$ is a Generalized Hyperbolic Skewed-$t$ Lévy process and the log asset values are Generalized Hyperbolic Skewed-$t$ distributed at discrete time intervals as

$$y_{i,t} = \log S_{i,t} \sim \text{GHST}(\tilde{\sigma}_{i,t}^2, \gamma_i, \nu).$$

Compared to the Student’s $t$ distribution, the GHST distribution is an asymmetric distribution with $\gamma_i$ as the skewness parameter. It is flexible enough to capture the most interesting data features with a limited set of parameters. The dynamic version of the GH distribution proposed in Chapter 2 can accommodate in addition the time-varying covariance matrices. In this paper we adopt the same framework, which is now used to model the correlated defaults in a large portfolio.

In the Merton model and also in our paper, a borrower $i$ defaults at time $t$ if $y_{i,t}$ falls below the firm specific default threshold $y_{i,*}$. Therefore, at time $t$, the firm’s marginal probability of default $p_{i,t}$ is given by

$$p_{i,t} = F(y_{i,t}^*),$$

where $F(\cdot)$ is the cumulative distribution function (CDF) of a standard univariate GHST distribution. Similarly, the joint default probability of two borrowers is

$$p_{1&2,t} = F_{\rho}\left(y_{1,t}^*, y_{2,t}^*\right),$$

where $F_{\rho}$ is the bivariate standard GHST distribution function with correlation $\rho$. 


If an estimate of a firm’s marginal default probability is available, say from Moody’s KMV EDF estimates, then (4.5) implicitly defines the corresponding threshold value \( y^*_i,t \).

With these thresholds, we are able to determine a distress region for the multivariate distribution. A firm defaults at time \( t \) when its asset value \( y_{i,t} \) fall into the region \((-\infty, y^*_i,t)\).

In this paper, we adopt EDF estimates as the estimated probability of default.

4.3 The model

4.3.1 The Dynamic GH skewed-\( t \) model

The risk measure we propose is the joint default probability for a large portfolio of \( k \) banks. In the multivariate case, the joint default probability can be inferred from the market by considering the interrelationship of equity returns. We assume the equity returns \( y_t = (y_{1,t}, \cdots, y_{k,t})' \) follow a multivariate dynamic Generalized Hyperbolic skewed-\( t \) (GHST) distribution. The GHST distribution can be obtained as a normal mean-variance mixture

\[
y_t = (s_t - \frac{\nu}{\nu - 2})\tilde{L}_t\gamma + \sqrt{s_t}\tilde{L}_t\epsilon_t, \tag{4.7}
\]

with a scalar random variable \( s_t \sim \text{InverseGamma}(\nu/2, \nu/2) \) where \( s_t \) is independent of \( \epsilon_t \), and \( k \)-dimensional \( \epsilon_t \sim N(0, I_k) \), and \( \tilde{L}_t \) is an \( k \times k \) loading matrix which defines the individual exposures to the common risk factor \( \epsilon_t \). The mixing structure introduces non-trivial clustering in the tails compared to the situation with only a Gaussian factor \( \epsilon_t \).

The GHST density of \( y_t \) is given by

\[
p(y_t; \tilde{\Sigma}_t, \gamma, \nu) = \frac{\nu^{\nu/2}\pi^{-k/2}}{K_{\nu/2}(\sqrt{\det(\tilde{\Sigma}_t)/s_t})} \cdot \frac{K_{\nu/2}\left(\sqrt{d(y_t)} \cdot (\gamma')\right)}{d(y_t)^{\nu/4+1} \cdot (\gamma')^{-\nu/4}}, \tag{4.8}
\]

\[
d(y_t) = \nu + (y_t - \tilde{\mu}_t)'\tilde{\Sigma}_t^{-1}(y_t - \tilde{\mu}_t), \tag{4.9}
\]

\[
\tilde{\mu}_t = -\frac{\nu}{\nu - 2}\tilde{L}_t\gamma, \tag{4.10}
\]

where \( K_a(b) \) is the modified Bessel function of the second kind, \( \tilde{\Sigma}_t = \tilde{L}_t\tilde{L}_t' \) is the scale matrix, see Bibby and Sørensen (2003).

\[
\tilde{L}_t = L_tT, \tag{4.11}
\]

\[
(T'\tilde{T})^{-1} = \frac{\nu}{\nu - 2}I + \frac{2\nu^2}{(\nu - 2)^2(\nu - 4)}\gamma'\gamma', \tag{4.12}
\]
4.3. THE MODEL

The matrix \( L_t \) characterizes the time-varying covariance matrix \( \Sigma_t = L_t L_t' \). We consider the time-varying covariance matrix of \( y_t \) as

\[
\Sigma_t = L_t L_t' = D_t R_t D_t,
\]

where \( D_t \) is a diagonal matrix holding the volatilities of \( y_{i,t} \) and \( R_t \) is the correlation matrix of equity returns \( y_t \). The marginal distribution for a multivariate Generalized Hyperbolic skewed-\( t \) distribution is a univariate Generalized Hyperbolic skewed-\( t \) distribution. The skewness variables can be different in each marginal.

We assume the dynamic covariance matrix \( \Sigma_t \) depends on the unobserved factor \( f_t \), where \( f_t \) follows the Generalized Autoregressive Score process as specified in Creal, Koopman and Lucas (2011, 2012) and Chapter 2.

\[
f_{t+1} = \omega + \sum_{i=0}^{p-1} A_i s_{t-i} + \sum_{j=0}^{q-1} B_j f_{t-j},
\]

\[
s_t = S_t \nabla_t,
\]

\[
\nabla_t = \frac{\partial \ln p_{GH}(y_t|\mathcal{F}_{t-1}; f_t, \theta)}{\partial f_t},
\]

\( \omega \) is a vector of fixed intercepts, and \( A_i \) and \( B_j \) are fixed parameter matrices. In order to obtain our result below, we define

\[
\text{vec}(L) = \mathcal{D}_k^0 \text{vech}(L)
\]

for a \( k \times k \) lower triangular matrix \( L \),

\[
\text{vech}(S) = \mathcal{B}_k \text{vec}(S)
\]

for a symmetric matrix \( S \), and the commutation matrix \( C_k \) for an \( k \times k \) matrix \( X \) as

\[
\text{vec}(X) = C_k \text{vec}(X').
\]

**Result 1** If \( y_t \) follows a GHST distribution \( p(y_t; \tilde{\Sigma}_t, \gamma, \nu) \), where the time-varying covari-
The dynamic score is

\[
\nabla_t = \Psi_t' \text{vec} \left( w_t \cdot y_t' - \tilde{\Sigma}_t - \left( 1 - \frac{\nu}{\nu - 2} w_t \right) \tilde{L}_t \gamma_t' \right),
\]

\[
w_t = \frac{\nu + k}{2d(y_t)} \frac{k'_{\nu+k}(\sqrt{d(y_t)} \cdot (\gamma' \gamma))}{2\sqrt{d(y_t)/(\gamma' \gamma)}},
\]

\[
\Psi_t = \frac{\partial \text{vech}(\Sigma_t)' \partial f_t}{\partial f_t},
\]

\[
H_t = (\tilde{\Sigma}_t^{-1} \otimes \tilde{\Sigma}_t^{-1})(\tilde{L}_t \otimes I_k)(T' \otimes I_k)D_k^0(\mathcal{B}_k(I_k2 + C_k)(L_t \otimes I_k)D_k^0)^{-1},
\]

where we define \( k_{\nu+k}(\cdot) = \ln K_{\nu+k}(\cdot) \) with first derivative \( k'_{\nu+k}(\cdot) \). The matrices \( \Psi_t \) and \( H_t \) are time-varying, parameterization specific, and depend on \( f_t \), but not on the data.

The dynamics driven by the score \( \nabla_t \) can be seen as a local improvement of the likelihood to the new data observed at time \( t \), and \( S_t \) is a scaling matrix for the score \( \nabla_t \). Typical choices for the scaling matrix \( S_t \) are the unit matrix or inverse (powers) of the Fisher information matrix \( \mathcal{I}_t \), where

\[
\mathcal{I}_t = \text{E} [\nabla_t \nabla_t' | y_{t-1}, y_{t-2}, \ldots].
\]

For example, \( S_t = \mathcal{I}_t^{-1} \) accounts for the curvature in the score \( \nabla_t \). With the choice of scaling matrix as the inverse Fisher information matrix, the GAS step \( s_t \) can be seen as a Gauss-Newton improvement step of the local fit of the model. As the Fisher information matrix for the GH distribution has no analytical expression, we instead use the inverse Fisher information matrix from the Student’s \( t \) in our current paper. Chapter 2 demonstrate that this results in a stable model that outperforms alternative models if the data are fat-tailed and skewed. We obtain

\[
S_t = \left\{ \Psi_t'(I \otimes \tilde{L}_t^{-1})' [gG - \text{vec}(I) \text{vec}(I)' (I \otimes \tilde{L}_t^{-1}) \Psi_t] \right\}^{-1},
\]

where \( g = (\nu + k)/(\nu + 2 + k) \), and \( G = \text{E}[x_t x_t' \otimes x_t x_t'] \) for \( x_t \sim N(0, I_k) \).

### 4.3.2 Estimation and restrictions

Chapter 2 show that the GAS dynamic structure has superior performance under skewed and fat-tailed distributions. However, evaluating the full covariance matrix in the full likelihood is cumbersome computationally if the dimension of the data is large. Therefore, we separate the estimation of the covariance matrix into volatility estimation and
correlation estimation procedures. The algorithm works in two steps.

1. Estimate the log-volatility $\log(\sigma_t)$ for each series with a univariate dynamic GHST model. The skewness parameter is estimated for each series separately, but the kurtosis parameter is fixed at 5. The motivation is to ensure that the marginal GHST distributions are internally consistent with the multivariate GHST distribution. The data at time $t$ is standardized by the volatility $\sigma_t$. The standardized data is tested for serial correlation using the F-test suggested in Engle (2002).

2. Estimate the correlation matrix $R_t$ of the standardized returns using the volatilities from the first step. The correlation matrix is driven by the factor $f_t$ from the multivariate dynamic GHST model. Again the kurtosis parameter is set ex ante as $\nu = 5$ and the skewness parameters are equal to those from the univariate distributions obtained from the first step. We need a parametrization as in Engle (2002) or Chapter 2 to ensure that $R_t$ actually is a correlation matrix.

In the univariate and the multivariate GH skewed-$t$ model, we fix the degrees of freedom parameter for all the marginal distributions at five. We can also estimate a GHST distribution in order to obtain a sensible degree of freedom. Interestingly when estimate static GHST model in a exploratory analysis, we find five a reasonable parameter value that ensures the distribution captures the tail behavior of the data.

The idea behind the algorithm is simple. We first use the dynamic GHST model as a filter for the volatility in the equity returns for each of the series. The standardized equity returns are then used in a multivariate dynamic GHST model model, where the covariance matrix is the correlation matrix. It is similar to the two-step procedure or the composite likelihood method in Engle (2002), Hu (2005), and other studies that are based on a multivariate GARCH framework.

If we want to work with a large dimensional dataset, we still need to impose some further restrictions to confront the computational difficulties. One difficulty arises from estimating the unconditional mean $\omega$ in Equation (4.14). In a dataset of $k$ time series, we have to estimate $k(k - 1)/2$ coefficient for the unconditional mean of factors $\omega$. In order to reduce the computational difficulty, we estimate the unconditional mean of the factors
\( \bar{f} \in \mathbb{R}^{k(k-1)/2} \) separately and estimate a scalar \( \omega \) in the equation (4.14),

\[
f_{t+1} = \omega \bar{f} + \sum_{i=0}^{p-1} A_i s_{t-i} + \sum_{j=0}^{q-1} B_j f_{t-j};
\]

(4.25)

The \( \omega \) is now defined as the levels of correlation coefficients proportional to the unconditional mean of our factors. We choose \( A \) and \( B \) as scalar parameters as in the DCC model. This reduces the total number of parameters in GAS model to three only, irrespective of the data’s cross-sectional dimension. In practice, we can also fix \( \omega \) at one, because the parameter estimate is usually close to one. It is sometimes called “correlation targeting” in the literature.

One of the attractive features of the GAS model is the possibility to introduce a latent factor structure to describe the time variation in the dynamic parameters we are interested in. We could impose the restriction that several time-varying parameters are driven by common factors. This is extremely useful to process high-dimensional data from a large system. In the next section, we introduce the block GAS-Equicorrelation model and the GAS-Equicorrelation model as two examples of such a framework.

4.3.3 The Block GAS-Equicorrelation model

With the two-step estimation procedure, the task of maximizing the multivariate GHST likelihood in a large system becomes more feasible. The computational burden is largely reduced due to the separation of the likelihood for volatilities and correlations. Still, this method is cumbersome if the data dimension becomes high, for instance around 100. The advantage of the factor structure in the GAS framework (4.14) underlying the dynamic correlation matrix makes it possible to address this problem by using common factors. We assume the factor dimension to be smaller than the number of correlations. This defines a multi-factor structure underlying the dynamic correlation model. In the literature, we call correlation matrices with such a structure a block dynamic equicorrelation matrix. Assume that \( k \) firms fall into \( m \) different groups according to their exposure to a common systematic risk factor. Firms have equicorrelation \( \rho_i^2 \) within each group and \( \rho_i \cdot \rho_j \) between groups \( i \) and \( j \). So we have \( k = n_1 + n_2 + \cdots + n_m \) random variables that follow a GH distribution with a correlation matrix that has a block equicorrelation structure, where
4.3. THE MODEL

\( n_i \) denotes the number of firms in group \( i \). The correlation matrix at time \( t \) is given by

\[
\mathbb{R}_t = \begin{bmatrix}
(1 - \rho_{1,t}^2)I_{n_1} & \cdots & \cdots & 0 \\
0 & (1 - \rho_{2,t}^2)I_{n_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & (1 - \rho_{m,t}^2)I_{n_m}
\end{bmatrix} + \begin{pmatrix}
\rho_{1,t} \ell_{1} \\
\rho_{2,t} \ell_{2} \\
\vdots \\
\rho_{m,t} \ell_{m}
\end{pmatrix} \cdot \begin{pmatrix}
\rho_{1,t} \ell_{1}' \\
\rho_{2,t} \ell_{2}' \\
\vdots \\
\rho_{m,t} \ell_{m}'
\end{pmatrix},
\]

(4.26)

where \( \ell_i \in \mathbb{R}^{n_i \times 1} \) is a column vector of ones and \(|\rho_{i,t}| < 1\) to ensure the positive-definiteness of \( \mathbb{R}_t \). The matrix \( L_t \) and the inverse of \( L_t \) can be calculated explicitly by assuming

\[
L_t = \begin{bmatrix}
a_{1,t}I_{n_1} & \cdots & \cdots & 0 \\
0 & a_{2,t}I_{n_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & a_{m,t}I_{n_m}
\end{bmatrix} + \begin{pmatrix}
b_{11,t}J_{11} & b_{12,t}J_{12} & \cdots & b_{1m,t}J_{1m} \\
b_{12,t}J_{21} & b_{22,t}J_{22} & \cdots & b_{2m,t}J_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
b_{1m,t}J_{m1} & b_{2m,t}J_{m2} & \cdots & b_{mm,t}J_{mm}
\end{pmatrix},
\]

(4.27)

where \( J_{ij} \in \mathbb{R}^{n_i \times n_j} \) is a matrix of ones \( J_{ij} = \ell_i \ell_j' \). We solve for all the parameters in the equation \( R_t = L_t L_t' \), where \( L_t \) is symmetric. The block equicorrelation model allows us to obtain analytical solutions for the determinant of \( R_t \). As a result of the Matrix Determinant Lemma (see Harville (2008)), the determinant of the matrix \( R_t \) is

\[
\det(R_t) = \det(\Xi_t + u_t u_t') = (1 + u_t' \Xi_t^{-1} u_t) \det(\Xi_t)
\]

\[
= \left[ 1 + \frac{n_1 \rho_{1,t}^2}{1 - \rho_{1,t}^2} + \cdots + \frac{n_m \rho_{m,t}^2}{1 - \rho_{m,t}^2} \right] (1 - \rho_{1,t}^2)^{n_1} \cdots (1 - \rho_{m,t}^2)^{n_m},
\]

with \( \Xi_t \) the diagonal matrix in the first term on the righthand side of (4.26) and \( u_t \) the vector in the second term, such that \( R_t = \Xi_t + u_t u_t' \). The determinant of matrix \( L_t \) is easy to find as the square root of this value. The analytic expressions facilitate the computation of the likelihood and GAS steps in high dimensions. The time-varying correlation coefficients \( \rho_{1,t}, \cdots, \rho_{m,t} \) are driven by the GAS factors from a GH skewed-\( t \) distribution. We can derive the GAS model with these restrictions.

**Result 2** If \( y_t \) follows a GH skewed-\( t \) distribution and the time-varying correlation matrix \( R_t \) has a block equicorrelation structure, the dynamic score follows Equation (4.20) and the matrix \( H_t \) stays the same as Equation (4.23). We denote the time-varying parameters in \( R_t \) as \( \Phi_t = (\rho_{1,t}, \cdots, \rho_{m,t})' = f_t \). The major difference is \( \frac{\partial \text{vec}(R_t)'}{\partial f_{i,t}} \) as part of \( \frac{\partial \text{vec}(\Sigma)'}{\partial f_{i,t}} \) in
CHAPTER 4. CREDIT RISK IN A LARGE BANKING SYSTEM

\[ \Psi_t, \]

\[
\frac{\partial \text{vec}(R_t)'}{\partial f_{i,t}} = -2\rho_{i,t} \cdot \text{vec} \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \text{I}_{n_i} & 0 \\ 0 & 0 & \ldots & 0 \end{bmatrix} \left( \begin{bmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{bmatrix} \otimes \begin{bmatrix} \ell_{n_i} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right) + \begin{bmatrix} \ell_{n_i} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{bmatrix} \right]
\]

(4.28)

The simplest case of the block GAS-Equicorrelation model is if we only have one block, which we call the GAS-Equicorrelation model. Following Engle and Kelly (2012), we then assume the correlation matrix \( R_t \) with the equicorrelation structure:

\[ R_t = (1 - \rho_t)\text{I} + \rho_t \ell \ell', \]

(4.29)

where \( \rho_t \in (\frac{-1}{N-1}, 1) \). Under such an assumption, the dynamic score equation stays the same as (4.20), but the matrix computations are simplified.

**Result 3** If we assume one equicorrelation structure for the correlation matrix, the GAS model works as in the equations in Section 4.3.1. The only difference is that \( \Psi_t \) simplifies to:

\[
\Psi_t = \left( \frac{d a_t}{d \rho_t} \text{vec}(1_k) + \frac{d a_t}{d \rho_t} d + \frac{d b_t}{d \rho_t} c + \frac{d b_t}{d \rho_t} k d k' \right) d'(f_t),
\]

(4.30)

\[
\frac{d a_t}{d \rho_t} = -\frac{1}{2\sqrt{1-\rho_t}},
\]

(4.31)

\[
\frac{d b_t}{d \rho_t} = \frac{1}{2k} \left( \frac{k-1}{\sqrt{1-\rho_t} + k\rho_t} + \frac{1}{\sqrt{1-\rho_t}} \right),
\]

(4.32)

where the scalar \( c = \frac{1}{\sqrt{\mu}}, d = \frac{\sqrt{\mu} - \sqrt{\mu^2/\gamma + \mu\sigma^2}}{\sqrt{\mu}}, \mu_\gamma = \frac{\nu}{\nu-2}, \) and \( \sigma^2 = \frac{2\nu^2}{(\nu-2)^2(\nu-4)} \).

The GAS-Equicorrelation model may seem too restrictive at first. In our application, however, the data we are dealing with are European financial institutions that have strong economic and financial links and the equicorrelation captures our salient parameter of interest: the systemic dynamic correlation in the entire system of banks considered. We compare the equicorrelation model with the full GAS model in Section 4.5.1 for a small system where we can still estimate both models. For the large system with more than 70 institutions, we only consider the GAS-Equicorrelation version of the model.
4.4 The risk measures in a large system

There are multiple ways to construct a financial sector stability measure. For example, a higher probability of at least a certain number of firms failing over the next year is a natural measure of systemic risk. Such a measure is for example constructed and tracked in the European Central Bank’s biannual Financial Stability Report, see for example ECB (2010). Here we use the same definition of a systemic risk measure. After estimating the conditional covariance matrix through the dynamic-GH model, the time-varying correlation and volatility mechanism are used to calculate the probability of failure of European financial firms. With this estimated multivariate density, we can thus produce a systemic risk measure. In this section, we calculate this measurement either by simulation or by analytic approximations. The latter are particularly useful for large cross-sectional dimensions.

The straightforward approach is based on simulations of equity returns. As discussed in Section 4.2, a firm default may happen if the equity return is too negative compared to pre-specified default threshold. In the multivariate distribution, these thresholds define a distress region. We can generate simulations and compute tail probabilities by counting the number of realizations in this pre-determined distress region. In this paper, we simulate from the estimated dynamic multivariate GHST distribution. The distress region is determined by the default thresholds transformed from Moody’s EDF estimates. This simulation based method is general enough for all different distributions and model specifications.

When the dimension of the dataset becomes too large, the simulation based risk measurements become inefficient. We need a large number of simulations. Interestingly, we are able to explore the advantage of the equicorrelation structure for the simplified correlation matrix. This is the alternative approach to produce the systemic risk in a large system. We consider the system of banks as homogenous portfolio of equities.\footnote{The homogeneity assumption is only used for exposition. Different $\gamma_i$ and $\rho_i$ in the block equicorrelation structure can easily be allowed for.} We can use a Law of Large Number (LLN) result in the context of credit risk as in Lucas et al. (2001). We define the Systemic Risk indicator as the probability that a certain number
of banks default in the same timespan. The number of defaults at time $t$ is

$$c_{N,t} = \frac{1}{N} \sum_{i=1}^{N} 1\{y_{i,t} < y_{i,t}^*|\kappa_t, s_t\}. \quad (4.33)$$

Given that the $1\{y_{i,t} < y_{i,t}^*\}$'s are conditionally independent, the Law of Large Numbers tells us if $N \to +\infty$,

$$c_{N,t} \approx \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}(1\{y_{i,t} < y_{i,t}^*|\kappa_t, s_t\}) \quad (4.34)$$

$$= \frac{1}{N} \sum_{i=1}^{N} \mathbb{P}(y_{i,t} < y_{i,t}^*|\kappa_t, s_t). \quad (4.35)$$

If the returns are GHST distributed and have a block equicorrelation structure as equation (4.26), we can model the banks' market values as:

$$y_t = (\zeta_t - \mu_t)\gamma + \sqrt{\sigma_t}z_t, \quad (4.36)$$

$$z_t = \eta_t\kappa_t + \Lambda_t\epsilon_t, \quad (4.37)$$

where $\kappa_t \sim N(0, 1)$ and $\epsilon_t \sim N(0, \mathbb{I}_k)$, $\eta_t$ is a vector of parameters $(\eta_{1,t}, \cdots, \eta_{k,t})'$, and $\Lambda_t$ is an $k \times k$ diagonal matrix with $(\lambda_{1,t}, \cdots, \lambda_{m,t})'$ on the diagonal. We are interested in finding the values of $\eta_t$ and $\Lambda_t$ such that $\text{Var}(z_t) = \mathbb{R}_t$. We know

$$\text{Var}(y_t) = \mathbb{R}_t + u_t u_t'$$

$$= \mu_t\Lambda_t^2 + \mu_t^2\eta_t^2 + \sigma_t^2\gamma\gamma'.$$ \quad (4.38)

So the parameters $\eta_t$ and $\Lambda_t$ should satisfy the following equations,

$$\lambda_{i,t} = \sqrt{1 - \rho_{i,t}^2}, \text{ for } i = 1, \cdots, m, \quad (4.39)$$

$$\mu_t\eta_t' = u_t u_t' - \sigma_t^2\gamma\gamma'. \quad (4.40)$$

This is a two-factor model with a common Gaussian factor $\kappa_t$ and a mixing factor $\zeta_t$. The stability measure in this setting is given by

$$p_t = \mathbb{P}(C_{n,t} > c_{p,t}), \quad (4.41)$$
where we can compute the measure conditional on the latent factors $\kappa_t$ and $\zeta_t$,

$$c_{p,t} = \frac{1}{k} \sum_{i=1}^{k} P[y_{i,t} < y_{i,t}^* | \kappa_t, \zeta_t],$$

$$P[y_{i,t} < y_{i,t}^* | \kappa_t, \zeta_t] = \Phi \left( \frac{(y_{i,t}^* + \mu_i \gamma_t - \zeta_t \gamma_t)/\sqrt{\eta_{i,t} \kappa_t}}{\lambda_{i,t}} \right).$$

The risk measure is related to the number of defaults as a proportion in the portfolio. Using equation (4.42), we rewrite the threshold common factor $\kappa_t = \kappa_t^*(c_p, \zeta)$ as a function of the default proportion $c_{p,t}$ and the mixing variable $\zeta_t$. We are able to compute the joint default probability numerically as

$$p_t = P(C_{N,t} > c_{p,t}) = \int P(\kappa_t < \kappa_t^*(c_{p,t}, \zeta_t)) p(\zeta_t) d\zeta_t.$$  (4.44)

Similarly, we can compute the probability of certain proportion $c_{p,t}^{-1}$ of the system excluding bank $i$ defaulting conditional on the event that bank $i$ fails.

$$P(C_{N^{-1},t} > c_{p,t}^{-1} | y_{i,t} < y_{i,t}^*) = \frac{P(C_{N^{-1},t} > c_{p,t}^{-1}, y_{i,t} < y_{i,t}^*)}{P(y_{i,t} < y_{i,t}^*)} = \frac{\int \Phi_2 \left( \frac{z_{i,t}^{*\gamma_t} \sqrt{\eta_{i,t}}}{\sqrt{1-\sigma_i^2 \gamma_t}}, \kappa_t^*(c_{1,t}^{*}, \zeta_t), \eta_{i,t} \right) p(\zeta_t) d\zeta_t}{\int P(\kappa_t < \kappa_t^*(c_{1,t}^{*}, \zeta_t)) p(\zeta_t) d\zeta_t},$$

where

$$z_{i,t}^{*} = \frac{y_{i,t}^* - (\zeta_t - \mu_i) \gamma_t}{\sqrt{\eta_{i,t}}}$$

from Equation (4.36), $\Phi_2(\cdot, \cdot, \eta_{i,t})$ is the bivariate normal CDF with correlation $\eta_{i,t}$, and $\kappa_t^*(c_{1,t}^{*}, \zeta_t)$ denotes the corresponding threshold common factor when bank $i$’s equity return fall below the threshold $y_{i}^*$. This conditional probability is close to the Multivariate extreme spillovers indicator of Hartmann et al. (2005).

We define the average of this conditional default probability over $N$ financial firms as the Systemic Risk Measure (SRM), as it measures the possibility that an individual credit event increases the level of systemic risk. We apply the two measurements proposed here in the empirical section.
4.5 Empirical application

In this section, we compute the banking stability measure in the European Union. We observe 73 major financial groups with complex interactions. The data contain monthly observations of equity prices and estimated EDFs for all 73 financial institutions. Our whole sample covers the period January 1992 to June 2010, but with missing observations of several names in the beginning of the sample. Dealing with missing values in our model’s setting is straightforward. Both the likelihood and the score steps in the dynamic GHST model adapt automatically if data are not observed at particular times and there are no sample selection issues.

The analysis in this section consists of two parts. To compare the dynamic GHST model with the block GAS Equicorrelation models, we choose a subsample consisting of ten European banks. The full multivariate model from Section 4.3.1 is estimated with a time-varying covariance matrix. We also show the estimation results for models in Section 4.3.3. These results are presented in Section 4.5.1. Second, we impose the GAS Equicorrelation structure in the dynamic GHST model for the whole sample of 73 financial institutions. The conditional Law of Large Numbers approximation is implemented to compute the Banking Stability Measure and the Systemic Risk Measure. Section 4.5.2 includes the results for this analysis.

4.5.1 The system of major European banks

In our first analysis, we select a geographically diversified sub-sample of 10 banks in the Euro Area: Bank of Ireland, BBVA, Santander, BNP Paribas, Commerzbank, Deutsche Bank, Societe Generale, ING, UniCredit, National Bank of Greece. To estimate the time-varying correlations and volatilities, we use monthly log returns from January 1994 to June 2010 from Bloomberg. The dataset contains 198 observations for each series. The EDF data used to compute the distress thresholds are provided by Moody’s KMV. From the descriptive statistics in Table 4.1 we see that all equity returns are skewed and fat-tailed. Commerzbank and ING Group stand out with a pronounced skewness of -1.10 and -1.64, and a kurtosis of 8.33 and 6.99, respectively. However, the Bank of Ireland has a large kurtosis of 16.053. We model the equity returns from all 10 banks with our skewed and heavy-tailed dynamic GH skewed-$t$ model.

We first estimate the full correlation matrix with forty-five pair-wised dynamic corre-
4.5. EMPIRICAL APPLICATION

Table 4.1: Sample Descriptive Statistics.

The descriptive statistics for the monthly equity returns between January 2000 and June 2010. The sample mean values are all very close to zeros. The standard deviations, minimum and maximum values are multiplied by 100 respectively in the table. All skewness and excess kurtosis are significantly different from 0.

<table>
<thead>
<tr>
<th>Bank</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of Ireland</td>
<td>0.000</td>
<td>1.309</td>
<td>-0.594</td>
<td>16.053</td>
<td>-113.917</td>
<td>106.153</td>
</tr>
<tr>
<td>BBVA</td>
<td>0.000</td>
<td>0.710</td>
<td>-0.512</td>
<td>3.220</td>
<td>-38.894</td>
<td>37.003</td>
</tr>
<tr>
<td>Santander</td>
<td>0.000</td>
<td>0.720</td>
<td>-0.725</td>
<td>3.758</td>
<td>-40.720</td>
<td>37.609</td>
</tr>
<tr>
<td>BNP Paribas</td>
<td>0.000</td>
<td>0.675</td>
<td>-0.502</td>
<td>3.261</td>
<td>-34.001</td>
<td>32.959</td>
</tr>
<tr>
<td>Commerzbank</td>
<td>0.000</td>
<td>0.940</td>
<td>-1.101</td>
<td>5.474</td>
<td>-67.779</td>
<td>45.536</td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>0.000</td>
<td>0.760</td>
<td>-0.421</td>
<td>3.906</td>
<td>-46.588</td>
<td>45.444</td>
</tr>
<tr>
<td>Societe Generale</td>
<td>0.000</td>
<td>0.777</td>
<td>-0.968</td>
<td>4.110</td>
<td>-53.679</td>
<td>29.201</td>
</tr>
<tr>
<td>ING</td>
<td>0.000</td>
<td>0.896</td>
<td>-1.647</td>
<td>8.939</td>
<td>-73.367</td>
<td>45.187</td>
</tr>
<tr>
<td>UniCredito</td>
<td>0.000</td>
<td>0.752</td>
<td>-0.048</td>
<td>3.282</td>
<td>-44.318</td>
<td>36.017</td>
</tr>
<tr>
<td>National Bank of Greece</td>
<td>0.000</td>
<td>0.938</td>
<td>0.336</td>
<td>2.324</td>
<td>-48.178</td>
<td>53.652</td>
</tr>
</tbody>
</table>

The estimated volatility series are plotted in separate panels in Figure 4.1. The volatility estimates are obtained via estimation of the GH skewed-t distribution for each individual time series. All parameters in the volatility models are significant at the 5% significant level, as shown in Table 4.2. From the graph, we see three highly volatile periods corresponding to either financial crises or global economic recessions. The most recent period with clearly high volatility begins in Sept. 2008, when the failure of Lehman Brothers brought down the stock prices of all banks. But the magnitude of this increase differs from one institution to the other. The most volatile time series is the Bank of Ireland’s equity return. In the midst of the Global Financial Crisis, the Irish Banking Crisis hits this largest Irish bank even harder. The Bank of Ireland was recapitalized by the Irish Government in February 2009 and further bailed-out by the ECB and IMF in 2010. The idiosyncratic shock to the Bank of Ireland, on top of the common shock from
the Lehman Brother's bankruptcy, drives up its volatility even higher.

We filter the equity returns with the estimated volatilities and apply a multivariate GH skewed-\( t \) model in the second step. The time-varying correlation matrices are assumed to follow the GAS model in Equations (4.14) and (4.16). We implement four dynamic GHST models imposing different parameterizations on the dynamic correlation matrix.

As a comparison, we estimate the dynamic GH skewed-\( t \) model with the GAS-Equicorrelation model (Equations (4.29)-(4.32)), and the two-Block GAS-Equicorrelation model (Equations (4.26)-(4.28)) on the same sample. The banks are separated into two groups. The first group contains the Bank of Ireland, BBVA, Santander, UniCredito and the National Bank of Greece. The second group includes the rest banks. The correlation estimates are plotted in the bottom panels in Figure 4.2. As benchmarks, we also include the average correlation from the Rolling Window (RW) method with the window size set to 12 months.

If we compare the Equicorrelation model outputs and the average correlation from the GAS model and RW method, the dynamic equicorrelation appears to be an average of the pairwise correlations. The flexible GAS-GHST model allows for more heterogenous dynamics on the pair-wise correlation coefficients. But we also see that the equicorrelation model picks up the most salient comovements in the data, such as the drop of correlation in 2001 and the increase after 2008 due to the financial crisis. In the model estimates from the two-block GAS-Equicorrelation matrix, we see that the three correlation estimates exhibit similar time-varying patterns as the equicorrelation dynamics. But we start to see differences in particular periods, for instance around the year 2008. It seems that the correlation of banks in the first group is higher in the crisis period. We provide the parameter estimates and log-likelihood values from the dynamic correlation models in Table 4.2.

With the estimated GH skewed-\( t \) distributions, either with the full model or with the equicorrelations and block equicorrelations, we can compute the Banking Stability Measure (BSM) and Systemic Risk Measure (SRM) given the default thresholds from inverting the GH skewed-\( t \) CDF at the observed EDF levels. The banking stability measure is defined as the joint probability of three or more banks defaulting. The Systemic Risk Measure is constructed with the conditional statement of two or more banks defaulting given bank \( i \) defaulted. With the estimated multivariate GH skewed-\( t \) distributions, we
### Table 4.2: The Estimation Results: Part I.

The parameter estimated in our GAS-GHST models for ten banks’ equity returns. We use univariate GAS-GHST models for the marginal volatility. With the filtered returns, we estimate three dynamic correlation models: the GAS Equicorrelation model, the Block GAS Equicorrelation model, and the GAS model with full correlation structure. All parameters are significant at the 5% level.

<table>
<thead>
<tr>
<th>Dynamic Volatility</th>
<th>A</th>
<th>B</th>
<th>$\omega$</th>
<th>$\gamma$</th>
<th>Log-lik</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank of Ireland</td>
<td>0.201</td>
<td>0.964</td>
<td>0.093</td>
<td>-0.206</td>
<td>-725.655</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBVA</td>
<td>0.154</td>
<td>0.902</td>
<td>0.220</td>
<td>-0.145</td>
<td>-701.432</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.010)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santander</td>
<td>0.196</td>
<td>0.884</td>
<td>0.256</td>
<td>-0.163</td>
<td>-696.317</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.011)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BNP Paribas</td>
<td>0.212</td>
<td>0.866</td>
<td>0.295</td>
<td>-0.152</td>
<td>-691.252</td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commerzbank</td>
<td>0.168</td>
<td>0.929</td>
<td>0.175</td>
<td>-0.167</td>
<td>-738.160</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.006)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>0.168</td>
<td>0.910</td>
<td>0.211</td>
<td>-0.105</td>
<td>-715.436</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.010)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Societe Generale</td>
<td>0.196</td>
<td>0.918</td>
<td>0.189</td>
<td>-0.134</td>
<td>-711.646</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.008)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ING</td>
<td>0.167</td>
<td>0.915</td>
<td>0.200</td>
<td>-0.224</td>
<td>-719.552</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.007)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UniCredito</td>
<td>0.126</td>
<td>0.969</td>
<td>0.071</td>
<td>-0.064</td>
<td>-708.966</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Bank of Greece</td>
<td>0.141</td>
<td>0.927</td>
<td>0.188</td>
<td>-0.060</td>
<td>-768.016</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.008)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dynamic Correlation</th>
<th>A</th>
<th>B</th>
<th>$\omega_1$</th>
<th>$\omega_2$</th>
<th>$\gamma$</th>
<th>Log-lik</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAS EquiCorr (1)</td>
<td>0.116</td>
<td>0.915</td>
<td>0.205</td>
<td>-0.071</td>
<td>(0.004)</td>
<td>-2050.956</td>
<td>4111.91</td>
<td>4126.45</td>
</tr>
<tr>
<td>GAS EquiCorr (2)</td>
<td>0.070</td>
<td>0.907</td>
<td>0.931</td>
<td>1.417</td>
<td>(0.002)</td>
<td>-2052.116</td>
<td>4114.23</td>
<td>4130.67</td>
</tr>
<tr>
<td>GAS Model</td>
<td>0.027</td>
<td>0.717</td>
<td>1.007</td>
<td>-1.952</td>
<td>(0.001)</td>
<td>-3914.40</td>
<td>3930.84</td>
<td>3953.84</td>
</tr>
</tbody>
</table>
Figure 4.1: Volatility estimations for the banks’ equities

The volatility estimates from the Dynamic GH Skewed-$t$ for all the banks’ stock return data. (BBVA stands for BBV.Argentaria and DB refers to Deutsche Bank.)
The correlation estimates from Dynamic GH Skewed-t model with banks’ stock returns. We selected the correlations of the Bank of Ireland’s with other banks in our sample. The last two panels are from a one-factor and two-factor equicorrelation model in the skewed-t distribution.
can use simulations to compute the risk indicators. We use 10,000,000 simulations at each time $t$ and count the number of banks under stress. As we obtain the simulations directly, we can compute the conditional and unconditional default probabilities. Alternatively if we use the GAS-Equicorrelation model, we can analytically calculate these measures under the LLN approximation suggested in Section 4.4. The analytical calculation is fast and less cumbersome than the simulation method.

From Figure 4.3, we see that the dynamic patterns of the risk indicators are very similar irrespective of the computation method used. The Banking Stability measures simulated/calculated from different correlation models are close to each other. The LLN approximated risk measure somewhat understates the risk in normal times and overestimates the risk in crisis times after the year 2008. This is because the number of banks is as small as 10 in our current setting, which makes the LLN approximation less accurate. Figure 4.4 plots the Systemic Risk Measure proposed in Section 4.4. The simulated (SIM) measure is computed with the straightforward simulation method and the correlation matrix is driven by the estimated GAS model in Result 1. The LLN approximated Systemic Risk Measure is calculated analytically based on the dynamic Equicorrelation estimates. We see the difference in the SRM between these two methods. The approximated SRM with the conditional Law of Large Numbers is always lower than the simulated SRM, but the pattern over time is similar. If we look at the average of the approximated indicator in the last panel, we see a break around the year 2002 in the mean for the analytical SRM. This may be attributed to the introduction of the Euro as a common currency, which tightened the interconnectedness of the European banks.

### 4.5.2 European large financial institutions

The task becomes more challenging with a few European large financial institutions. These financial institutions are large and possibly systemically important, as their failure would likely spread and have adverse implications for financial markets or other financial institutions operating within the system.

The datasets we use are monthly equity returns from 73 financial institutions. These institutions are European banks, insurance companies and investment companies. In Table 4.3, we provide a full list of the names in our sample. The sample skewness and kurtosis for each time series is also included in the table. Most equity return series exhibit
The Banking Stability Measure constructed from the Dynamic GH skewed-$t$ models. A comparison study is provided here with two different correlation assumptions. The top left and bottom left panel contains the BSM with Dynamic Equicorrelation, but the top one is calculated with the analytical computation and the other one is simulated. The top right plot shows the simulated BSM with the full model correlation result. These measures are defined as the probability of three or more firm defaults.
The Systemic Risk Measure constructed from the Dynamic GH skewed-$t$ models. We show the result of simulated SRM with correlation estimates from a Full GAS model, as well as the LLN approximated SRM from a Dynamic Equicorrelation model. The last panel contains the average of the SRM measure over all firms. SRM is defined as the probability of two or more firms defaulting given firm $i$ failing.
negative skewness and fat-tailness.

Table 4.3: Sample Skewness and Kurtosis Statistics.

Descriptive statistics for the CRSP stock returns between January 1970 and June 2010. All observations are monthly log returns. All names are large European financial firms including banks, insurance companies and investment firms.

<table>
<thead>
<tr>
<th>Name</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Name</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKERMANS &amp; VAN HAAREN</td>
<td>-0.10</td>
<td>3.92</td>
<td>DEUTSCHE BANK (XET)</td>
<td>-0.36</td>
<td>6.55</td>
</tr>
<tr>
<td>AEGON</td>
<td>-1.13</td>
<td>6.75</td>
<td>DEUTSCHE BOERSE (XET)</td>
<td>-0.30</td>
<td>3.98</td>
</tr>
<tr>
<td>AGEAS (EX-FORTIS)</td>
<td>-3.78</td>
<td>30.21</td>
<td>DEUTSCHE POSTBANK (XET)</td>
<td>-1.39</td>
<td>8.42</td>
</tr>
<tr>
<td>ALLIANZ (XET)</td>
<td>-0.58</td>
<td>5.77</td>
<td>DEXIA</td>
<td>-0.83</td>
<td>7.56</td>
</tr>
<tr>
<td>ALLIED IRISH BANKS</td>
<td>-2.16</td>
<td>13.41</td>
<td>EFG EUROBANK ERGASIAS</td>
<td>-0.21</td>
<td>5.19</td>
</tr>
<tr>
<td>ALPRA BANK</td>
<td>-0.42</td>
<td>4.36</td>
<td>ERSTE GROUP BANK</td>
<td>-0.61</td>
<td>9.86</td>
</tr>
<tr>
<td>ALLIANZ</td>
<td>-0.83</td>
<td>5.40</td>
<td>EURAZEO</td>
<td>-0.45</td>
<td>5.00</td>
</tr>
<tr>
<td>ATRIUM EUROPEAN RLST.</td>
<td>-0.32</td>
<td>10.60</td>
<td>FONCIERE DES REGIONS</td>
<td>-0.85</td>
<td>8.33</td>
</tr>
<tr>
<td>AXA</td>
<td>-1.05</td>
<td>6.67</td>
<td>GECINA</td>
<td>-0.33</td>
<td>7.49</td>
</tr>
<tr>
<td>AZIMUT HOLDING</td>
<td>-0.26</td>
<td>3.43</td>
<td>GBL NEW</td>
<td>-0.85</td>
<td>5.06</td>
</tr>
<tr>
<td>BANK OF IRELAND</td>
<td>-0.32</td>
<td>13.30</td>
<td>SOCIETE GENERALE</td>
<td>-0.72</td>
<td>4.72</td>
</tr>
<tr>
<td>BANKINTER 'R'</td>
<td>0.09</td>
<td>4.97</td>
<td>HANNOVER RUCK. (XET)</td>
<td>-0.85</td>
<td>6.65</td>
</tr>
<tr>
<td>BANCA CARIGE</td>
<td>-1.36</td>
<td>8.54</td>
<td>ICADÉ</td>
<td>-0.29</td>
<td>3.76</td>
</tr>
<tr>
<td>BANCA MONTE DEI PASCHI</td>
<td>-0.95</td>
<td>5.76</td>
<td>IMMOFINANZ</td>
<td>-2.75</td>
<td>19.24</td>
</tr>
<tr>
<td>BANCA POPOLARE DI MILANO</td>
<td>-0.61</td>
<td>4.37</td>
<td>ING GROEP</td>
<td>-1.36</td>
<td>9.58</td>
</tr>
<tr>
<td>BANCA PPO. DI SONDRO</td>
<td>-0.28</td>
<td>3.71</td>
<td>INTESA SANPAOLO</td>
<td>-0.96</td>
<td>5.40</td>
</tr>
<tr>
<td>BANCA PPO. EMILIA ROMAGNA</td>
<td>-1.02</td>
<td>7.33</td>
<td>KBC GROUP</td>
<td>-0.99</td>
<td>9.54</td>
</tr>
<tr>
<td>BANCA ARGENTARIA</td>
<td>-0.33</td>
<td>4.47</td>
<td>KLEIPERIE</td>
<td>-0.97</td>
<td>6.20</td>
</tr>
<tr>
<td>BANCO COMR. PORTUGUES 'R'</td>
<td>-0.50</td>
<td>4.10</td>
<td>MAPFRE</td>
<td>-0.40</td>
<td>4.94</td>
</tr>
<tr>
<td>BANCO DE VALENCIA</td>
<td>-0.25</td>
<td>4.27</td>
<td>MARFIN INV GP HDG.</td>
<td>0.19</td>
<td>3.78</td>
</tr>
<tr>
<td>BANCO ESPRITO SANTO</td>
<td>-1.03</td>
<td>6.19</td>
<td>MEDIOBANCA</td>
<td>0.11</td>
<td>4.36</td>
</tr>
<tr>
<td>BANCO POPOLARE</td>
<td>-0.99</td>
<td>8.06</td>
<td>MUENCHENER RUCK. (XET)</td>
<td>-0.77</td>
<td>10.20</td>
</tr>
<tr>
<td>BANCO POPULAR ESPANOL</td>
<td>-0.34</td>
<td>6.42</td>
<td>NATIONAL BK. OF GREECE</td>
<td>-0.37</td>
<td>4.70</td>
</tr>
<tr>
<td>BANCO DE SABADELL</td>
<td>-0.24</td>
<td>3.99</td>
<td>NATTIXIS</td>
<td>0.26</td>
<td>8.21</td>
</tr>
<tr>
<td>BANCO SANTANDER</td>
<td>-0.66</td>
<td>4.73</td>
<td>BANK OF PIRAEUS</td>
<td>-0.45</td>
<td>3.62</td>
</tr>
<tr>
<td>BNP PARIBAS</td>
<td>-0.66</td>
<td>6.56</td>
<td>POJIOLA PANKKI A</td>
<td>-1.77</td>
<td>13.97</td>
</tr>
<tr>
<td>BOLSAS Y MERCADOS ESPANOLÉS</td>
<td>-0.07</td>
<td>3.85</td>
<td>RAFFEISEN INTL. BK. HLDC.</td>
<td>-0.99</td>
<td>5.89</td>
</tr>
<tr>
<td>CATTOLICA ASSICURAZIONI</td>
<td>-0.35</td>
<td>5.45</td>
<td>SAMPO 'A'</td>
<td>-0.48</td>
<td>3.63</td>
</tr>
<tr>
<td>CNP ASSURANCES</td>
<td>-0.65</td>
<td>4.01</td>
<td>SCOR SE</td>
<td>-2.58</td>
<td>17.83</td>
</tr>
<tr>
<td>COPINIMMO</td>
<td>-1.43</td>
<td>7.86</td>
<td>SOFINA</td>
<td>-0.86</td>
<td>5.01</td>
</tr>
<tr>
<td>COMMERZBANK (XET)</td>
<td>-0.92</td>
<td>6.31</td>
<td>UBI BANCA</td>
<td>-0.86</td>
<td>7.05</td>
</tr>
<tr>
<td>CIE NALE A PTF.</td>
<td>-0.35</td>
<td>3.20</td>
<td>UNIBAIL-RODAMCO</td>
<td>-0.78</td>
<td>3.54</td>
</tr>
<tr>
<td>CINON</td>
<td>-0.84</td>
<td>4.47</td>
<td>UNICREDIT</td>
<td>-0.45</td>
<td>7.71</td>
</tr>
<tr>
<td>CREDIT Agricole</td>
<td>-0.66</td>
<td>4.17</td>
<td>VIENNA INSURANCE GROUP A</td>
<td>-0.66</td>
<td>13.37</td>
</tr>
<tr>
<td>CREDITO VAUTELLINES</td>
<td>0.16</td>
<td>5.09</td>
<td>WEDDEL</td>
<td>-0.93</td>
<td>4.58</td>
</tr>
<tr>
<td>CRITERIA CAIXACORP</td>
<td>-0.73</td>
<td>4.05</td>
<td>WERELDHAVE</td>
<td>-0.19</td>
<td>2.69</td>
</tr>
<tr>
<td>DELTA LLOYD GROUP</td>
<td>-0.32</td>
<td>1.70</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The sample covers the period between January 1992 and June 2010. But the length of time series differs for each financial institution. The longest time series contains 488 observations and the shortest one has 10 observations. We modified our model to adapt to this structure. We assume the time-varying equicorrelation matrix is driven by one common factor that follows the GAS process. The correlation between two institutions starts to load this dynamic factor once the equity returns become available for both names. So the size of the correlation matrix is also changing over time and reaches 73 at the maximum. There are two approaches to compute the stability measure for this large dimensional dataset. One is the simulation method proposed in Section 4.5.1. The drawback is that it takes a long time to generate enough simulations for all possible stressed scenarios. The alternative way is to use the law of large numbers (LLN) rule to approximate the probability, as in Section 4.4. This approach is numerically easier and
still sensible if the main purpose of the study is a joint risk analysis as demonstrated in the previous subsection.

We assume our 73 institutions form a homogenous portfolio. That means all the institutions have the same skewness and kurtosis coefficients $\gamma$ and $\nu$ in the multivariate dynamic distribution for their equity returns. With the volatilities estimated from marginal GAS-GH skewed-$t$ model, we standardize the equity returns and focus on the modeling of dynamic correlations. A multivariate GHST distribution is estimated with the equicorrelation restriction. The parameter estimation results are shown in Table 4.4. The correlation coefficient plotted in Figure 4.5 hovers around 0.3 over time. Compared with a rolling window correlation series (the window size is 12), the GAS equicorrelation is more persistent over time. But the means of these two correlation series are similar.

We compute the financial risk measures analytically given the multivariate GHST model and the probability of default from the expected default frequency (EDF) of Moody’s KMV. We numerically evaluate the integral (4.44) to compute the Banking Stability Measure, defined as the probability of more than 10% financial institutions defaulting. The risk measure is plotted in Figure 4.5. From the figure, the LLN result for the default probability does not move too much before 2008. But it appears that the period of 2008-2010 is quite special: the failure probability increases to more than five times the historical mean. We also compute the same measure with the simulation method. The approximated risk indicator is the same as the simulated one. So we did not include that in the graph. We plot the LLN approximated risk indicators, the Banking Stability Measure and the Systemic Risk Measure in Figure 4.6. From the graphs, we see the large influence of the recent financial crisis, which drives up the two risk measures in that period. Note that the systemic risk indicator shoots up to 0.60 around the failure of Lehman Brothers.

### 4.6 What factors drive the bank equity correlation?

A natural extension of the statistical model so far is to relate the correlation dynamics in the banking system to observed factors. There are extensive discussions about the common factors underlying stock returns and stock return correlations, for example Hou et al. (2011). Also research has been done on idiosyncratic volatility factors, see Bekaert et al. (2010). However to the knowledge of the authors, less attention has been paid to
4.6. WHAT FACTORS DRIVE THE BANK EQUITY CORRELATION?

Figure 4.5: The Banking Stability Measures in whole sample

The Banking Stability Measure defined as the probability of more than 10% firms defaulting under the Law of Large Numbers approximation result. The upper-right panel shows the dynamic correlation estimated in the GAS-Equicorrelation model. And the bottom-left panel plots the average of pairwise rolling window correlation coefficients. As a comparison, the bottom-right panel shows the two correlation estimates jointly.
The Banking Stability Measure (BSM) and Systemic Risk Measure (SRM) under the LLN approximation from the GAS-GHST Equicorrelation model. The BSM indicator is defined as the probability of more than 10% firms defaulting at time $t$. The SRM indicator is the average of the default probability of more than seven other firms defaulting conditional on firm $i$ defaulting.
study which observed factors determine dynamic correlations in the banking system. This is essential for the purpose to differentiate contagion effects and interdependence in the comovement of banks’ equity returns. In this section, we select a set of economic variables ranging from global macroeconomic factors to market and country specific observed indicators to capture the dynamics in correlations.

We allow the correlations \( \rho_t \) to depend on both observed \( (X_t) \) and unobserved \( (f_t) \) factors. Define \( X = (X_1, \cdots, X_t) \) as a \( k \times \tau \) matrix of economic variables and \( \beta \) a \( 1 \times k \) vector of parameters,

\[
\rho_t = \vartheta(f_t + \beta X_t). \tag{4.47}
\]

where \( f_t \) has the familiar GAS dynamics. The monthly economic variables we choose for \( X_t \) are: the European Volatility Index (VSTOXX), the Euribor-EONIA (ECB Overnight Interest) spread, and the S&P European stock market index return. VSTOXX is a popular measure of the implied volatility of European index options. It is considered to be a good measure of the short term volatility and thus an indicator for market turbulence. The Euribor-EONIA spread is a measurement of liquidity in the banking sector. The S&P European stock market index tracks the health of European equity markets and corporate profitability conditions as perceived by financial markets. It describes the condition of European stock markets and also reflects the state of the economy. The time series are plotted in Figure 4.7. The sample including the explanatory variables \( X_t \in \mathbb{R}^3 \) covers the period January 2000 to June 2010. The parameters in the Dynamic GHST framework and the new regression coefficients \( \beta \) are estimated via maximizing the log likelihood.

The coefficient estimates and standard errors are presented in Table 4.4. All coefficients are significant at the level 5%. It appears that the stock market index significantly explains the correlation movement. The negative sign coincides with the past observations about downside risk and rising correlations during crises. As the Euribor-EONIA spread is a measure of a lack in funding liquidity for the banking sector, we see the positive coefficient for this variable as an indication that the correlation increases in times of market turbulence and reduced liquidity. The coefficient for the VSTOXX is positive. It means that the correlation is high when the market volatility is also high, as in times of financial crisis. We plot the estimated equicorrelation with extra economic variables in Figure 4.8.

By adding observed economic variable to the GAS dynamics, the correlation estimates
Figure 4.7: The economics variables.

The plots of the economic variables used to explain the variation of correlation over time: Euribor-EONIA, EU stock index return and VSTOXX.
4.6. **WHAT FACTORS DRIVE THE BANK EQUITY CORRELATION?**

The parameter estimates in the GAS-GHST models with extra economic variables and the GAS-GHST Equicorrelation model. These models are estimated with the filtered returns data. All parameters are significant at the 5% level. Note that the two models are applied to different datasets. Due to the availability of the economic variables, the Augmented GAS model uses a shorter dataset starting in January 2000 and ending in June 2010.

<table>
<thead>
<tr>
<th></th>
<th>Augmented GAS Model</th>
<th>GAS EqCorr Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td>0.270</td>
<td>0.239</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.018)</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>0.903</td>
<td>0.897</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.012)</td>
</tr>
<tr>
<td><strong>ω</strong></td>
<td>-1.279</td>
<td>-1.214</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.004)</td>
</tr>
<tr>
<td><strong>γ</strong></td>
<td>-0.039</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Euribor-EONIA</td>
<td>0.126</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>S&amp;P European stock index</td>
<td>-0.869</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td></td>
</tr>
<tr>
<td>VSTOXX</td>
<td>0.160</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td><strong>Log-lik</strong></td>
<td>-9498.91</td>
<td>-9502.36</td>
</tr>
</tbody>
</table>
The correlation coefficient estimated over time with the Dynamic GHST density but with different correlation models. The top-left panel plots the correlation estimates from the Equicorrelation model. The top-right panel contains the time-series plot of the correlation in the Equicorrelation model with extra economics variables. The bottom two panels shows the correlation and the GAS factors under the Equicorrelation model and the Equicorrelation model with extra variables. The left panel plots the correlation estimates and the right panel is the plot of factors.
are less persistent over time. They are still close to the previous GAS correlation estimates. The major difference emerges before and after the Lehman Brothers’ bankruptcy on 2008. The correlation produced by the augmented GAS model is higher before the bankruptcy. Also note that the GAS factors $f_t$ in the augmented model are lower than the factors in the GAS model. The reason is that the observed economic variables capture part of the correlation movements. The time-varying pattern in correlations is driven partly by the market perception of risk and liquidity. However, a substantial part of the correlation dynamics are still unexplained with the three variables considered in our current analysis. Further macroeconomic variables need to be added to the model to enhance its explanatory power and capture more of the underlying economic mechanisms for changing dynamic (systemic) correlations.

4.7 Conclusion

In this paper, we develop the dynamic GHST model with GAS-Equicorrelation or block GAS-Equicorrelation structure. These models are applicable to large dimensional problems. We also propose two risk measures with a large panel of multiple European financial institutions. The Banking Stability measure we developed indicates the joint default risk in the system. The Systemic Risk Measure takes the average of conditional default probabilities to test the interconnectedness of the financial system. The full dynamic multivariate model with the GH skew-$t$ distribution is used to simulate the possible distress scenarios for the banks. Based on the Monte Carlo simulation, we can analyze the joint and conditional credit risk in individual financial institutions. Another risk measuring model originates from the conditional Law of Large Numbers approximation method. With the application of a Dynamic Equicorrelation model in a large system of financial firms, the approximated risk indicator provides a good measure of credit risks for an unbalanced large panel.

We further showed the explanatory power of some commonly used economic variables (VSTOXX index, Euribor-EONIA spread and European stock market index) to explain systemic correlation dynamics. By introducing these new variables in our dynamic system, the correlation becomes less persistent compared to the pure GAS dynamic model. The residual GAS factor decreases due to the explanatory power of the extra economic variables. It appears that we still miss one or a few more factors to explain the variation
in correlation dynamics. Moreover, we might miss a few important firm specific variables, such as the leverage ratio. The current model also enable us to measure the systematic contribution of each bank by looking at the conditional probability in the multivariate GH skewed-$t$ distribution. We leave this for future research.
Appendix to Chapter 4

The GH skewed-$t$ distribution is a subclass of the GH distribution family which preserves much of the flexibility of GH distribution, but with less parameters. With the observed stock return for bank $i$ defined as $y_{it}$ following the GH skewed-$t$ distribution, the model is

$$y_{it} \sim p(\tilde{\Sigma}_t, \nu, \gamma), \quad (4.48)$$

$$\tilde{\Sigma}_t = L_t(TT')L_t', \quad (4.49)$$

$$\Sigma_t = L_tL_t' = R_t, \quad (4.50)$$

where $\gamma$ collects the skewness parameters and the matrix $T$ satisfies the condition

$$(T'T)^{-1} = \frac{\nu}{\nu-2}I_k + \frac{2\nu^2}{(\nu-2)^2(\nu-4)}\gamma'\gamma. \quad (4.51)$$

The deco-Dynamic-GH model defines the correlation matrix as

$$R_t = (1 - \rho_t)I_k\rho_t\ell_k\ell_k', \quad \rho_t \in \left(-\frac{1}{k-1}, 1\right), \quad (4.52)$$

where $\Sigma_t = R_t$ is the dynamic conditional correlation matrix we are interested in and $\ell_k \in \mathbb{R}^k$ is a vector of ones. In this model, we define $L_t$ as a symmetric matrix. Further, we parameterize $\rho_t$ as a GAS model

$$\rho_t = \partial(f_t), \quad (4.53)$$

$$f_{t+1} = \omega + As_t + Bf_t. \quad (4.54)$$

We also define the scale matrix as $\tilde{\Sigma}_t = \hat{L}_t\hat{L}_t'$. The variable $T$ links these two matrices such that $\hat{L}_t = L_tT$.

The innovation term $s_t$ is the scaled observation density score as in Chapter 2. Note that the matrix $L_t$ is symmetric.

$$s_t = \mathcal{S}_t\nabla_t, \quad (4.55)$$

$$\nabla_t = \Psi_t'H_t'\text{vec}\left(w_t \cdot y_t'y_t' - \tilde{\Sigma}_t - \left(1 - \frac{\nu}{\nu-2}w_t\right)\tilde{L}_t'\gamma y_t'\right), \quad (4.56)$$

$$\mathcal{S}_t = \left\{\Psi_t'(I \otimes \tilde{L}_t^{-1})'[gG - \text{vec}(I)\text{vec}(1)'](I \otimes \tilde{L}_t^{-1})\Psi_t\right\}^{-1}, \quad (4.57)$$

$$H_t = ((I_{k^2} + C_k)(\Sigma_t \otimes \tilde{\Sigma}_t))^{-1}, \quad (4.58)$$

$$\Psi_t = \frac{\partial\text{vec}(\Sigma_t)'}{\partial f_t}, \quad (4.59)$$

$$\hat{L}_t = L_tT.$$
where \( g = \frac{\nu + d}{\nu + 2 + d} \) and \( G \) is defined as in Creal et al. (2011).

From the derivation, it is clear that we have to take inverses and compute the determinants of matrices in a large dimension. If we have the matrices in blocks or in the form of the equicorrelation model, we can obtain the inverse and determinant in analytical form which will help to speed up the computations. To get the matrix \( L_t \) in an easy-to-operate form, we have

\[
\tilde{L}_t = a_t I_k + b_t \ell_k \ell'_k,
\]

(4.60)

where \( a_t = \sqrt{1 - \rho_t} \) and \( b_t = (\sqrt{1 - \rho_t + k\rho_t} - \sqrt{1 - \rho_t})/k \). The condition for the correlation matrix to be positive definite does not change.

\[
T^{-1} = c I_k + d \ell_k \ell'_k,
\]

(4.61)

where

\[
c = 1 \frac{1}{\sqrt{\mu_\xi}},
\]

\[
d = \sqrt{\mu_\xi} - \sqrt{\sigma^2_\xi \gamma' \gamma + \mu_\xi},
\]

\[
\mu_\xi = \frac{\nu}{\nu - 2},
\]

\[
\sigma^2_\xi = \frac{2\nu^2}{(\nu - 2)^2(\nu - 4)}.
\]

So we have

\[
L_t = \tilde{L}_t T^{-1} = a_t c I_k + (a_t d + b_t c + k b_t d) \ell_k \ell'_k.
\]

(4.62)

It is straightforward to derive the inverse and determinant as

\[
L_t^{-1} = \frac{1}{a_t c} I_k - \frac{a_t d + b_t c + k b_t d}{a_t c(a_t c + k(a_t d + b_t c + k b_t d))} \ell_k \ell'_k,
\]

(4.63)

\[
det(\Sigma_t) = \det(L_t)^2 = (a_t c)^{2(k-1)} \left( a_t c + k(a_t d + b_t c + k b_t d) \right)^2.
\]

(4.64)

In the application with the whole sample, it appears that the computation of the scale matrix and score matrix takes too much time. One reason is the inversion of a large \( k \times k \) matrix in equation (4.59). In order to reduce the burden for calculation, we manage to
derive $\Psi_t$ analytically, which would speed up the computational speed.

\[
\Psi = \frac{\partial \text{vec}(L_t)}{\partial f_t} = \left( \frac{\partial a_t}{\partial \rho_t} \text{vec}(I_k) + \left( \frac{\partial a_t}{\partial \rho_t} d + \frac{\partial b_t}{\partial \rho_t} c + \frac{\partial b_t}{\partial \rho_t} k d \right) \ell_{k2} \right) \varphi'(f_t), \tag{4.65}
\]

\[
\frac{da_t}{d\rho_t} = -\frac{1}{2\sqrt{1 - \rho_t}}; \tag{4.66}
\]

\[
\frac{db_t}{d\rho_t} = \frac{1}{2k} \left( -\frac{k - 1}{\sqrt{1 - \rho_t + k \rho_t}} + \frac{1}{\sqrt{1 - \rho_t}} \right). \tag{4.67}
\]

This does help in getting out the correlation simply. The scale matrix $S_t$ is the inverse Fisher information matrix from the symmetric $t$ distribution, as explained in Chapter 2.
Chapter 5

Conclusion

This chapter concludes the thesis. Section 5.1 present a summary of the main findings. Section 5.2 contains possible directions for future research.

5.1 Summary of the main conclusions

This thesis focuses on modeling the time variation in financial systemic risk, which is of great interest in times of financial crises. This section summarizes the main findings and conclusions.

In Chapter 2, we presented and motivated a novel multivariate time series model for time-varying volatilities and correlations under a non-Gaussian distribution family named Generalized Hyperbolic distributions. The model is the first to allow the asymmetric and fat-tailed shape of the observation distribution to directly affect the transition mechanism of the volatilities and correlations. As the model is a dynamic recursion driven by a local density score, the parametric assumption of the distribution influences not only the likelihood, but also the dynamics of the time-varying parameter. The volatilities and correlations updating mechanism reweights the observed observations according to the parametric assumptions imposed. This mitigates the impact of outliers and incidental influential observations on the time variation in the covariance matrix. With Monte Carlo simulations, we show that the model does a better job at estimating the time-varying correlation than competing models if the error distribution is fat-tailed and skewed. Using the U.S. stock returns, we find the estimated volatilities and correlations based on the new model following a robust pattern over time than their counterparty based on the normal distribution and GARCH type dynamics. We also developed the Expectation-
Maximization procedures for the new observation driven model. This chapter opens the possibility for modeling financial returns and risks with a general time-varying distribution that fits empirical data well.

In Chapter 3, we constructed a new measure of joint and conditional failure probabilities for Euro Area sovereigns. The risk assessment framework uses sovereign Credit Default Swap (CDS) spread data to calibrate sovereign default probabilities at each point in time. Applying the time series model in Chapter 2, the dependence between countries is estimated with a skewed heavy-tailed multivariate distribution with non-trivial tail-dependence. Our sample of the Euro Area sovereigns CDS covers the period between 2008 and 2011, during which the default risk for one or more sovereigns has been rising. Over recent years, the CDS market experienced a turbulent time with many jumps and asymmetric movements. As the model uses a dynamic distribution that accounts for the non-Gaussian properties in the data, the framework is very suitable to extract robust risk measures from the CDS market. We document substantial increases of conditional default probabilities for several EU sovereigns conditional on a Greek failure. Comparing the model results with other parametric models, we find that it is important to capture higher-order time-varying moments for conditional risk statements. Joint risk statements are less sensitive to the distribution we use. Our model also allows us to investigate the impact of central bank policy interventions. With a study on the effect of the rescue package announced on May 9, 2010, we find that the market perceived the new policy as an attempt that lowered joint risks, but not necessarily decreased the joint dependence.

In Chapter 4, we developed a multivariate risk model to measure the credit risk in a large banking system. Using a normal mean-variance mixture, we could exploit the non-Gaussian properties of financial datasets to drive the time-varying pattern in the dependence structure. By imposing restrictions on the mixing construction, we could reduce the computational burden of estimating a large dimensional covariance matrix. We also proposed two risk measures for a large panel of European financial institutions. The Banking Stability measures we presented capture the joint default risk in the financial system. The Systemic Risk measure takes the average of conditional default probabilities to test the interconnectedness of the financial system. Based either on simulations or on a conditional Law of Large Numbers approximation, we analyzed the joint and conditional risk in the system of financial institutions. We further extended the study to relate the
correlation dynamics in the banking system to observed risk factors. The results showed that some commonly used economic variables (VSTOXX index, Euribor-EONIA spread and European stock market index) can partly explain systemic correlation dynamics.

5.2 Direction for future research

The main chapters of the thesis established a general framework to model financial risks in an interconnected system. The last chapter already provided two topics for future research. The study on explaining the systemic correlation with extra economic variables shows that we still miss one or more factors to explain the variation in correlation dynamics. It would be interesting to link up the real economic variables to the time variation in correlation and systemic risk. From this extension, we may be able to forecast systemic risk from the fundamental economic development. Secondly the model from Chapter 4 also enables us to measure the systematic risk contribution of each bank by looking at the conditional probability from the multivariate non-Gaussian distribution.

This thesis studies the financial systemic risk and sovereign credit risk in isolation. Acharya et al. (2011) argues that financial systemic risk and sovereign credit risk are intimately linked. Governments have to provide the bank bailout plans in the distressed system, which is funded by diluting existing government bondholders. As a result, there is a deterioration of the sovereign’s creditworthiness. Because of the feedback between financial and sovereign credit risk and it is not enough to study the two risk separately. In future work, we plan to extend our econometric models to identify and measure these risk transfer channels.
Bibliography


BIBLIOGRAPHY


Samenvatting (Summary in Dutch)

Systeemrisico vormt een grote bedreiging voor de stabiliteit van moderne financiële systeem. In het onderling verbonden stelsel van financiële instituties en soevereine staten houdt onderzoek naar systeemrisico in dat uitdrukkelijk gekeken wordt naar zowel de structuur van de onderlinge afhankelijkheden als naar de tijdsvariatie daarin. Beide kunnen leiden tot systeemrelevant schokken. Dit proefschrift verschaf een coherent econometrisch raamwerk waarmee financieel systeemrisico op een algemene en consistente manier gemeten kan worden. Het raamwerk is toegepast op financiële markten, gebruik makend van rendementen op aandelen en veranderingen in credit default swap premies. Beiden zijn typisch asymmetrisch verdeeld met dikstaartige verdelingen. De bijbehorende varianties en correlaties veranderen ook over tijd, zeker tijdens een crisis. Het in dit proefschrift ontwikkelde model is algemeen genoeg om deze niet-Gaussische elementen en tijdsvariërende conditionele covarianties te beschrijven. We rapporteren sterk bewijs voor tijdsvariatie in systeemrisico’s, en onderschrijven het belang van hogere-orde momenten in het modelleren van systeemrisico.

Hoofdstuk 2 introduceert een nieuw model voor dynamische volatiliteit en correlaties van scheve en dikstaartige data. Ons model voegt tijdsvariërende parameters, gedreven door de score van de geobserveerde kansdichtheidsfunctie, toe aan niet-normale verdelingen. De belangrijkste vernieuwing in onze aanpak is het feit dat de scheve en dikstaartige vorm van de verdeling het dynamische gedrag van de tijdsvariërende parameters direct beïnvloedt. Dit onderscheidt onze procedure van bekende alternatieve modellen waarin aannames over de verdeling wel de aannemelijkheidsfunctie maar niet de parameterdynamiek beïnvloeden. We presenteren een aangepast expectation-maximization algoritme om het model te schatten. Simulaties en empirisch bewijs tonen aan dat dit model beter werkt dan alternatieve gangbare modellen als scheefheid en dikstaartigheid relevante karakteristieken van de data zijn.
In Hoofdstuk 3 presenteren we een nieuw empirisch raamwerk om de aannemelijkheid van gezamenlijke en conditionele faillisementen te onderzoeken voor soevereine Eurolanden. Ons model is gebaseerd op een dynamisch, niet-normalecopula, die alle belangrijkkeelementen van de data beschrijft, inclusief niet-normaal verdeelde veranderingen in de prijs van credit default swap (CDS), alsook dynamische volatiteiten en correlaties die ervoor zorgen dat de afhankelijkheid van falen kan toenemen in risicovolle tijden. We passen het raamwerk toe op de veranderingen in CDS spreiding van Eurolanden tussen 2008 en mid-2011. Onze resultaten tonen significante tijdsvariatie in risicoafhankelijkheid aan en een grote rol voor overloop-effecten in de aannemelijkheid van faillisementen van individuele staten. We onderzoeken tevens deze afhankelijkheid rondom een belangrijke beleidsaankondiging door leiders van Eurolanden op 9 mei 2010, en tonen het belang aan van het modelleren van hogere-orde tijdsvariërende momenten in crisistijd.

In Hoofdstuk 4 berekenen we twee nieuwe indicatoren voor financieel systeemrisico, gebaseerd op tijdsvariërende conditionele en onconditionele kansen van gezamenlijk falen van meerdere financiële instellingen. Deze risicomaatstaven zijn afgeleid vanuit een multivariaat model met asymmetrische niet-normaal verdeelde veranderingen in de marktwaardes van de koersen van financiële instellingen. Ons model kan gezien worden als een Merton-model op basis van gecorreleerde Lévy processen. Het model bevat dynamische volatiliteiten en afhankelijkheidsindicatoren en gebruikt de volledige informatie over de vorm van de multivariate verdeling. Onze correlatieschattingen zijn robuust tegen uitschieters en invloedrijke waarnemingen. Voor zeer grote cross-sectionele dimensies bieden we een benadering, gebaseerd op een conditionele wet van grote aantallen om extreme gezamenlijke kansen op faillissement te berekenen. We passen het model toe op het meten van het risico dat meerdere financiële instellingen in de Europese Unie tijdens de crisis failliet gaan. Door enkele economische variabelen die indicatief zijn voor systeem-stress toe te voegen aan het model tonen we aan dat het toevoegen van extra economische variabelen kan helpen de systeemrelevante correlatiedynamiek te verklaren.

Samenvattend biedt dit proefschrift een multivariaat econometrisch raamwerk aan om systeemrisico en bankroetrizico te meten. Het is een gedisaggregeerd risicomodel met een flexibele tijdsvariërende afhankelijkheidsstructuur. Omdat de financiële rendementen niet normaal verdeeld zijn, gebruiken we een algemene niet-Gaussische verdeling die alle belangrijke karakteristieken van de data beschrijft. De verschillende toepassingen laten
zien hoe het modelraamwerk in de praktijk kan worden geoperationaliseerd.
The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

500  J. GARCÍA, The moral herd: Groups and the Evolution of Altruism and Cooperation
501  F.H. LAMP, Essays in Corporate Finance and Accounting
502  J. SOL, Incentives and Social Relations in the Workplace
503  A.I.W. HINDRAYANTO, Periodic Seasonal Time Series Models with applications to U.S. macroeconomic data
504  J.J. DE HOOP, Keeping Kids in School: Cash Transfers and Selective Education in Malawi
505  O. SOKOLINSKIY, Essays on Financial Risk: Forecasts and Investor Perceptions
506  T. KISELEVA, Structural Analysis of Complex Ecological Economic Optimal Management Problems
507  U. KILINC, Essays on Firm Dynamics, Competition and Productivity
508  M.J.L. DE HEIDE, R&D, Innovation and the Policy Mix
509  F. DE VOR, The Impact and Performance of Industrial Sites: Evidence from the Netherlands
510  J.A. NON, Do ut Des: Incentives, Reciprocity, and Organizational Performance
511  S.J.J. KONIJN, Empirical Studies on Credit Risk
512  H. VRIJBURG, Enhanced Cooperation in Corporate Taxation
513  P.ZEPPINI, Behavioural Models of Technological Change
514  P.H.STEFFENS, It’s Communication, Stupid! Essays on Communication, Reputation and (Committee) Decision-Making
515  K.C. YU, Essays on Executive Compensation - Managerial Incentives and Disincentives
516  P. EXTERKATE, Of Needles and Haystacks: Novel Techniques for Data-Rich Economic Forecasting
517  M. TYSZLER, Political Economics in the Laboratory
518  Z. WOLF, Aggregate Productivity Growth under the Microscope
520  P.R. KOSTER, The cost of travel time variability for air and car travelers
Y. ZU, Essays of nonparametric econometrics of stochastic volatility

B. KAYNAR, Rare Event Simulation Techniques for Stochastic Design Problems in Markovian Setting

P. JANUS, Developments in Measuring and Modeling Financial Volatility

F. P. W. SCHILDER, Essays on the Economics of Housing Subsidies

S. M. MOGHAYER, Bifurcations of Indifference Points in Discrete Time Optimal Control Problems

C. ÇAKMAKLı, Exploiting Common Features in Macroeconomic and Financial Data

J. LINDE, Experimenting with new combinations of old ideas

D. MASSARO, Bounded rationality and heterogeneous expectations in macroeconomics

J. GILLET, Groups in Economics

R. LEGERSTEE, Evaluating Econometric Models and Expert Intuition

M. R. C. BERSEM, Essays on the Political Economy of Finance

T. WILLEMS, Essays on Optimal Experimentation

Z. GAO, Essays on Empirical Likelihood in Economics

J. SWART, Natural Resources and the Environment: Implications for Economic Development and International Relations

A. KOTHIYAL, Subjective Probability and Ambiguity

B. VOOGT, Essays on Consumer Search and Dynamic Committees

T. DE HAAN, Strategic Communication: Theory and Experiment

T. BUSER, Essays in Behavioural Economics

J. A. ROSERO MONCAYO, On the importance of families and public policies for child development outcomes

E. ERDOGAN CIFTCI, Health Perceptions and Labor Force Participation of Older Workers

T. WANG, Essays on Empirical Market Microstructure

T. BAO, Experiments on Heterogeneous Expectations and Switching Behavior

S. D. LANSDORP, On Risks and Opportunities in Financial Markets

N. MOES, Cooperative decision making in river water allocation problems

P. STAKENAS, Fractional integration and cointegration in financial time series

M. SCHARTH, Essays on Monte Carlo Methods for State Space Models

J. ZENHORST, Macroeconomic Perspectives on the Equity Premium Puzzle

B. PELLOUX, the Role of Emotions and Social Ties in Public On Good Games: Behavioral and Neuroeconomic Studies

N. YANG, Markov-Perfect Industry Dynamics: Theory, Computation, and Applications

R. R. VAN VELDHUIZEN, Essays in Experimental Economics