HEC Montréal
Affiliée à l'Université de Montréal

Three Essays on Modeling Asset Co-movement Asymmetries Using Copula Functions

par

Denitsa Stefanova

Département de Finance et Chaire de recherche du Canada
en gestion des risques

HEC Montréal

Thèse présentée en vue de l'obtention du grade de Philosophiæ Doctor
(Ph.D.) en Administration

Septembre, 2008

HEC Montréal
Affiliée à l'Université de Montréal

Cette thèse intitulée:

Three Essays on Modeling Asset Co-movement Asymmetries Using Copula Functions

Présentée par:

Denitsa Stefanova

a été évaluée par un jury composé des personnes suivantes:

Président-rapporteur:	Dr. Chantal Labbé HEC Montréal
Président du comité de surveillance:	Dr Georges Dionne HEC Montréal
Co-directeur:	Dr Tony Berrada Université de Genève
Membre du jury:	Dr. Bruno Rémillard HEC Montréal
Examinateur externe:	Dr. Marcel Rindisbacher Boston University School of Management

Résumé

Parmi les caractéristiques clés de tout modèle d'allocation de portefeuille se trouvent les hypothèses concernant la structure de dépendance des facteurs de risque sous-jacents. Une modélisation inappropriée peut entamer une interprétation inadéquate de l'exposition au risque et ainsi mener à des choix de portefeuille sous-optimaux. La corrélation linéaire a été l'outil traditionnel dans l'analyse de la dépendance dans un cadre statique ou dynamique, mais des études récentes ont démontré son incapacité de refléter la dépendance entre des événements extrêmes, dont l'asymétrie prononcée est devenue un fait stylisé de la distribution des rendements : les titres ont la tendance d'évoluer dans la même direction lorsque le marché est en baisse que lorsque le marche est en hausse (Poon et al., 2004). Au contraire, la théorie des copules fournit un environnement approprié pour la recherche de mesures de dépendance qui conviennent mieux à l'asymétrie des co-mouvements extrêmes. Le caractère parcimonieux des fonctions de copules les rend appropriées pour la modélisation de risques multiples et ainsi pour les problèmes de choix de portefeuille.

Malgré la prolifération d'articles sur les copules, qui sont concentrés majoritairement sur la spécification, l'estimation et les tests de validité de l'ajustement, il n'y a pas beaucoup de recherche faite sur la modélisation de la dépendance spatiale des processus stochastiques. Les applications sur les copules existantes sont majoritairement concentrées sur la modélisation conditionnelle de processus stochastiques en temps discret dans le contexte des modèles GARCH, et les modèles de choix de portefeuille correspondants traitent plutôt l'allocation inconditionnelle sur la prochaine période (Patton, 2004; Jondeau and Rockinger, 2002, 2005). Par conséquent, le but de la thèse présente a deux dimensions: proposer un processus stochastique en temps continu capable de refléter les asymétries dans les co-mouvements des facteurs de risque sous-jacents, et examiner les implications d'un tel processus multidimensionnel sur la couverture inter-temporelle d'un portefeuille.

Dans le premier chapitre de la thèse je propose une extension multi-variée de la construction d'un processus stochastique avec une distribution stationnaire donnée, à la base des fonctions des copules. Dans le contexte uni-varié les processus de prix construits à partir d'une distribution stationnaire pré-spécifiée ont été largement étudiés, et leur succès dans la réplication des faits stylisés des rendements en termes de leurs propriétés dynamiques et leur structure de dépendance statique a été établi. L'extension multi-varié, que je propose

utilise la relation entre la mesure stationnaire du processus et sa spécification de diffusion (Hansen and Scheinkman, 1995). La dépendance asymétrique dans les queues est prise en considération à l'aide d'une mixture de fonctions de copules qui proviennent des familles Elliptiques ou de Valeurs Extrêmes. Elle est isolée du comportement marginal, modélisé avec la classe flexible de distributions généralisées hyperboliques.

Le processus de diffusion proposé est un processus hautement non-linéaire ce qui pose des sérieux problèmes d'estimation. Malgré le fait que la distribution stationnaire du processus soit connue explicitement, ceci n'est pas vrai pour la densité de transition, ce qui rend inapplicables les techniques d'estimation standards comme le maximum de vraisemblance, sauf si on fait recours à la discrétisation. Je propose d'utiliser les techniques d'estimation MCMC basées sur l'augmentation de l'espace d'états afin de diminuer le biais de discrétisation dans le sens de Durham and Gallant (2002), Roberts and Strammer (2001) et Golightly and Wilkinson (2006a).

Dans le deuxième chapitre de la thèse j'étudie l'effet de la dépendance asymptotique sur le choix de portefeuille dans un contexte de marché complet, où une solution explicite des termes de couverture inter-temporelle du portefeuille est obtenue en utilisant la théorie des martingales et l'application du calcul de Malliavin (Detemple et al., 2003). Je compare l'évolution des termes de couverture inter temporels au fil du temps, induits par une structure de dépendance asymétrique dans les extrêmes à celles d'un modèle asymptotiquement indépendant. Un exercice utilisant les vraies données et un autre basé sur les simulations suggèrent un déplacement de la composition du portefeuille vers l'actif sans risque lorsque la dépendance entre les événements de queue pour les actifs risqués est prise en compte. J'évalue également l'importance économique de ce changement des parts de portefeuille, principalement motivée par la nécessité de se couvrir contre les variations stochastiques des variables d'état lorsque le processus générant les données incorpore un comportement asymétrique dans les queues. A cet effet je calcule les coûts de l'équivalent certain engendrés par le fait d'ignorer la dépendance entre les valeurs extrêmes et je trouve que la prise en considération de ce fait stylisé mène à des gains économiques importants.

Dans le troisième chapitre de la thèse j'aborde le problème d'allocation optimale de portefeuille dans un cadre dynamique lorsque la corrélation conditionnelle des rendements des actifs risqués est modélisée avec des facteurs observables, ce qui me permet d'isoler la

demande de couverture contre le risque de corrélation. Ainsi, je suis en mesure d'analyser séparément l'impact de la dépendance de queues à travers la distribution inconditionnelle et ce de la corrélation conditionnelle sur les parts optimales de portefeuille. Avec ces deux approches différentes de modéliser la dépendance je réplique le fait stylisé d'une dépendance accrue pendant des périodes de chutes extrêmes du marché, de volatilité croissante et d'aggravation des conditions macro-économiques. Je trouve que les termes de couverture contre le risque de corrélation ainsi que ceux engendrés par la dépendance dans les queues ont un impact distinct sur le portefeuille optimal et ne peuvent pas agir comme des substituts les uns des autres. De plus, le fait d'ignorer la dynamique de la corrélation conditionnelle ou la dépendance extrême engendre des coûts économiques non-négligeables.

Mots clés : Markov Chain Monte Carlo, allocation dynamique de portefeuille, simulation Monte Carlo, dépendance extrême, fonctions de copules, diffusion stationnaire multivariée, corrélation dynamique conditionnelle.

Abstract

A key feature of any portfolio allocation model is the assumption concerning the dependence structure of the underlying risk factors. Its inappropriate modeling could lead to a misunderstanding of the risk exposure and thus to suboptimal portfolio choices. Linear correlation has been the traditional tool for describing dependence in both static and dynamic settings, but recent studies have demonstrated its inability to capture dependence between extreme events, whose pronounced asymmetry has turned into an established stylized fact for asset returns: assets tend to move together in extreme market downturns to a greater extent than in extreme market upturns (Poon et al., 2004). Instead, copula theory provides a natural environment for the search of dependence measures that are better suited for capturing extreme co-movement asymmetries. The highly parsimonious nature of copula functions makes them suitable for high-dimensional models, as those encountered in portfolio choice problems.

Despite of the abundant literature on copulas, focusing mainly on their specification, estimation and goodness-of-fit tests, not much research has been done for the multivariate dependence modeling of stochastic processes. Applications are mainly centered on conditional modeling of discrete time processes within a GARCH framework, while portfolio choice applications based on copulas mostly treat the unconditional one-period-ahead allocation (Patton, 2004; Jondeau and Rockinger, 2002, 2005). The aim of the present thesis is thus twofold: propose a continuous-time stochastic process that is able to accommodate co-movement asymmetries in the underlying risk factors, and investigate its implications for the hedging behaviour in a dynamic portfolio allocation setting.

In the first chapter of this thesis I propose a multivariate extension to the construction of a stochastic process with a given stationary distribution, based on copula functions. In the univariate setting, price processes with a prespecified marginal distribution have been largely studied and proven successful in replicating stylized features of asset returns in terms of their dynamic properties and static dependence structure. The multivariate extension I propose exploits the relationship between the stationary measure and the diffusion specification of the process (see Hansen and Scheinkman, 1995). The asymmetric tail dependence is captured by a mixture copula of the Elliptic and the Extreme Value families, which is isolated from the marginal behavior, modeled by the flexible Generalized Hyperbolic class of distributions.

The proposed diffusion process that is highly non-linear poses a serious estimation problem. Even though the stationary measure of the process is explicitly known, the same is not true for the transition density, thus rendering standard maximum likelihood estimation impossible without resorting to discretisation. I propose a sequential MCMC estimation of the process that relies on increasing the state space in order to subdue any discretisation bias in the lines of Durham and Gallant (2002), Roberts and Strammer (2001) and Golightly and Wilkinson (2006a).

In the second chapter of this thesis I study the effect of asymptotic extreme value dependence on portfolio choice in a complete market setup where optimal allocation rules are obtained analytically under the Martingale technique using Malliavin calculus in the lines of (Detemple et al., 2003). I compare the evolution of the intertemporal hedging terms over time induced by a data generating process that allows for asymmetric dependence in the extremes to those of an asymptotically independent model. A real-data experiment and a simulation exercise both suggest a shift in the portfolio composition towards the risk-free asset when dependence between tail events for the risky assets is accounted for. I further assess the economic importance of this change in portfolio shares, mainly driven by the need to hedge against changes in the stochastic opportunity set when the data generating process incorporates the above-mentioned asymmetric tail behavior, through the certainty equivalent cost of ignoring extreme value dependence and find that taking it into account leads to significant economic gains.

In the third chapter of the thesis I address the problem of solving for optimal portfolio allocation in a dynamic setting, where conditional correlation is modeled using observable factors, which allows me to isolate the demand for hedging correlation risk. I am able to analyze separately the impact of tail dependence through the unconditional distribution and that of conditional correlation on portfolio holdings. With those distinct ways of modeling dependence I aim at replicating the stylized fact of increased dependence during extreme market downturns, rising market-wide volatility, or worsening macroeconomic conditions. I find that both correlation hedging demands and intertemporal hedges due to increased tail dependence have distinct portfolio implications and cannot act as substitutes to each other. As well, there are substantial economic costs for disregarding both the dynamics of conditional correlation and the dependence in the extremes.

Keywords: Markov Chain Monte Carlo, dynamic portfolio allocation, Monte Carlo simulation, tail dependence, copulas, multivariate stationary diffusion, dynamic conditional correlation.

Table des matières

Liste des tableaux

Liste des figures

Liste des Abréviations

GH	....	Generalized Hyperbolic
NIG	....	Normal Inverse Gaussian
MCMC	....	Markov Chain Monte Carlo
SF	....	Simulation Filter
MV	....	Mean Variance
MPRH	....	Market Price of Risk Hedge
IRH	....	Interest Rate Hedge
CIR	....	Cox Ingersoll Ross (process)
CRRA	....	Constant Relative Risk Aversion
HARA	....	Hyperbolic Absolute Risk Aversion
MD	....	Malliavin derivative
DCC	....	Dynamic Conditional Correlation
CCC	....	Constant Conditional Correlation

To my parents who offered me unconditional love and support throughout the course of this thesis.

To my brother who was beside me every step of the way.

To Kris, whose love was my source of motivation and inspiration.

Remerciements

Many people contributed to this thesis in innumerable ways and I wish to express my gratitude to all of them. Foremost, I would like to thank my advisors, Georges Dionne and Tony Berrada, for guiding me through the various meanderings of this project. It would be difficult to overstate my gratitude to Prof. Dionne, a great mentor who saw me through the whole course of my graduate studies, entrusting me with his advice and wisdom, guiding me with his insights and experience, boosting my confidence whenever I lacked it. I thank him for giving me freedom in research, trusting me with many a challenging idea I chose to explore, while always being there for guidance and support.

This project would have not been possible without the enthusiastic encouragement and sound advice of Tony, whose valuable suggestions shaped the direction of my research. He was always there to ask challenging questions, and share insightful thoughts. Many ideas came to life during my visit to HEC Lausanne, where I managed to put my thesis on a fast-track, and I am greatly indebted to him for that.

I have also been very fortunate to have had many great professors in Montreal who made my doctoral student experience a gratifying one. I am especially thankful to Michèle Breton for her continuous support, to Bruno Remillard for various helpful and encouraging suggestions, to Peter Christoffersen for helping me shape quite insightful ideas during his course and providing beneficial comments as a member of my thesis committee.

I am very much indebted to two exceptional people, who in turn played a crucial role in determining my path. To Renée, my very first mentor, who taught me that there are no boundaries that one is not strong enough to surpass. To Jean Bénéteau, whose visionary attitude was the ignition key for my doctoral studies, and whose trust and support never faltered.

A special thanks goes to Claire Boisvert, whose kind and cheerful disposition made my days at HEC so enjoyable, being always there to help, and to Lise Cloutier-Delage, who seamlessly facilitated all administrative procedures with a smile and friendly advice in store. And of course I give my thanks to Mohamed Jabir, always ready to help at the lab during the times of my endless simulations.

I gratefully acknowledge financial support from the *Agence Universitaire de la Francophonie*, HEC Montréal, the Canada Research Chair in Risk Management, the Center for Research on E-finance (CREF), and the *Institut de Finance Mathématique de Montréal* (IFM2).

I wish to further express my gratitude to my friends and colleagues who made my time in Montreal highly enjoyable and productive. Many thanks to all friends from IFAG who joined me here at various times and created a sense of continuity throughout my studies, providing support through hardships and sharing moments of joy. My thanks to Cristi for countless discussions, and to Oana, who was always there to listen to my doubts and share suggestions. Special thanks to my faithful colleague Redouane, with whom so many ideas were bounced back and forth, checked and re-checked, blessed or condemned. I am grateful to my fellow colleagues from the Chair with whom I shared the ups and downs of the life of a graduate student: Khemais, who generously shared with me the office, a bunch of courses and a lot of ideas, Nadia, Thouraya, Hela and Nabil, with whom we have been together from the very beginning, as well as Olfa, Karima, Oussama, Abdelhakim, thanks to all of you. And a special thanks from all my heart to Maria, my office-mate and great friend, who was always there for me when I had doubts or when I needed encouragement.

And finally, this thesis would have never seen the light of day, had it not been for the utter support of my family, who had to deal with 8,000 kilometers separating us during all those years, and who did so admiringly well. I am deeply grateful to my parents for their unfaltering love and encouragement, and to my brother who was always there when I needed him, offering his unconditional support. And to Kris, for being the incredible person he is, accompanying me every step of the way, patiently awaiting the promised life after the thesis, and never letting my heart (or computer) down even through the toughest moments. To you I dedicate my thesis.

Chapter 1

Stock Market Asymmetries: A Copula Diffusion Model

1.1 Introduction

There is wide spread evidence that the distribution of financial asset returns deviates from the assumption of normality both in terms of univariate properties of the data such as excess kurtosis or thick tails, as well as the dependence structure: multivariate normality imposes independence between extreme realizations of the variables, whereas returns are known to be highly correlated during large market downfalls. In a study of several major international market indices Longin and Solnik (2001) provide strong evidence of increased correlation of tail events of asset returns, especially during bear markets.

For risk management applications, multivariate option pricing or portfolio choice decisions it is important to introduce relatively parsimonious models that can capture the above mentioned features of the data. There has been a proliferation of studies in recent literature that propose models for incorporating the asymmetric response of conditional correlation to returns, mainly building upon the Dynamic Conditional Correlation model of Engle (2002). Alternatively, in a jump diffusion framework, Das and Uppal (2004) model high correlation across large changes in asset returns and study their effect on portfolio allocation. Ang and Chen (2002) compare several discrete time models in terms of their ability to reproduce the asymmetric dependence pattern present in stock return data. None of the models, however, succeeds in either picking up the extremal dependence pattern of the data or explaining the degree of correlation asymmetry.

In this chapter we propose a model that is able to accommodate this extremal dependence structure, based on the construction of a multivariate diffusion with a pre-specified stationary distribution that relies on copula theory. Despite of the abundant literature on copulas, not much research has been done for the multivariate dependence modeling of stochastic processes. The study of the dynamic multivariate spatial dependence structure of stochastic processes has found several model applications in a discrete time setting (Jondeau and Rockinger, 2002; Patton, 2004; Fermanian and Wegkamp, 2004). However, the spatial dependence structure of multivariate diffusions in the wider copula context has not been extensively studied. Kunz (2002) proposes a framework for modeling extremes in multivariate diffusions of the gradient field type via the use of copula functions, but limits the attention to a specification with a constant diffusion term that inevitably restricts the ability of the model to account for certain dynamic properties of the data, while fitting the stationary distribution. Instead, we propose a more general model for which the above mentioned construction is a special case. We aim at answering the following issues:

(a) The stochastic process for asset prices should be able to imply a dependence structure that allows for increased dependence between extreme realizations, but should be flexible enough to include the case of asymptotic independence (as implied by the Gaussian distribution). The latter condition comes from the concern, raised by Poon et al. (2004) that using a model which precludes independence in the tails may lead to serious overestimation of the joint risks. To this end we use a mixture of copula functions to tailor the multivariate distribution, as they allow for flexibility in terms of choice of the marginals, and can also be modeled to allow or not for dependence between tail realizations. Based on the copula decomposition between the dependence structure and the marginal distributions, we build a multivariate diffusion with a pre-specified stationary density. Its construction relies on restricting the drift for a given specification of the diffusion term and the stationary density (see Hansen and Scheinkman, 1995; Chen et al., 2002). Thus we obtain a flexible process for asset prices that is able to accommodate a wide array of dependence structures.

(b) While replicating different types of dependence patterns, we would like our model to keep track of univariate properties of asset returns, such as a leptokurtic univariate distribution as compared to the normal, semi-heavy tails (Barndorff-Nielsen, 1995), or volatility clustering expressed as serial correlation of squared log returns. To achieve this, the copula construction leaves us with the flexibility to chose the appropriate marginals. We turn to the Generalized Hyperbolic family of distributions, as their ability to replicate the tail behavior of asset returns, as well as certain dynamic properties as persistence in auto-correlation in squared returns, has been recorded in literature in the context of univariate diffusion modeling (Eberlein and Keller, 1995; Rydberg, 1999; Bibby and Sorensen, 2003). We show that this property is retained in the multivariate model we propose. As well, Jaschke (1997) points out that one could obtain a process for returns with a Generalized Hyperbolic stationary distribution with stochastic volatility as a weak limit of a GARCH model in the sense of Nelson (1990).

The copula functions that we study are the tail-independent Gaussian copula, the symmetric tails Student's t copula and the extreme value Gumbel copula that allows for asymmetric tail behavior in combination with its survival counterpart. The tail dependence coefficients that we estimate point towards a structure with extremal dependence, and the Gaussian diffusion is rejected in favour of an alternative that takes into account tail dependence.

While the stationary distribution of the proposed process is known in closed form, the same cannot be said in general for the transition density, which raises a serious estimation challenge, as an exact likelihood approach cannot be applied. Instead of relying on approximations of the likelihood function in the spirit of Aït-Sahalia (1999) and Aït-Sahalia (2003), as such an approach may prove to be too computationaly intensive when explicit solutions for the density approximation coefficients cannot be obtained, we resolve to a Markov Chain Monte Carlo (MCMC) method to estimate model parameters, following a sequential inference procedure of Golightly and Wilkinson (2006a) in the spirit of Roberts and Strammer (2001) and Durham and Gallant (2002). Further discussion on the available estimation approaches is provided in the subsequent sections.

We also address the question of model selection, using the traditional Bayesian approach based on the marginal likelihood functions of alternative models. Results suggests that models that disregard asymmetric dependence between extreme realizations are rejected in favour of those that take these particular features of the dependence structure into account.

The remainder of the chapter is organized as follows. Section 1.2 discusses the issue of modeling dependence through the use of copula functions. Section 1.3 introduces the process for asset prices, its construction and the particular assumptions on the univariate marginals as well as the dependence structure. Section 1.4 reviews the estimation methodology of the proposed multivariate diffusion based on copula functions using an MCMC estimation algorithm. Section 1.5 discusses the estimation results, focusing on the degree of tail dependence that could be achieved under the proposed model specification, and Section 1.6 concludes.

1.2 Copula functions and dependence modeling

The pitfalls of using the linear correlation coefficient as a dependence measure have been largely discussed in literature. Linear correlation fully describes the dependence patterns only in the elliptical class of distributions that are inevitably characterized by symmetry. It is also an inadequate tool for discerning dependence when it comes to extreme events. Among the deficiencies of linear correlation comes the fact that second moments have to be finite in order for it to be defined. As well, it is not invariant under non-linear strictly increasing transformations of the variables (a transformation that is known to leave the dependence structure unchanged). In contrast, all concordance measures of dependence depend only on the copula property and are invariant to increasing changes in the marginals, while the tail dependence coefficient characterizes the extreme dependence using only the copula specification.

Thus, copula theory provides a natural environment for the search of dependence measures that are better suited for capturing extreme co-movement asymmetries. The main concept behind copulas is the separation of the distribution structure from the univariate marginals, as they are functions that link marginals to their multivariate distribution, following Sklar's theorem. Their parsimonious nature makes them suitable for high-dimensional models, as the ones encountered in portfolio selection problems, while their functional specification could be flexible enough to allow for asymptotic extreme (in)dependence: dependence structures range from those generated by elliptical copulas that maintain the validity of the mean-variance framework , to copulas that are able to express extreme value dependence (like the Gumbel copula, consistent with multivariate extreme value theory). Various dependence measures useful for financial applications (comonotonicity, concordance, quadrant (orthant) and tail dependence) can be expressed in terms of copulas.

Copula functions are a useful tool to construct multivariate distributions. They are used to disentangle the information contained in the marginal distributions from that pertaining to the dependence structure. As they are defined as multivariate distribution functions, they contain all the relevant information with respect to the dependence structure. As well, as copulas are defined over transformed uniform marginals, they contain the information on dependence regardless of the marginal distributions, as these transformed variables are $Uniform\,(0,1)$ regardless of the particular marginal distributions. This last feature makes copulas particularly suitable for developing flexible models based on different univariate distributions that best suit the marginal properties of the data, while leaving the freedom to define separately the most appropriate dependence structure.

1.2.1 Copulas and the dependence structure

A standard treatment of copulas can be found in the monographs of Joe (1997), Nelsen (1999), Embrechts et al. (2002), and Frees and Valdez (1998). Cherubini et al. (2004) offer a comprehensive review of the application of copula functions in finance. The main concept behind them is the separation of the distribution structure from the univariate marginals. A copula can be viewed as a multivariate distribution function on the unit cube, with uniformly distributed marginals. Alternatively, it can be defined as a function $C : [0,1]^n \to [0,1]$ with the following properties:

(P1) for every u in $[0,1]^n$, $C\,(u) = 0$ if at least one coordinate of u is 0; $C\,(u) = u_k$ if all coordinates of u except u_k equal 1;

(P2) C is n-increasing if for each $a, b \in [0,1]^n$ such that $a \leq b$, the volume of the hypercube with corners a and b is positive, that is $V_C\,([a,b]) = \sum sgn\,(c)\,C\,(c) \geq 0$ where c are the vertices of $[a,b]$, and $sgn\,(c) = 1$ if $c_k = a_k$ for even k, $sgn\,(c) = -1$ if $c_k = a_k$ for odd k. For the bivariate case this translates into $V_C\,([u_1,u_2] \times [v_1,v_2]) \equiv C\,(u_1,v_1) + C\,(u_2,v_2) - C\,(u_1,v_2) - C\,(u_2,v_1) \geq 0$ for all $u_1,u_2,v_1,v_2 \in [0,1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$.

An important result concerning copulas is Sklar's representation theorem (Sklar, 1959):

For a multivariate joint distribution function F with marginals $F_1, ..., F_n$, there exists an n-copula C, such that for all x in $\mathbb{R}^n$ we have that:

$$F\,(x_1, ..., x_n) = C\,(F_1\,(x_1)\,, ..., F_n\,(x_n)) \tag{1.2.1}$$

The copula is uniquely determined if all marginal distributions $F_1, ..., F_n$ are continuous, otherwise C is unique on $RangeF_1 \times ... \times RangeF_n$. The converse statement also holds, i.e. for a given copula C with marginals $F_1, ..., F_n$, the function F defined above is an n-dimensional multivariate distribution function. Sklar provides the following corollary: for a multivariate joint distribution function F with continuous margins $F_1, ..., F_n$ and copula C, satisfying the above theorem, and for any $u \in [0,1]^n$, the following holds:

$$C(u_1, ..., u_n) = F\left(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)\right) \tag{1.2.2}$$

In the subsequent sections we will use the copula density decomposition formula that follows from (1.2.1):

$$f(x_1, ..., x_n) = c(F_1(x_1), ..., F_n(x_n)) \prod_{i=1}^{n} f_i(x_i)$$

where $c(\cdot)$ is the copula density and $f_i(\cdot)$ are the univariate PDFs.

A key property of copulas, that makes them particularly well suited for dependence structure modeling, is their invariance under strictly increasing transformations of the marginals. However, this property is true for the linear correlation as a dependence measure only for affine strictly increasing transformations. In particular, if we consider the functions $\alpha(X)$ and $\beta(Y)$ of two random variables X and Y, then the following transformations change the copula functions in a deterministic way (see Nelsen, 1999):

(i) if α, β are strictly increasing, then $C_{\alpha(X),\beta(Y)}(u,v) = C_{X,Y}(u,v)$;

(ii) if α is strictly increasing and β is strictly decreasing, then $C_{\alpha(X),\beta(Y)}(u,v) = u - C_{X,Y}(u, 1-v)$;

(iii) if α, β are both strictly decreasing, then $C_{\alpha(X),\beta(Y)}(u,v) = u + v - 1 + C_{X,Y}(1-u, 1-v)$.

If C is an n-dimensional copula, then it has a known upper and lower bound (the Frechet-Hoeffding bounds):

$$\begin{aligned} L_n(u) &\leq C(u) \leq U_n(u) \\ \text{where } L_n(u) &= \max\left(\sum_{i=1}^{n} u_i - n + 1, 0\right) \\ U_n(u) &= \min(u_1, ..., u_n) \end{aligned} \tag{1.2.3}$$

For $n = 2$ the upper and the lower bound are copulas, but for $n \geq 3$, L_n is the lower bound in the sense that for any $u \in [0,1]^n$ there exists such a copula C, that $C(u) = L_n(u)$ (see Nelsen, 1999).

Following Drouet-Mari and Kotz (2001), the continuity of a copula can be established for each $u, v \in [0,1]^n$, if it satisfies the stronger Lipschitz condition:

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1| \tag{1.2.4}$$

Further on, as $C(u)$ is increasing and continuous in u, it is differentiable almost everywhere, and the following holds:

$$0 \leq \frac{\partial}{\partial u_i} C(u) \leq 1, \qquad i = 1, ..., n$$

For each copula we can define a survival function: $\overline{C}(u, v) = 1 - u - v + C(u, v)$ for the bivariate case, and more generally:

$$\overline{C}(u_1, ..., u_n) = \Pr(U_1 > u_1, ..., U_1 > u_1)$$

Below we discuss briefly several dependence concepts in a copula framework. Following the Frechet-Hoeffding inequality, it can be shown that the upper and the lower bound are both copulas in the bivariate case, and can be thought of as the joint distribution functions of two couples of univariate vectors: $(U, 1 - U)$ for the lower bound and (U, U) for the upper bound. Thus, the lower bound describes the state of perfect negative dependence (two vectors having this copula are said to be countermonotonic), whereas the upper bound corresponds to the state of perfect positive dependence (and the two vectors having this copula are comonotonic).

Following Embrechts et al. (2002), a proper dependence measure δ should have the following properties:

(i) δ should be defined for every pair X, Y;

(ii) $\delta(X,Y) = \delta(Y,X)$;

(iii) $-1 \leq \delta(X,Y) \leq 1$;

(iv) $\delta(X,Y) = 1$ iff X and Y are comonotonic, and $\delta(X,Y) = -1$ iff X and Y are counter-monotonic;

(v) $\delta(\varphi(X), Y) = \delta(X,Y)$ for a strictly increasing function φ, and $\delta(\varphi(X), Y) = -\delta(X,Y)$ for a strictly decreasing function φ.

(vi) $\delta(X,Y) = 0$ iff X, Y are independent.

As there is no dependence measure that satisfies properties (v) and (vi), then we should modify the following properties if we require (vi):

(iii-a) $0 \leq \delta(X,Y) \leq 1$;

(iv-a) $\delta(X,Y) = 1$ iff X and Y are co/counter-monotonic;

(v-a) $\delta(\varphi(X), Y) = \delta(X,Y)$ for a strictly monotone function φ.

Concordance measures can also be defined in terms of the copula. Following Embrechts et al. (2002), if (X,Y) and $\left(\widetilde{X}, \widetilde{Y}\right)$ are two couples of independent vectors with common marginals, then the difference between the probability of concordance and discordance (Q) can be expressed in terms of their corresponding copulas:

$$
\begin{aligned}
\text{If } Q &\equiv \Pr\left[\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) > 0\right] - \Pr\left[\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) < 0\right], \\
\text{then } Q &= Q\left(C, \widetilde{C}\right) = 4 \int\!\!\int_{[0,1]^2} \widetilde{C}(u,v)\, dC(u,v) - 1
\end{aligned}
$$

Kendall's tau $\tau(X,Y)$ and Spearman's rho $\rho_S(X,Y)$ are two measures of concordance that also have copula representation:

$$\tau(X,Y) \equiv Q(C,C) = 4 \int\int_{[0,1]^2} C(u,v)\, dC(u,v) - 1 \tag{1.2.5}$$

$$\rho_S(X,Y) \equiv 3Q(C,\Pi) = 12 \int\int_{[0,1]^2} uv dC(u,v) - 3 \tag{1.2.6}$$

where $\Pi^n(u) = u_1 u_2 ... u_n$ is the independence copula.

When both Kendall's tau and Spearman's rho are equal to $1(-1)$, then the copula of the two vectors is the upper (lower) Frechet bound.

As we are interested in modeling dependence asymmetries in the tails of the distribution, then the tail coefficient, as a measure of dependence in the lower and the upper tail is of particular interest. The coefficient of upper tail dependence is defined as the probability of an extreme event in Y, conditional on an extreme event in X:

$$\begin{aligned} \tau^U &= \lim_{u\to 1} \Pr\left(Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u)\right) \\ &= \lim_{u\to 1} \frac{\Pr\left(Y > F_Y^{-1}(u), X > F_X^{-1}(u)\right)}{\Pr\left(X > F_X^{-1}(u)\right)} \end{aligned} \tag{1.2.7}$$

provided that the limit exists. If $\tau^U \in (0,1]$ then the two vectors of random variables are said to be asymptotically dependent in the right tail. Asymptotic independence is reached for the case of $\tau^U = 0$. Joe (1997) shows that the concept of tail dependence can be related to that of the copula by the following alternative definition of the coefficient for upper tail dependence of a bivariate copula, for which the following limit exists:

$$\tau^U = \lim_{u\to 1} \frac{1 - 2u + C(u,u)}{1-u} \tag{1.2.8}$$

The coefficient of lower tail dependence can be derived in a similar fashion:

$$\begin{aligned} \tau^L &= \lim_{u\to 0} \Pr\left(Y \le F_Y^{-1}(u) \mid X \le F_X^{-1}(u)\right) \\ &= \lim_{u\to 0} \frac{\Pr\left(Y \le F_Y^{-1}(u), X \le F_X^{-1}(u)\right)}{\Pr\left(X \le F_X^{-1}(u)\right)} \\ &= \lim_{u\to 0} \frac{C(u,u)}{u} \end{aligned} \tag{1.2.9}$$

and the notions of asymptotic dependence and independence are analogous to those in the right tail. Having in mind the relationship between a copula and its survivor copula, it can be shown that the coefficient of upper tail dependence of a copula is in fact the coefficient of lower tail dependence of the survivor copula. We will rely on this property in the subsequent modeling of the extreme-value diffusion process.

Despite these asymptotic measures of dependence, we are interested as well in the behavior of random variables as they approach the extremes. This 'near' tail dependence measure is called quantile dependence and it is defined in the following way for quantiles q:

$$\tau(q) = \begin{array}{ll} \Pr\left[U \leq q \mid V \leq q\right] & \text{if } q \leq 0.5 \\ \Pr\left[U > q \mid V > q\right] & \text{if } q > 0.5 \end{array} \tag{1.2.10}$$

1.2.2 Degree of tail dependence asymmetry in the data

In order to get an impression of the degree of tail dependence asymmetry present in the data, consider daily CRSP US stock capitalization decile indices for the period 1990-2005. These indices represent yearly rebalanced portfolios based on market capitalization. The stock universe includes stocks listed on NYSE, AMEX, and NASDAQ. All ten capitalization decile indices were grouped in three sub-categories: small-cap (deciles 1-3), mid-cap (deciles 4-7), and large-cap (deciles 8-10).

The degree of 'near' tail dependence for all three couples of data is displayed using quantile plots on Figure 1.2.1. The dependence does not decay to zero as we go further in the left tail as it would be the case under bivariate normality. As well, for the Large-Mid cap couple quantile dependence is high for both tails, while for the other couples of data it tends towards zero for the right tail, pointing towards asymmetric ('near') tail dependence.

In order to test the significance in the differences in correlation patterns between the left and the right tail, we use the model-free test of dependence symmetry, developed by Hong et al. (2003). The test statistic under a null hypothesis of symmetry exploits the estimates of the exceedence correlations $\left(\rho_q^-, \rho_q^+\right)$ at different quantile levels q and their variance covariance matrix Ω:

$$J = n\left(\rho^+ - \rho^-\right)\Omega^{-1}\left(\rho^+ - \rho^-\right) \xrightarrow{d} \chi_m^2$$

Figure 1.2.1: Quantile dependence plots

Plots of quantile dependence for all three couples of de-trended log-prices of the three CRSP indices formed on the basis of size deciles for the period 1986-2005 (small-cap (deciles 1-3), mid-cap (deciles 4-7), and large-cap (deciles 8-10)).

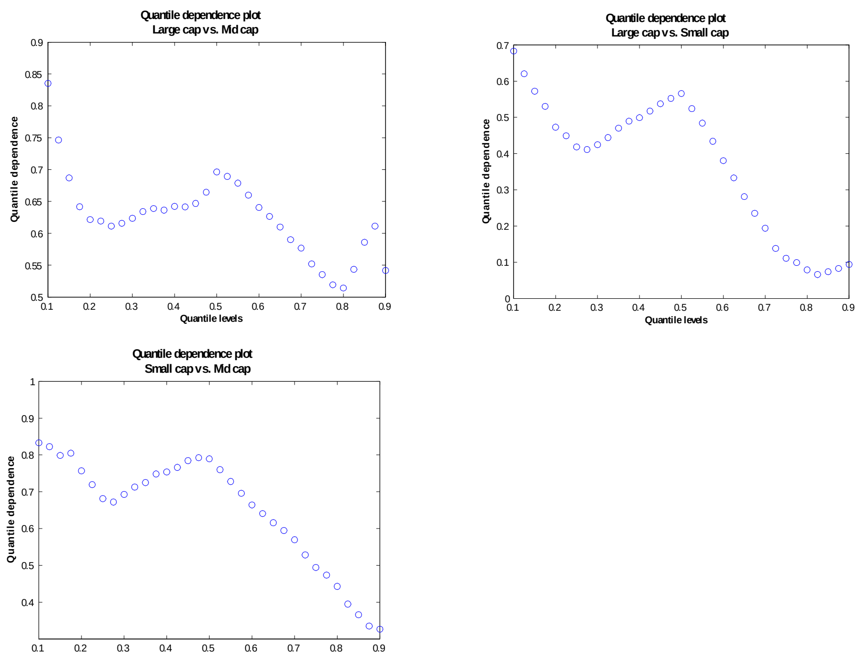


Table 1.2.1: Test of symmetry in the exceedence correlations

The Hong et al. (2003) test of exceedence correlations symmetry in the lower and upper quartiles for the de-trended log-prices of the three CRSP indices formed on the basis of size deciles for the period 1986-2005 (small-cap (deciles 1-3), mid-cap (deciles 4-7), and large-cap (deciles 8-10)). The test statistic is given by:

$$J = n\left(\rho^{+} - \rho^{-}\right)\Omega^{-1}\left(\rho^{+} - \rho^{-}\right) \xrightarrow{d} \chi^2_m$$

where ρ^{+} and ρ^{-} are the exceedence correlations calculated at the corresponding quantile levels, n is the sample size and m is the number of quantile levels considered. Results for three quantile levels (0.85, 0.90, 0.95) are given below:

	Large vs. Mid cap	Large vs. Small cap	Small vs. Mid cap
Test statistic (J)	1.9351	17.6046	13.3933
p-values	(0.5860)	(5.3065e-004)	(0.0039)

where n is the sample size and m is the number of quantile levels considered. Table 1.2.1 summarizes the results of the test, rejecting symmetry for all but the Mid-Large cap couple, for which the quantile dependence plots indicated as well high dependence in both tails.

In the sections that follow we will build a diffusion process that accounts for those dependence features of the data with the help of copula functions. It also accommodates desirable univariate properties of asset returns such as volatility clustering, heavy tails, and slowly decaying autocorrelation function of squared returns, without reverting to a stochastic volatility specification or the introduction of jumps.

1.3 The multivariate copula diffusion model

In the discrete time literature there exist numerous models that are able to replicate both stylized facts of univariate asset returns series, such as thick-tailed asymmetric marginals, volatility clustering, slowly decaying autocorrelation function of squared returns, and asymmetric dependence structure in the extremes of the multivariate distribution. Copula functions have become increasingly popular in multivariate discrete time models, as in Patton (2004), Jondeau and Rockinger (2002) among others. Astonishingly, much less effort has been spent in this respect in continuous time modeling, except for scalar diffusions. Examples include stochastic volatility models (Heston, 1993) or diffusions with jumps in returns and volatility (Eraker et al., 2003), hyperbolic diffusions (Bibby and Sorensen, 1997), generalized hyperbolic diffusions (Rydberg, 1999), time-changed Lévy processes (Carr and Wu, 2004). However, the multivariate spatial dependence structure modeling of diffusions has attracted much less attention. Here we propose a construction of a multivariate diffusion with pre-specified stationary density with arbitrary marginals, coupled by a sufficiently parsimonious copula dependence function that avoids the curse of dimensionality problem, normally encountered in modeling multivariate datasets. The aim is to provide a sufficiently flexible treatment of the univariate return series that is able to accommodate the stylized features of the data, as well as to allow for possible asymmetries in the tail dependence of the multivariate distribution via the copula function.

1.3.1 Constructing a diffusion with a pre-specified stationary distribution

We assume that uncertainty is driven by a d-dimensional standard Brownian motion and that the price of the risky asset can be expressed as [1]:

$$S_{it} = \exp\left(\phi_i\left(t\right) + X_{it}\right), i = 1, ..., d \tag{1.3.1}$$

for some deterministic function of time $\phi_i\left(t\right)$, which we assume to be linear in t, $\phi_i\left(t\right) = k_i t$ with a linear trend parameter k_i, and where

$$dX_t = \mu\left(X_t\right)dt + \Lambda\left(X_t\right)dW_t \tag{1.3.2}$$

Thus, applying Itô's lemma we obtain for the price process for $i = 1, ..., d$:

$$\begin{aligned} dS_{it} &= S_{it}\mu_i^S\left(\ln S_{it} - k_i t\right)dt + S_{it}\sum_{j=1}^{d}\Lambda_{ij}\left(\ln S_{it} - k_i t\right)dW_{jt} \qquad (1.3.3) \\ \text{where } \mu_i^S\left(X_t\right) &= \mu_i\left(X_t\right) + k_i + \frac{1}{2}\sum_{j=1}^{d}\overline{\sigma}_{ij}\left(X_t\right)^2 \end{aligned}$$

where $\overline{\sigma}_{ij}$ are entries of the matrix Λ in the diffusion term of the process for the de-trended log-price X. As pointed out in Bibby and Sorensen (1997), there is empirical evidence that the increments of the process for the log-price are nearly uncorrelated but not independent, which motivates the specification in (1.3.1). It is chosen as the most straightforward

[1] Following the parametrization of Bibby and Sorensen (1997) and Rydberg (1999)

generalization of the Black Scholes model. The exact parametrization of the drift and the diffusion term will be discussed in the subsequent sections, where we present a method to construct a diffusion with a pre-specified stationary distribution.

Before proceeding to the specific construction of the multivariate diffusion process for the state variables X, we fix several conditions that the model specification should satisfy. First, we would like to be able to allow for possibly different univariate processes for each of the state variables. Second, the dependence structure should be constructed independently from the margins, and it should allow for asymptotic dependence and independence. Third, we would desire that the dependence structure be modeled parsimoniously, in order to allow for the treatment of a highly multivariate dataset. The copula construction we pursue allows answering all these conditions.

Following Chen et al. (2002), we construct a multivariate stationary diffusion by exploiting the relationship that exists between the invariant density, the drift and the diffusion term for the process in (1.3.2):

$$\begin{aligned} \mu_j &= \frac{1}{2q} \sum_{i=1}^{d} \frac{\partial \left(v_{ij} q \right)}{\partial x_i} \\ \Sigma &= \Lambda \Lambda^{\intercal} \quad \text{with entries } v_{ij} \end{aligned} \tag{1.3.4}$$

where Λ is a lower triangular matrix, q is a strictly positive continuously differentiable multivariate density function, and Σ is a continuously differentiable positive definite matrix. Using this construction, q is the stationary density of the Markov process, and the drift vector μ is determined by the choice of q and the volatility matrix Σ. Thus, in order to model the stationary diffusion (1.3.2), we need to specify its invariant density and its diffusion term. For the latter, we propose a constant conditional correlation specification, given by:

$$\begin{aligned} v_{ij} &= \rho_{ij} \sigma_i^X \sigma_j^X \\ \sigma_i^X &= \sigma_i \left[\widetilde{f}^i \left(x_i \right) \right]^{-\frac{1}{2}\kappa_i} \end{aligned} \tag{1.3.5}$$

which is a multivariate generalization of the diffusion term proposed by Bibby and Sorensen

(2003) for the case of univariate diffusion, where $\left(\sigma_i^X\right)^2 > 0$ and $\kappa_i \in [0,1], i = 1, ..., d$. The function $\widetilde{f^i}(x_i) \propto f^i(x_i)$, i.e. it is proportional to the ith univariate marginal distribution whose choice will be discussed in the subsequent section. If all parameters κ are set equal to zero, and the correlation matrix is assumed to be diagonal with unit entries, this will reduce the diffusion to one with a constant volatility term, which is the case discussed in Kunz (2002).

In what follows, we will discuss the particular choice for the marginal distributions and the way they are joined using the copula function to obtain the stationary multivariate distribution q that determines the drift (1.3.4) of the diffusion in (1.3.2).

Choice of the marginal distributions

A much exploited distribution specification for the univariate return series in recent literature has been that of the family of the Generalized Hyperbolic distributions. Introduced by Barndorff-Nielsen (1977) for studying the particle-size distribution of wind-blown sand, it has consequently found application in numerous fields, including finance. Distributions in that family have been successfully fitted to financial time series, while stochastic processes, built on the basis of generalized hyperbolic laws, have been proposed to model the dynamics of financial markets. Eberlein and Keller (1995) introduce the hyperbolic Levy motion in modeling the dynamic behavior of asset returns. Their model is further extended in Prause (1999) to the generalized hyperbolic case. Bibby and Sorensen (1997) fit a hyperbolic diffusion model to individual stock price data, while Rydberg (1999) proposes a one-dimensional Normal Inverse Gaussian diffusion that accommodates thick tails in log returns, and Bauer (2000) investigates the usefulness of hyperbolic distributions for risk management in the context of VaR modeling. As the family of Generalized Hyperbolic distributions covers a vast spectrum of tail behavior (from Gaussian to power tails), it is particularly suited for modeling the marginal distributions in the present context of investigating the extremal behavior of return series.

Form, properties and subclasses of the univariate Generalized Hyperbolic (GH) family of distributions. The family of GH distributions is constructed as normal mean-variance mixtures with the Generalized Inverse Gaussian (GIG) as the mixing distribution. Thus, the density function for the GH distribution is expressed as:

$$f_{GH}\left(x;\alpha,\beta,\delta,\mu\right)=\int\limits_0^\infty N\left(x;\mu+\beta s,s\right)GIG\left(s;\lambda,\delta^2,\alpha^2-\beta^2\right)ds \tag{1.3.6}$$

where $N\left(\cdot\right)$ is the normal density with mean $\mu+\beta s$ and variance s, and the GIG density has the form:

$$\begin{aligned} GIG\left(x;\lambda,\chi,\psi\right) &= \frac{\left(\psi/\chi\right)^{\lambda/2}}{2K_\lambda\left(\sqrt{\psi\chi}\right)}x^{\lambda-1}e^{-\frac{1}{2}\left(\chi x^{-1}+\psi x\right)} \\ x &> 0,\lambda\in\mathbb{R},\psi,\chi\in\mathbb{R}_+ \end{aligned} \tag{1.3.7}$$

where K_λ is the modified Bessel function of the third kind with index λ, whose integral representation, following Barndorff-Nielsen and Blaesid (1981) is given by:

$$K_\lambda\left(x\right)=\frac{1}{2}\int\limits_0^\infty y^{\lambda-1}e^{-\frac{x}{2}\left(y+y^{-1}\right)}dy,\quad x>0$$

The fact that the GH class of distributions is obtained via this convolution operation is exploited when simulating random GH variables.

Solving this integral form gives the following probability density function of the univariate GH distribution:

$$\begin{aligned} f_{GH}\left(x;\alpha,\beta,\delta,\mu\right) &= c\left(\lambda,\alpha,\beta,\delta\right)\left(\delta^2+\left(x-\mu\right)^2\right)^{\frac{\lambda-1/2}{2}}\times \\ & \quad K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+\left(x-\mu\right)^2}\right)e^{\beta\left(x-\mu\right)} \\ \text{where } c\left(\lambda,\alpha,\beta,\delta\right) &= \frac{\left(\alpha^2-\beta^2\right)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^\lambda K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \\ x &\in \mathbb{R} \end{aligned} \tag{1.3.8}$$

$c\left(\lambda,\alpha,\beta,\delta\right)$ is the normalizing constant, and the parameters have the following interpretations in terms of the distribution: α determines the shape, β the skewness, μ is a location parameter and δ is a scaling parameter. The parameter domain is:

$$\begin{aligned}
\delta &\geq 0, \alpha > |\beta| \text{ for } \lambda > 0 \\
\delta &> 0, \alpha > |\beta| \text{ for } \lambda = 0 \\
\delta &> 0, \alpha \geq |\beta| \text{ for } \lambda < 0 \\
\mu &\in \mathbb{R}
\end{aligned}$$

GH distributions have semi-heavy tails, given by[2]:

$$\lim_{x \to \pm\infty} f_{GH}\left(x; \lambda, \alpha, \beta, \delta, \mu\right) \sim |x|^{\lambda - 1} \exp\left\{\left(\mp\alpha + \beta\right) x\right\} \tag{1.3.9}$$

Thus the class can easily accommodate any tail behavior ranging from power to exponential decline, and can account for tail asymmetries.

The GH family of distributions has the normal distribution as a limiting case for $\delta \to \infty$, $\delta/\alpha \to \sigma^2$, and the Student's t distribution as a limit for $\lambda < 0$, $\alpha = \beta = \mu = 0$ (Barndorff-Nielsen, 1978; Prause, 1999). The tail behavior for those limiting cases is as follows. For the normal distribution we have very thin exponential tails:

$$\lim_{x \to \pm\infty} f_{Ga}\left(x\right) \sim c \exp\left(-\frac{x^2}{2}\right)$$

while for the Student's t distribution with υ degrees of freedom we have power tails:

$$\lim_{x \to \pm\infty} f_{t}\left(x\right) \sim c\left|x\right|^{-\upsilon - 1}$$

Various special cases can be obtained for a different parameterization of the GH distribution. For $\lambda = 1$ the hyperbolic distribution is obtained:

$$\begin{aligned}
f_H\left(x; \alpha, \beta, \delta, \mu\right) &= c\left(\alpha, \beta, \delta\right) e^{-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)} \\
\text{where } c\left(\alpha, \beta, \delta\right) &= \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1\left(\delta\sqrt{\alpha^2 - \beta^2}\right)} \\
x &\in \mathbb{R}
\end{aligned} \tag{1.3.10}$$

[2] See Prause (1999) and Barndorff-Nielsen and Blaesid (1981).

where $\delta > 0$, $\alpha > |\beta|$, $\mu \in \mathbb{R}$. This parametrization has been widely exploited in literature because of the ease of implementation, as the Bessel function appears only in the normalizing constant. However, it limits the possible tail behavior cases one could obtain, as the tails are allowed exponential decay: $\lim_{x\to\pm\infty} f_H(x;\alpha,\beta,\delta,\mu) \sim e^{(\mp\alpha+\beta)x}$, but nevertheless it has proved to be successful in modeling the dynamic behavior of financial time series.

Another subclass of the GH family is that of the Normal Inverse Gaussian (NIG) distribution, obtained for $\lambda = -1/2$, whose density is given by:

$$\begin{aligned} f_{NIG}(x;\alpha,\beta,\delta,\mu) &= c(\alpha,\delta)\left(\delta^2+(x-\mu)^2\right)^{-\frac{1}{2}} \times \\ & K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right) e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)} \\ \text{where } c(\alpha,\delta) &= \frac{\alpha\delta}{\pi} \\ x &\in \mathbb{R} \end{aligned} \tag{1.3.11}$$

where $\delta > 0$, $\alpha \geq |\beta| \geq 0$, $\mu \in \mathbb{R}$. This specification has been successfully used as the stationary measure of a univariate diffusion in Rydberg (1999) for modeling US stock price data. It has a somewhat richer specification for the tail decay as compared to the hyperbolic distribution: $\lim_{x\to\pm\infty} f_{NIG}(x;\alpha,\beta,\delta,\mu) \sim |x|^{-3/2} e^{(\mp\alpha+\beta)x}$. Also, it is one of the two members of the GH class that are closed under convolution (the other one being the Variance Gamma distribution), so that for the sum of two independent random variables $X_i \sim NIG(x;\alpha,\beta,\delta_i,\mu_i), i = 1,2$ we have that $X_1 + X_2 \sim NIG(x;\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2)$. This property is exploited in Rydberg (1999) when modeling log prices as NIG diffusions in that log returns are expected to be also approximately NIG distributed as the time horizon goes to infinity, provided that there is almost no autocorrelation in the increments of log prices.

The moment generating function for the Generalized Hyperbolic distribution is given by Prause (1999):

$$\begin{aligned} M(u) &= e^{u\mu}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\left(\delta\sqrt{\alpha^2-(\beta+u)^2}\right)}{K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \\ |\beta+u| &< \alpha \end{aligned} \tag{1.3.12}$$

The characteristic function takes the form:

$$\varphi(u) = e^{i\mu u}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\frac{\lambda}{2}}\frac{K_\lambda\left(\delta\sqrt{\alpha^2-(\beta+iu)^2}\right)}{K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \tag{1.3.13}$$

The mean and variance in this class of distributions are given by:

$$\begin{aligned} E[X] &= \mu+\frac{\delta\beta K_{\lambda+1}(\delta\gamma)}{\gamma K_\lambda(\delta\gamma)} \\ Var(X) &= \frac{\delta K_{\lambda+1}(\delta\gamma)}{\gamma K_\lambda(\delta\gamma)}+\frac{\delta^2\beta^2}{\gamma^2}\left(\frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)}-\frac{K^2_{\lambda+1}(\delta\gamma)}{K^2_\lambda(\delta\gamma)}\right)\end{aligned} \tag{1.3.14}$$

where $\gamma^2 = \alpha^2-(\beta+x)^2$. These expressions have a particularly simple form for the NIG distribution, following the property of the Bessel function that:

$$K_{n+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}}e^{-e}\left(1+\sum_{i=1}^{n}\frac{(n+i)!}{(n-i)!i!}(2x)^{-i}\right), n = 0, 1, 2, ...$$

So that for NIG we obtain[3]:

$$\begin{aligned} E[X] &= \mu+\frac{\delta\beta}{\gamma} \\ Var(X) &= \frac{\delta\alpha^2}{\gamma^3} \\ Skew(X) &= 3\delta\alpha^2\beta\gamma^{-5} \\ Kurt(X) &= 3\delta\alpha^2\left(\alpha^2+4\beta^2\right)\gamma^{-7}\end{aligned}$$

In our empirical application we choose the general form of the GH distribution, or its special case – the NIG distribution, because of the general tail behavior allowed under these specifications.

Univariate diffusion specifications with Generalized Hyperbolic stationary distribution. Before turning to the more complex case of multivariate dependence modeling, let us first concentrate on diffusion specifications for the state variables governing the market price of risk process that are susceptible of reproducing features of univariate return data

[3]See Bibby and Sorensen (2003).

like fat tailedness and distribution asymmetry.

We thus consider the construction of a univariate diffusion process with a prespecified stationary distribution. The typical construction of a scalar diffusion exploits the relationship between the stationary density and the densities of the speed and the scale measure. If we assume a univariate state variable diffusion process defined on $S = (l, u) \in \mathbb{R}$, given by the stochastic differential equation (SDE):

$$dX_t = \mu\left(X_t|\theta\right) dt + \overline{\sigma}\left(X_t|\theta\right) dW_t$$

where W is a standard Wiener process, and the drift and the diffusion terms $\mu(x|\theta)$ and $\overline{\sigma}(x|\theta)$, for a parameter vector $\theta \in \Theta$, are such that a unique weak solution exists. We also assume that $\overline{\sigma}(x|\theta) > 0$ for all $x \in S$. Then under certain conditions on the density of the speed and the scale measure of the solution to the above process (see Bibby and Sorensen (2001)) and for a function $\tilde{f}$ that is integrable on S, the process is ergoric and its invariant density μ_θ is proportional to $\tilde{f}$. More specifically, the scale measure of the solution to the above SDE has density with respect to the Lebesgue measure defined as:

$$s(x) = \exp\left(-\int_{x^*}^{x} \frac{2\mu(u)}{v(u)} du\right), \qquad x \in S \tag{1.3.15}$$

for some $x^* \in S$, where $v(x) = \left[\overline{\sigma}(x)\right]^2$, and where we have suppressed the conditioning on the parameter vector θ for the sake of brevity. The speed measure of the solution to the above SDE has density with respect to the Lebesgue measure given by:

$$m(x) = \frac{1}{v(x)\, s(x)}, \qquad x \in S$$

Then under the assumptions that $m(x)$ is finite and that $\int_{x^*}^{u} s(x)\, dx = \int_{l}^{x^*} s(x)\, dx = \infty$ the process X is ergodic and its invariant measure has density proportional to $m(x)$. Further, the relationship between the function $\tilde{f}(x)$, proportional to the invariant density, and the drift and diffusion coefficients of the SDE can be shown to verify:

$$2\mu(x) - v'(x) = v(x) \frac{\tilde{f}'(x)}{\tilde{f}(x)} \tag{1.3.16}$$

This allows us to construct a stationary univariate diffusion with a prespecified invariant density. As this construction leaves either the drift or the diffusion coefficient free to be specified, once the form of the stationary density has been chosen, Bibby and Sorensen (2003) suggest the following specification of the drift, that holds for any diffusion coefficient:

$$\mu(x) = \frac{1}{2} v(x) \frac{d}{dx} \ln\left(v(x) \widetilde{f}(x)\right) \tag{1.3.17}$$

It can be shown that this drift specification is a special case of the multivariate drift restriction (1.3.4) for the univariate case.

Notice that the relationship determining the drift of the stationary diffusion depends only on the ratio $\frac{f'(x)}{f(x)}$, thus it is sufficient to specify the invariant density up to a constant of proportionality. Thus we consider the function $\widetilde{f}(x) \propto f(x)$ that is proportional to the density of the univariate GH distribution (1.3.8). This function enters the volatility term (1.3.5) of the multivariate specification.

If we remain in the univariate context, then the volatility term is given by $\overline{\sigma}(x) = \sigma \widetilde{f}(x)^{-\frac{1}{2}\kappa}$, and we obtain the general form of a stationary diffusion process for the state variable X as the one proposed in Bibby and Sorensen (2003):

$$dX_t = \frac{1}{2}\sigma^2 (1-\kappa) \left[\widetilde{f}(X_t)\right]^{-\kappa-1} \frac{\partial \widetilde{f}(X_t)}{\partial X_t} dt + \sigma \left[\widetilde{f}(X_t)\right]^{-\frac{1}{2}\kappa} dW_t \tag{1.3.18}$$

This specification nests the special cases of a zero drift diffusion (in the case of $\kappa = 1$) or constant diffusion term (in the case of $\kappa = 0$).

The above mentioned models in the family of the Generalized Hyperbolic diffusions are preferred to an alternative specification of Normal Inverse Gaussian Levy processes, proposed in Barndorff-Nielsen (1995), that have grown considerably popular in modeling log returns, because the latter suffer from the deficiency of being incapable of replicating the persistence in correlation in absolute and squared log returns because of the independent Levy increments. This is not shared by the Generalized Hyperbolic diffusions, as pointed out in Rydberg (1999).

In our empirical application we use (1.3.18) to model the de-trended log-prices with the aim of fitting the potentially heavy tails and distributional asymmetry of asset returns in the scalar diffusion case. This univariate treatment will later be used as a guidance as to which marginal distributions to choose when modeling the multivariate dependence structure. This is particularly important, because the use of a given copula function is sensitive to the correct choice of the marginals, and failing to do so would entail model misspecification.

Univariate model validation. In order to check the fit of the proposed model, we proceed to a formal validation procedure for the scalar diffusions, proposed in Pedersen (1995) and applied in Rydberg (1999), which is based on the univariate residuals:

$$u_{t_i} = F\left(t_i, X_{t_i} \mid t_{i-1}, X_{t_{i-1}}; \psi\right) \tag{1.3.19}$$

where $F(\cdot)$ is a transition function $F(x, t \mid y, s; \psi)$ for a given parameter vector ψ that can be estimated via simulation using the dynamic probability transform for a discretized sample of the process $\{X_{\Delta t}\}_{i=1}^{n}$ over the period $t = 1, ..., n$ with a discretization step Δ:

$$\widehat{u}_t = \int_{-\infty}^{X_{t\Delta}} f\left(t\Delta, x \mid (t-1)\Delta, X_{(t-1)\Delta}\right) dx \tag{1.3.20}$$

Under the hypothesis of correct model specification, the series $\{\widehat{u}_t\}_{t=1}^{n}$ is $i.i.d.U(0,1)$.

Having chosen the appropriate univariate model, we can now proceed to the problem of building the multivariate distribution with the use of copula functions. However, let us note that there exists a multivariate version of the GH distribution that could be a potential candidate for a straightforward generalization to higher dimensions. Still, there is one important caveat: it is tail-independent. Thus we proceed to the multivariate diffusion construction based on copula functions that allow us to address the problem of modeling different dependence structures independently of the marginals that could differ across the separate univariate data series (another feature that could not be addressed by the multivariate form of the GH distribution).

Choice of the copula and the multivariate stationary diffusion

We now turn to the construction of the n-variate diffusion that has its spatial dependence structure in the stationary density modeled by a specific parametric copula. Following Sklar's theorem, we define the invariant density as:

$$q\left(x_1, ..., x_n\right) \equiv \widetilde{c}\left(x_1, ..., x_n\right) \prod_{i=1}^{n} \widetilde{f}_i\left(x_i\right) \tag{1.3.21}$$

where $\widetilde{c}\left(x_1, ..., x_n\right) = c\left(F^1\left(x_1\right), ..., F^n\left(x_n\right)\right)$, $\widetilde{f}_i\left(\cdot\right)$is proportional to the univariate GH distribution (1.3.8), and $F^i\left(\cdot\right)$ is its corresponding CDF.

In order to account for different degrees of upper and lower tail dependence, we consider several parametric families of copulas that have either no tail dependence (Gaussian copula), or symmetric tail dependence (Student's t copula), or that allow for different degrees of dependence in the left and in the right tail (several Archimedean copulas). Below we discuss the form and properties of the copula functions that we consider for the stationary distribution of the multivariate diffusion for the state variables driving asset prices, and the alternatives we have to build truly multivariate copula functions (of a dimension higher than two) that maintain some degree of parsimony.

Elliptic copulas. We consider two elliptical copulas, the Gaussian and the t copula, that are characterized by symmetry in the dependence structure. Our benchmark model relies on the Gaussian copula. In this case, dependence is governed by the correlation matrix R_{Ga}. Its CDF is defined as:

$$\begin{aligned} & C^{Ga}\left(u_1, u_2, ..., u_d \mid R_{Ga}\right) \\ = & \int_{-\infty}^{\Phi^{-1}(u_1)} ... \int_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{2\pi\left|R_{Ga}\right|^{1/2}} \exp\left\{-\frac{1}{2} x^{\intercal} R_{Ga}^{-1/2} x\right\} dx_1 ... dx_d \end{aligned} \tag{1.3.22}$$

where $\Phi^{-1}\left(u\right)$ denotes the inverse of the univariate standard normal CDF. The Gaussian copula generates a multivariate normal distribution iff the marginal distributions are also normal. It has no upper or lower tail dependence for imperfectly correlated random variables. Thus, for any pair(u_i, u_j) the bivariate tail dependence coefficients are zero: $\tau_{Ga}^{U} = \tau_{Ga}^{L} = 0$.

The Student's t copula, however, allows for both upper and lower tail dependence, but the tail dependence coefficients are equal. Its CDF is given by:

$$C^t\left(u_1,u_2,...,u_d \mid R_T,\nu\right) \tag{1.3.23}$$
$$= \int\limits_{-\infty}^{t_\nu^{-1}(u_1)} ... \int\limits_{-\infty}^{t_\nu^{-1}(u_d)} \frac{\Gamma\left(\frac{\nu+d}{2}\right)\left|R_T\right|^{1/2}}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{d/2}}\left(1+\frac{1}{\nu}x^\intercal R_T^{-1}x\right)^{-\frac{\nu+d}{2}} dx_1...dx_d$$

where ν is the degrees of freedom parameter,R_T is the correlation matrix, and $t_\nu^{-1}(u)$ is the inverse of the univariate CDF of the Student's t distribution with ν degrees of freedom. For any pair (u_i,u_j) the tail dependence coefficient is given by $\tau_T^U = \tau_T^L = 2t_{\nu+1}\left(-\sqrt{v+1}\sqrt{1-\rho_{ij}}/\sqrt{1+\rho_{ij}}\right)$, where ρ_{ij} is the $(i,j)^{th}$ off-diagonal element of R_T. Thus, the tail dependence coefficient decreases for higher values of the degrees of freedom parameter and in the limit it goes to zero as $\nu \to \infty$ (in this case the Student's t copula converges to the Gaussian copula). One interesting property of the t copula is the fact that it can still show tail dependence even if the correlation is zero.

Archimedean copulas. Copulas in this family are constructed using a continuous, decreasing and convex generator function $\varphi(u) : [0,1] \to [0,\infty)$ that has a defined pseudo-inverse $\varphi^{[-1]}(\varphi(u)) = u$ for all u in $[0,1]$:

$$\varphi^{[-1]}(u) = \left\{ \begin{array}{c} \varphi^{-1}(u) \ \text{ for } 0 \leq u \leq \varphi(0) \\ 0 \text{ for } \varphi(0) \leq u \leq \infty \end{array} \right\}$$

The pseudo-inverse is given by the usual inverse for the cases when we have a strict generator function φ. Then the Archimedean copulas are defined in terms of the generator function as follows:

$$C(u_1,u_2,...,u_n;\alpha) = \varphi^{-1}\left(\varphi(u_1;\alpha)+\varphi(u_2;\alpha)+...+\varphi(u_n;\alpha)\right) \tag{1.3.24}$$

for a given dependence parameter α. The density of Archimedean copulas for the bivariate case is given by (see Nelsen, 1999):

$$c(u_1,u_2) = \frac{-\varphi'\left(C(u_1,u_2)\right)\varphi'(u_1)\varphi'(u_2)}{\left(\varphi'\left(C(u_1,u_2)\right)\right)^3}$$

Archimedean copulas have the useful property that most dependence measures, including the coefficients of upper and lower tail dependence, can be expressed in terms of the generator function. Joe (1997) provides the following result with respect to tail dependence: for a strict generator $\varphi(u)$, if $\varphi'(0)$ is finite and different from zero, then the copula has no tail dependence. The copula has upper tail dependence for $1/\varphi'(0) = -\infty$, given by:

$$\tau^U = 2 - 2\lim_{z \to 0_+} \frac{\varphi'(z)}{\varphi'(2z)}$$

and lower tail dependence, given by:

$$\tau^L = 2\lim_{z \to +\infty} \frac{\varphi'(z)}{\varphi'(2z)}$$

Kendall's tau also has a representation in terms of the generator function, given by Genest and MacKay (1986):

$$\tau = 4\int\limits_{[0,1]} \frac{\varphi(z)}{\varphi'(z)} dz + 1$$

A member of the Archimedean family of copulas that we consider is the Gumbel copula, introduced by Gumbel (1960). It is a parsimonious one-parameter copula, whose generator is given by $\varphi(x) = (-\log(x))^{\frac{1}{\alpha}}, \alpha \in (0,1]$, so that its CDF can be expressed as:

$$C_\alpha^G(u_1, u_2, ..., u_n) = \exp\left(-\left(\sum_{i=1}^n (-\log u_i)^{\frac{1}{\alpha}}\right)^\alpha\right), \quad \alpha \in (0,1] \tag{1.3.25}$$

Its Kendall's tau is given by $\rho_\tau = 1 - \alpha$, and the coefficient of upper tail dependence is given by $\tau_G^U = 2 - 2^\alpha$, while the coefficient of lower tail dependence is zero. Independence is achieved for $\alpha = 1$, in this case both tail dependence coefficients are zero.

As we are particularly interested in the lower tail dependence, we have to use the survival Gumbel copula to allow for it. The survival copula for the bivariate case can be defined in terms of the copula function (see Theorem 4.7 in Cherubini et al. (2004) for dimensions bigger than 2):

$$\begin{aligned} \overline{C}_{\overline{\alpha}}^G(u,v) &= u + v - 1 + \exp\left(-\left[(-\log(1-u))^{\frac{1}{\overline{\alpha}}} + (-\log(1-v))^{\frac{1}{\overline{\alpha}}}\right]^{\overline{\alpha}}\right) \\ \overline{\alpha} &\in (0,1] \end{aligned} \tag{1.3.26}$$

Its Kendall's tau is given by $\rho_\tau = 1 - \overline{\alpha}$, and the coefficient of lower tail dependence is given by $\tau^L_{SG} = 2 - 2^{\overline{\alpha}}$ while the coefficient of upper tail dependence is zero.

The symmetrized Joe-Clayton (SJC) copula. A bivariate copula function that has both upper and lower tail dependence is the 'BB7' copula of Joe (1997), also known as the Joe-Clayton copula. It is given by:

$$\begin{aligned}
& C^{JC}\left(u_1, u_2 \mid \tau^L, \tau^U\right) \\
&= 1 - \left\{1 - \left[\left(1-(1-u_1)^{\kappa}\right)^{-\gamma} + \left(1-(1-u_2)^{\kappa}\right)^{-\gamma} - 1\right]^{-\frac{1}{\gamma}}\right\}^{\frac{1}{\kappa}} \\
\text{where } \kappa &= \frac{1}{\log_2\left(2-\tau^U\right)} \\
\gamma &= -\frac{1}{\log_2\left(2-\tau^L\right)} \\
\tau^U &\in (0,1)\,, \quad \tau^L \in (0,1)
\end{aligned}$$

The two parameters of the Joe-Clayton copula are indeed the coefficients of upper (τ^U) and lower (τ^L) tail dependence. As it is claimed in Patton (2004), this copula function suffers from the drawback that even if both parameters are equal, there is still some residual asymmetry in the copula due to its functional form. So we consider instead its 'symmetrized' version, proposed by Patton (2004), whose form is as follows:

$$\begin{aligned}
& C^{SJC}\left(u_1, u_2 \mid \tau^L, \tau^U\right) \\
&= \frac{1}{2}\left[C^{JC}\left(u_1, u_2 \mid \tau^L, \tau^U\right) + C^{JC}\left(1-u_1, 1-u_2 \mid \tau^L, \tau^U\right) + u_1 + u_2 - 1\right]
\end{aligned}$$

However, this copula function has only a bivariate representation, while the one-parameter Archimedean copulas can easily be extended to higher dimensions. In what follows, we consider a nested version of the Archimedean copula that provides an extension to higher dimensions without imposing excessive restrictions on the dependence structure.

Nested Archimedean copulas. Popular approach in literature consists in choosing the same dependence parameter for all univariate marginals, as in (1.3.24), but this seems an implausible restriction on the dependence structure for more than two dimensions, as the pairwise dependence between each couple of random variables would be exactly the same.

A remedy to this problem has been proposed in recent literature in terms of the so-called nested copula construction for the family of Archimedean copulas (Whelan, 2004; Embrechts et al., 2002). The idea of the construction is as follows. Instead of using (1.3.24), we could construct a multivariate Archimedean copula by repeatedly nesting bivariate copulas. For the tri-variate case the copula will thus have the form:

$$C(u_1, u_2, u_3) = \varphi_2^{-1}\left(\varphi_2\left(\varphi_1^{-1}\left(\varphi_1(u_1) + \varphi_1(u_2)\right)\right) + \varphi_2(u_3)\right) \tag{1.3.27}$$

where each generating function $\varphi_i(u_i)$ has its own dependence parameter α_i. With this construction we achieve $(n-1)$ different pairs of variables, which are still below the general case, but this is a considerable improvement compared to the simple form in (1.3.24). However, there are certain conditions that the parameters should satisfy in order for the nested copula to be a valid copula function (see Embrechts et al. (2002) for a discussion). For the parameterization of the Gumbel copula it can be shown that the parameters in each generating function have to satisfy the condition $\alpha_1 \leq \alpha_2$, i.e. dependence should be higher in the more deeply nested copulas[4].

The parsimonious structure of the Gumbel copula makes it a suitable candidate for a nested copula, so we consider it in our application. However, it allows for only upper tail dependence, so we combine it in a mixture copula with its survival counterpart in order to allow for lower tail dependence as well, as we describe over the following lines.

The mixture copulas. Combining both Gumbel and survival Gumbel copulas in a mixture copula, where each function is assigned a certain weight, is a way to construct a copula that has both lower and upper tail dependence with different tail dependence coefficients. However, if we mix only extreme value copulas, we will implicitly assume asymptotic tail dependence for all cases where $\alpha \neq 1$. Following the Poon et al. (2004) critique, and in order to allow for asymptotic tail independence, we include the Gaussian copula in this mixture model, to obtain:

[4] Usually the Gumbel copula parameter is defined as $\gamma = \frac{1}{\alpha}$, $\gamma \in (1, \infty)$, and higher dependence will translate in higher levels of γ. But for estimation purposes, we chose the alternative parametrization, using $\alpha \in (0, 1]$, so that higher dependence requires a lower level of α.

$$C_m^{Ga}\left(u; R_{Ga}, \alpha, \overline{\alpha}, \omega, \varpi\right) = \omega C^G\left(u; \alpha\right) + \varpi \overline{C}^G\left(u; \overline{\alpha}\right) + \left(1 - \omega - \varpi\right) C^{Ga}\left(u; R_{Ga}\right) \quad (1.3.28)$$

or the Student's t copula:

$$C_m^t\left(u; R_T, \upsilon, \alpha, \overline{\alpha}, \omega, \varpi\right) = \omega C^G\left(u; \alpha\right) + \varpi \overline{C}^G\left(u; \overline{\alpha}\right) + \left(1 - \omega - \varpi\right) C^T\left(u; R_T, \upsilon\right) \quad (1.3.29)$$

where we are mixing the two extreme value copulas: the nested Gumbel copula $C^G(u;\alpha)$, where α is the vector of dependence parameters α_i that determine upper tail dependence, the nested survival Gumbel copula $\overline{C}^G(u;\overline{\alpha})$, where $\overline{\alpha}$ is the vector of dependence parameters $\overline{\alpha}_i$ that determine lower tail dependence, with two elliptic copulas: the Gaussian copula $C^{Ga}(u;R_{Ga})$ with correlation matrix R_{Ga} in (1.3.28), or the Student's t copula with a correlation matrix R_T and a degrees of freedom parameter υ in (1.3.29). The key difference between the two mixture copulas consists in the fact that the one based on the Gaussian copula allows for tail independence by setting the extreme value copula weights to zero, while for the Student's t case there is still some degree of tail dependence, even if the correlation parameter of the Student's t copula is zero. Thus, we achieve varying degrees of tail dependence or asymmetry. Further, $u = (u_1, u_2, ..., u_n)^\intercal$ is the vector of marginal CDFs of the random variables, and $\{\omega, \varpi\} \in [0,1], \omega + \varpi \leq 1$ are the corresponding weights for the Gumbel and the survival Gumbel copulas.

So far we have collected all the building blocks of the multivariate diffusion for the state variables driving the stock price process, so in what follows we turn to the task of its estimation.

1.4 MCMC estimation of the multivariate copula diffusion

The above construction of a stationary diffusion with a prespecified stationary density (1.3.1)-(1.3.5) poses a serious estimation problem. Even though the invariant density is explicitly known, this cannot be said for the conditional density of the state variables. Thus, exact likelihood estimation cannot be applied in this case. There is a variety of methods proposed in literature to deal with the estimation of a diffusion with an unknown conditional distribution. Aït-Sahalia (2003) proposes closed-form expansions of the likelihood function both for univariate and multivariate discretely sampled diffusions, based on Hermite polynomials and Taylor expansion of some fixed order. While this method seems well suited for the problem at hand, it could become too computationaly intensive in the cases where no explicit solutions for the coefficients of the density approximation can be found. Bibby and Sorensen (1995) and Rydberg (1999) propose another estimation technique that relies on approximating the conditional density by a normal density and applying a martingale estimation technique. However, even though the martingale estimator is consistent and asymptotically normally distributed, it rests inefficient.

To solve this problem, Tse et al. (2004) propose an alternative way of dealing with the problem of unknown transition density - the MCMC estimation for a hyperbolic diffusion. Relying on a discretization of the underlying diffusion, they apply a random-walk Metropolis Hastings algorithm in order to estimate parameters. However they assume that the discrete time intervals given by observation times are accurate enough to approximate the transition density. If the available data is not fine enough, this approach would introduce discretization bias.

A suitable alternative to deal with the problem of the discretization bias for a highly non-linear (multivariate) diffusion that we apply in this setting, is data augmentation, i.e. introducing latent data points between each pair of observations. This technique has been used in Pedersen (1995) for simulated maximum likelihood estimation of diffusions, or in Jones (2003), Elerian et al. (2001), Roberts and Strammer (2001), and Eraker (2001) for MCMC analysis. The simulated maximum likelihood method relies on a discretization scheme such as the Euler scheme to approximate the one-period-ahead transition density. The MCMC approaches on the other hand propose simulated paths of latent data that bridge two consecutive observations, constraining both ends of the simulated path to be equal to the actual data. Thus, conditioning on both the beginning and the end of each observation sub-period reduces the variance of the simulated latent data and augments the efficiency of the algorithm.

However, augmentation schemes are susceptible to causing slow rates of convergence of the resulting Markov chain due to the dependence between the latent data points and the volatility of the diffusion as the degree of augmentation increases (known as the Roberts and Strammer (2001) critique). There have been several remedies to this issue proposed in recent literature, as the particular transformation of the diffusion process to one with constant volatility proposed by Roberts and Strammer (2001), the simulation filter for multivariate diffusions of Golightly and Wilkinson (2006a) that builds upon the sequential parameter estimation procedure of Johannes et al. (2004) for discrete-time stochastic volatility models, or the Gibbs sampler of Golightly and Wilkinson (2006b) that iterates between updates of parameter and states and relies upon conditioning on the Brownian increments instead of the underlying latent data in order to overcome the dependence with volatility parameters.

The estimation scheme we propose to apply in the present setup relies on an MCMC estimation algorithm with data augmentation for both the univariate and the multivariate diffusion specifications. It follows the sequential inference procedure of Golightly and Wilkinson (2006a) and is closely related to the work of Roberts and Strammer (2001) and Durham and Gallant (2002). As the augmentation of the parameter and state space with latent data points is the corner stone in each MCMC algorithm for diffusion estimation, we will first discuss the particular scheme that was chosen and the motivation behind it.

1.4.1 Data augmentation

Consider a d-dimensional Itô diffusion given by:

$$dY_t = \mu\left(Y_t\right)dt + \sigma\left(Y_t\right)dW_t \tag{1.4.1}$$

Let data be observed at times $t_0 < t_1 < ... < t_{n-1} < t_n$ with a time increment $\Delta\tau = t_{i+1} - t_i$. We divide each subinterval between observations in m equidistant points, so that we obtain an augmented data matrix:

$$Y^{aug} = [\ \overline{Y}_{t_0,0} \quad Y_{t_0,1} \quad ... \quad Y_{t_0,m} \quad \overline{Y}_{t_1,0} \quad ... \quad \overline{Y}_{t_{n-1},0} \quad ... \quad Y_{t_{n-1},m} \quad \overline{Y}_{t_n,0}\ ],$$

where $Y_{t_i,j}$ is a d-dimensional vector of latent data points at time $t_i + j\Delta\tau/\left(m+1\right)$ and $\overline{Y}_{t_i,0}$ is the vector of observations at time t_i. The augmented data matrix could also consist of unobservable state variables, whose treatment would be similar to that of the latent data. Working with the Euler discretization of the process, the joint posterior of data and model parameters θ is given by:

$$\pi\left(Y;\theta\right)\propto\pi\left(\theta\right)\prod_{t=t_0}^{t_n-1}\left\{\prod_{j=0}^{m}\pi\left(Y_{t,j+1}\mid Y_{t,j};\theta\right)\right\} \tag{1.4.2}$$

where $\pi\left(\theta\right)$ is the prior density for the parameter vector, and $\pi\left(Y_{t,j+1}\mid Y_{t,j};\theta\right)$ comes from the Gaussian transition density implied by the Euler discretization:

$$\pi\left(Y_{i+1}\mid Y_i;\theta\right)=\phi\left(Y_i+\mu\left(Y_i\right)\Delta t,\sigma\left(Y_i\right)\sigma\left(Y_i\right)^{\intercal}\Delta t\right)$$

where $\Delta t=\Delta\tau/\left(m+1\right)$ is the time increment between successive data points in the augmented vector Y^{aug}, and $\phi\left(\widetilde{\mu},\widetilde{\sigma}\right)$ denotes the Gaussian density with mean $\widetilde{\mu}$ and covariance matrix $\widetilde{\sigma}$.

Inference procedures that rely on a Gibbs sampler use the conditional posterior for parameters given data and the conditional posterior of missing data given parameters and observations, rather than the joint posterior (1.4.2), and iteratively propose parameters and missing data from each one of them, so that the obtained simulated sequence of parameters and missing data (after an initial burn-in stage) forms a Markov chain whose stationary distribution is the posterior in question. The alternative approach that we apply is the joint update of parameters and states, which overcomes the problem of increasing correlation between the volatility parameters and latent data as the degree of augmentation becomes large. But as it is virtually infeasible to update all latent data in one single block, this sampling scheme can be applied in a sequential manner, updating parameters and unobserved state variables as each observation becomes available.

A straightforward procedure for sampling the latent data points has been proposed by Eraker (2001). It can easily deal with high-dimensional problems, including unobserved state variables. It consists of designing an Accept-Reject Metropolis Hastings algorithm for updating one column of data at a time, where the conditional posterior of one column of missing data is defined as:

$$\pi\left(Y_i\mid Y_{\setminus i};\theta\right)\propto p\left(Y_i\mid Y_{i-1},Y_{i+1};\theta\right)$$

following the Markov property of the diffusion. At each iteration h the algorithm proposes a latent data point Y_i^* from some proposal density (Eraker uses a normal proposal $q\left(\cdot\mid Y_{i-1}^h,Y_{i+1}^{h-1};\theta^h\right)$), which is then accepted or not following the acceptance procedure of the Accept-Reject Metropolis Hastings algorithm of Tierny (1994). The sampling scheme, proposed by Elerian et al. (2001), is essentially the same, but instead of updating one column vector at a time, they propose updating blocks of missing data with random size. However, increasing the number of imputed data points m, while reducing the discretization bias of the Euler approximation, seems to adversely affect the mixing properties of the algorithm because of the increasing correlation of the diffusion parameters and the simulated path as m increases. In fact, when the number of latent data points tends to infinity, one could very precisely estimate the diffusion term by the quadratic variation, so that when updating the diffusion parameter, its posterior distribution given the simulated latent path tends towards a point mass at its previous iteration value, rendering it impossible to update the parameter.

Roberts and Strammer (2001) propose a reparametrization of the missing data that circumvents the problem of reducible data augmentation. The basic idea behind their scheme is a construction of the latent path that does not depend on the diffusion term. They apply the sampling algorithm on a univariate diffusion with constant diffusion term, as well as on a reducible diffusion in the sense of Aït-Sahalia (2003) that has a deterministic time-varying diffusion term, and that could be transformed to a constant volatility diffusion following the Doss transformation.

Their methodology could easily be extended to the estimation of a reducible multivariate diffusion, such as the constant volatility specification considered in Kunz (2002), that is a special case of the model we propose, but for a general multivariate diffusion as in (1.3.5) it is almost impossible to solve for the volatility transformation. Therefore, we use a more promising approach that is applicable for the multivariate specification we are proposing, which consists in the joint update of parameters and states following the sequential MCMC method of Golightly and Wilkinson (2006a). It does not rely on a volatility transformation for the diffusion and at the same time overcomes the Roberts and Strammer critique to data augmentation. As a direct draw from the joint posterior of the model's parameters and the latent state variables is virtually impossible due to the dimension of the state space, a solution to proceed is to revert to Bayesian sequential filtering, devising an MCMC scheme that updates parameters as each new observation becomes available. This idea has been exploited in Stroud et al. (2004), Johannes et al. (2004), Liu and West (2001) among others.

In what follows, we will briefly discuss the algorithm that has been applied in Golightly and Wilkinson (2006a) for the estimation of a general multivariate diffusion. It has proved to have better convergence properties than the standard Gibbs sampler that iteratively updates parameters and states.

1.4.2 The sequential parameter and state estimation scheme

Let us consider that we are at time $t_{j+m} = t_M$ and that we observe $\overline{Y}_{t_{j+m}} = \overline{Y}_{t_M}$, and also suppose that we have a sample of size MC from the marginal parameter posterior distribution $\pi\left(\theta \mid \overline{Y}_{t_j}\right)$,where $\overline{Y}_{t_j}$ denotes all the observed data up to time t_j. As we are interested in sampling the set of parameters from their marginal posterior $\pi\left(\theta \mid \overline{Y}_{t_M}\right)$, we could do so by formulating the joint posterior for parameters and latent data $\pi\left(\theta, Y_{t_M} \mid \overline{Y}_{t_M}\right)$ and then integrating out the latter, where Y_{t_M} denotes all the latent data points up to time t_M. Notice that the marginal parameter posterior at time t_M can be rearranged as follows:

$$\begin{aligned}\pi\left(\theta \mid \overline{Y}_{t_M}\right) &= \int\limits_{Y_{t_M}^{aug}} \pi(\theta) \prod_{i=0}^{M-1} \pi\left(Y_{t_{i+1}}^{aug} \mid Y_{t_i}^{aug}; \theta\right) \\ &= \pi\left(\theta \mid \overline{Y}_{t_j}\right) \int\limits_{Y_{t_M}^{aug} \setminus \left\{Y_{t_j}^{aug}\right\}} \prod_{i=j}^{M-1} \pi\left(Y_{t_{i+1}}^{aug} \mid Y_{t_i}^{aug}; \theta\right)\end{aligned} \tag{1.4.3}$$

So that our target density at time t_M would be

$$\pi\left(\theta \mid \overline{Y}_{t_M}\right) = \pi\left(\theta \mid \overline{Y}_{t_j}\right) \prod_{i=j}^{M-1} \pi\left(Y_{t_{i+1}}^{aug} \mid Y_{t_i}^{aug}; \theta\right)$$

with the augmented data for the interval (t_j, t_M) integrated out.

In order to sample from this target density, we need to devise a Metropolis-Hastings algorithm that will propose parameter and latent data points and will accept or reject those proposals given a certain probability.

The parameter proposal

The approach taken by Golightly and Wilkinson (2006a) that we apply here, and also used in Liu and West (2001), consists in proposing the parameter set θ using a kernel density estimate of the marginal parameter posterior $\pi\left(\theta \mid \overline{Y}_{t_j}\right)$ with the kernel shrinkage correction of Liu and West (2001) that takes care of the over-dispersion of the kernel density function compared to the posterior sample. Thus, we draw the proposal sample of parameters from the following density:

$$
\begin{aligned}
\theta^* &\sim \phi\left(\alpha\theta_u + (1-\alpha)\,\overline{\theta}, h^2 V\right) \\
\alpha^2 &= 1 - h^2 \\
h^2 &= 1 - \left((3\delta - 1)/2\delta\right)^2
\end{aligned}
\qquad (1.4.4)
$$

for a discount factor δ, where ϕ denotes the Gaussian density, and u is an integer that has been drawn uniformly from $\{1, 2, ..., MC\}$. This parameter proposal scheme simplifies considerably the expression for the acceptance probability, as at each observation time t_j we sample from the previous posterior density $\pi\left(\theta \mid \overline{Y}_{t_j}\right)$, so that it will enter both the target posterior density and the proposal, and thus be cancelled out in the calculation of the acceptance probability.

The latent data points proposal

The idea behind the proposal density q from which the proposal latent data points will be sampled is that it should satisfy $\sup(q) \subseteq \sup(p)$ where p denotes the target density π in its unnormalized form. A good proposal would be one that makes the ratio p/q as close to a constant as possible. This is especially important for independence samplers, as the one used in this setting, as pointed out in Tierny (1994), in order to avoid that the algorithm spends too much time in a certain region of parameter space that it explores.

We apply a Modified Diffusion Bridge proposal for the latent data, based on an Euler scheme for the transition density. The idea behind it is quite simple: a Brownian bridge is in fact a Brownian motion that is conditioned upon terminating at a specific value within the interval of interest, that is, it bridges the values at each end of the interval. Using such a Brownian bridge is a way to reduce variance in Monte Carlo integration and Durham and Gallant (2002) show that it compares nicely to other transition density approximations like the Milstein scheme. Thus, the proposal for the latent data points takes the form:

$$\begin{aligned} q\left(Y_{t_{i+1}} \mid Y_{t_i}, \overline{Y}_{t_M}; \theta\right) &= \phi\left(Y_{t_{i+1}}, Y_{t_i} + \widetilde{\mu}_i, \widetilde{\sigma}_i\right) \\ \text{where } \widetilde{\mu}_i &= \frac{1}{M-i}\left(\overline{Y}_{t_M} - Y_{t_i}\right) \\ \widetilde{\sigma}_i &= \Delta t \frac{1}{M-i}\left(M-i-1\right)\sigma\left(Y_{t_i}\right) \end{aligned} \tag{1.4.5}$$

where ϕ denotes the Gaussian density and $\sigma(Y_{t_i})$ is the volatility term of the process for Y from (1.4.1). Thus for each iteration $s = 1, ..., MC$ we sample a latent data path $Y^*_{t_j}, ..., Y^*_{t_{M-1}}$, so we have the joint proposal sample

$$\left(Y^*_{t_j}, ..., Y^*_{t_{M-1}}; \theta\right) \sim \pi\left(\theta \mid \overline{Y}_{t_j}\right) \prod_{i=j}^{M-2} q\left(Y^*_{t_{i+1}} \mid Y^*_{t_i}, Y^*_{t_i}, \overline{Y}_{t_M}; \theta\right) \tag{1.4.6}$$

A Metropolis-Hastings algorithm moves as follows: provided that we have obtained the proposed sample at iteration s and that we have a parameter and latent state sample obtained from the previous iteration $s-1$, we decide whether to keep the parameters and latent data from the previous iteration or alternatively replace them with the ones from the proposal. To this end we form the ratio

$$A = \frac{p\left(Y^*_s, \theta^*_s\right)\widetilde{q}\left(Y_{s-1}, \theta_{s-1}\right)}{p\left(Y_{s-1}, \theta_{s-1}\right)\widetilde{q}\left(Y^*_s, \theta^*_s\right)}$$

where $(Y^*_s, \theta^*_s) = \left(Y^*_{t_j}, ..., Y^*_{t_{M-1}}; \theta^*\right)_s$ is the proposed sample at iteration s, $(Y_{s-1}, \theta_{s-1}) = \left(Y_{t_j}, ..., Y_{t_{M-1}}; \theta\right)_{s-1}$ is the previously accepted sample at iteration $s-1$, p denotes the target posterior density in its unnormalized form, and $\widetilde{q}$ is the proposal density (1.4.6). Replacing all terms in the expression, we obtain for the ratio A:

$$A = \frac{\prod_{i=j}^{M-1} \pi\left(Y^*_{t_{i+1}} \mid Y^*_{t_i}; \theta^*\right) \prod_{i=j}^{M-2} q\left(Y_{t_{i+1}} \mid Y_{t_i}, Y_{t_i}, \overline{Y}_{t_M}; \theta\right)}{\prod_{i=j}^{M-1} \pi\left(Y_{t_{i+1}} \mid Y_{t_i}; \theta\right) \prod_{i=j}^{M-2} q\left(Y^*_{t_{i+1}} \mid Y^*_{t_i}, Y^*_{t_i}, \overline{Y}_{t_M}; \theta^*\right)} \tag{1.4.7}$$

The standard Metropolis Hastings algorithm then accepts the new draw with probability $\alpha = \min(1, A)$, or else the draw is rejected and the last accepted draw is retained.

The algorithm

The algorithm for carrying out the Metropolis-Hastings scheme for sampling from the conditional posterior of parameters and latent data can be summarized as follows:

Initialization. Set $j = 0$. Initialize the augmented data points for each of the $s = 1, ..., MC$ iterations by linearly interpolating between observations for the first interval. Initialize the parameter set for all s by sampling from a prior density $\pi(\theta)$.

1. For each $s = 1, ..., MC$:

- Propose the parameters θ^* using (1.4.4)

- Propose the latent data Y^* for the interval (t_j, t_{j+m}) using (1.4.5) for each $i = j + 1, ..., M - 1$

- Accept the parameter and latent data proposal with probability $\alpha = \min(1, A)$ with A given by (1.4.7), and set $(Y_s, \theta_s) = (Y_s^*, \theta_s^*)$, or else set $(Y_s, \theta_s) = (Y_{s-1}, \theta_{s-1})$.

2. Set $j = j + m$ and go to (1).

The resulting draws of latent data and parameters form a Markov chain, whose stationary distribution after an initial burn-in period is given by (1.4.2). The number of imputed data points that are needed could be determined by running the sampler for low values of m and consequently increasing the discretization points until there is no significant change in the posterior parameter samples.

Convergence

In order to assess the accuracy of the parameter estimates obtained as ergodic averages of the form:

$$\widehat{\theta}_{MC} = \frac{1}{MC} \sum_{i=1}^{MC} (\theta^i)$$

we estimate their variance σ_θ^2 using the batch-mean approach (see Roberts, 1996; Tse et al., 2004). To this end, we run the MCMC scheme for $MC = m \times n$ iterations with m batches of n draws each. We compute the mean of each batch $k = 1, ..., m$ with:

$$\widehat{\theta}_k = \frac{1}{n} \sum_{i=(k-1)n+1}^{kn} \left(\theta^i\right)$$

Then we obtain an estimate of σ_θ^2 using:

$$\widehat{\sigma}_\theta^2 = \frac{n}{m-1} \sum_{k=1}^{m} \left(\widehat{\theta}_k - \widehat{\theta}_{MC}\right)^2 \tag{1.4.8}$$

and the Monte Carlo standard errors are obtained as $\sqrt{\frac{\widehat{\sigma}_\theta^2}{MC}}$.

As well, as a diagnostic tool that allows us to see how well the Markov chain mixes, we compute the simulation inefficiency factor (SIF) (see Kim et al., 1998), estimated as the variance of the ergodic averages σ_θ^2, divided by the variance of the sample mean from a hypothetical sampler that draws independent random variables from the parameter posterior. In order to compute the latter variance, we use the output of the MCMC runs, as in Tse et al. (2004), and obtain:

$$\overline{\sigma}_\theta^2 = \frac{1}{MC-1} \sum_{i=1}^{MC} \left(\theta^i - \widehat{\theta}_{MC}\right)^2$$

so that the SIF is estimated as:

$$SIF = \frac{\widehat{\sigma}_\theta^2}{\overline{\sigma}_\theta^2} \tag{1.4.9}$$

Model comparison through Bayes factors

In order to compare the estimated multivariate diffusion models of asset returns, we follow the traditional Bayesian approach that makes use of the marginal likelihood of each (potentially nonnested) model. The marginal likelihood is obtained by integrating the likelihood function of each model $\mathcal{M}_i$ with respect to the prior density:

$$p\left(Y \mid \mathcal{M}_i\right) = \int p\left(Y \mid \theta_i, \mathcal{M}_i\right) p\left(\theta_i \mid \mathcal{M}_i\right) d\theta_i$$

where θ_i are the parameters, corresponding to model $\mathcal{M}_i$. Then the Bayes factors for comparing model $\mathcal{M}_i$ against $\mathcal{M}_j$ are simply the ratio of the marginal likelihoods:

$$B_{ij} = \frac{p\left(Y \mid \mathcal{M}_i\right)}{p\left(Y \mid \mathcal{M}_j\right)} \tag{1.4.10}$$

We use the Laplace-Metropolis estimator of the marginal likelihood, proposed by Lewis and Raftery (1997) that relies on the posterior simulation output from the individual estimation of each model and approximates the integral using the Laplace method. Let us denote by θ_i^* the posterior parameter mean (or any other high density point of the parameter posterior). Then the logarithm of the marginal likelihood is estimated as:

$$\log\left(p\left(Y \mid \mathcal{M}_i\right)\right) \approx \frac{d}{2}\log\left(2\pi\right) + \frac{1}{2}\log\left(\left|H^*\right|\right) + \log\left(p\left(\theta_i^*\right)\right) + \log\left(p\left(Y \mid \theta_i^*, \mathcal{M}_i\right)\right)$$

where d is the dimension of the diffusion, $p\left(\theta_i^*\right)$ is the parameter prior under model $\mathcal{M}_i$, H^* is the inverse Hessian of $\log\left(p\left(\theta_i^*\right)p\left(Y \mid \theta_i^*, \mathcal{M}_i\right)\right)$, $\left|H^*\right|$ is its determinant, and $p\left(Y \mid \theta_i^*, \mathcal{M}_i\right)$ is the likelihood function, evaluated at θ^*.

Lewis and Raftery (1997) propose to estimate H^* by the sample covariance matrix of parameters from the MCMC output, so the only quantity that is left to be estimated is the likelihood function. The most straightforward estimator would be the one proposed by Pedersen (1995) that consists in averaging over the transition density implied by the Euler discretization. But as estimation was done by exploiting the information in both ends of each observation interval, we revert to a more efficient approach that is similar to the Metropolis-Hastings update used for latent data. Thus, the importance sampling estimator of the likelihood function has the following form[5]:

$$p\left(\overline{Y}_{t_M} \mid \overline{Y}_{t_j};\theta\right) = \int \frac{p\left(\overline{Y}_{t_M}, Y_{t_M} \mid \overline{Y}_{t_j};\theta\right)}{q\left(Y_{t_M} \mid \overline{Y}_{t_j}, \overline{Y}_{t_M};\theta\right)} q\left(Y_{t_M} \mid \overline{Y}_{t_j}, \overline{Y}_{t_M};\theta\right) dY_{t_M}$$

for an interval between two successive observations $\overline{Y}_{t_M}$ and $\overline{Y}_{t_j}$. Thus, the modified Brownian bridge proposal density that we used for the Metropolis-Hastings update could be used in this setup as the importance density q, which leads us to the following estimator of the likelihood function:

$$\widehat{p}\left(\overline{Y}_{t_M} \mid \overline{Y}_{t_j};\theta\right) = \frac{1}{M}\sum_{k=1}^{M} \frac{p\left(\overline{Y}_{t_M}, Y_{t_j}^k \mid \overline{Y}_{t_j};\theta\right)}{q\left(Y_{t_j}^k \mid \overline{Y}_{t_j}, \overline{Y}_{t_M};\theta\right)}$$

where $Y_{t_j}^k, k = 1, ..., M$ is a set of latent vectors between each pair of observations.

[5] See Elerian et al. (2001)

1.5 Estimation results

Although a joint estimation of each of the multivariate models is feasible, we propose to use a two step procedure, as this allows us to choose the appropriate marginal distribution for each data series. Such a two-step approach is commonly used in discrete-time copula models (Patton, 2004), as it allows to avoid copula model misspecifications. A misspecified univariate model would directly entail copula misspecifications, as the latter relies on the probability integral transform of each univariate series to model the dependence structure. A two step approach is possible in our continuous time setup as well, as a system of independent univariate diffusions is obtained under the product copula, assuming an identity correlation matrix for the diffusion term. Thus we first estimate each marginal diffusion separately, and then the obtained parameters are plugged in the multivariate model in order to estimate the dependence parameters pertaining to the chosen copula function, as well as the conditional correlation parameters for the multivariate diffusion.

1.5.1 Univariate diffusion

As the copula construction leaves us the freedom to choose any marginal distribution that should not be the same across all univariate series, we chose to estimate a NIG stationary distribution for all series, except the Mid caps, for which the more general GH construction appears to be appropriate (a NIG diffusion for the Mid caps is rejected on the basis of the uniform residuals obtained by the probability integral transform discussed in section 1.3.1; as well Bayes factors comparison between a NIG and a GH diffusion point towards the latter for Mid caps). The chosen marginal distribution is thus given by (1.3.11) for Small and Large caps, with the Bessel function parameter λ set to -0.5, while the marginal distribution for Mid caps is given by (1.3.8) with a free λ parameter. Table 1.5.1 summarizes the estimation results for the parameters specific to each univariate series.

It is interesting to note that the parameter κ for all three series of data is significantly different from 0 or 1, which would correspond to either a constant volatility diffusion for the state variables ($\kappa = 0$) or a zero drift diffusion ($\kappa = 1$). Thus we retain that more general specification for the subsequent multivariate diffusion modeling. As well, correct modeling of both the drift and the diffusion term for the state variables would be crucial for filtering out the market price of risk, which in turn could have a significant impact on

Table 1.5.1: Parameter estimates for the univariate series

The table summarizes the posterior parameter estimates from the MCMC output. Monte Carlo standard errors are reported in parenthesis (multiplied by a factor of 1000) (obtained using the batch-mean approach). SIF refers to the simulation inefficiency factor for each parameter (its integrated autocorrelation time).

	Smallcap	Midcap	Largecap
α	3.0502	18.7839	10.6904
(MC s.e.)	(0.1616)	(0.5220)	(0.2193)
(SIF)	(0.0938)	(0.6694)	(0.6912)
β	-0.5911	0.4476	-1.5737
(MC s.e.)	(0.6329)	(2.9453)	(1.5404)
(SIF)	(0.1104)	(1.5392)	(1.7637)
δ^2	0.0301	0.0721	0.0410
(MC s.e.)	(0.0024)	(0.0011)	(0.0031)
(SIF)	(0.1219)	(1.0535)	(1.8122)
μ	6.7059	6.3101	6.5360
(MC s.e.)	(0.0249)	(0.0129)	(0.0102)
(SIF)	(0.1038)	(0.5407)	(0.4991)
σ^2	0.0406	0.0400	0.0082
(MC s.e.)	(0.0022)	(0.0030)	(0.0006)
(SIF)	(0.1142)	(1.4686)	(1.2930)
κ	0.6490	0.4670	0.5102
(MC s.e.)	(0.0373)	(0.0235)	(0.0850)
(SIF)	(0.0955)	(1.4322)	(1.7551)
λ	-0.5	-1.4295	-0.5
(MC s.e.)	-	(0.0519)	-
(SIF)	-	(1.1704)	-

portfolio decisions based on this model for stock prices.

Further analysis of the MCMC output is offered on Figure 1.5.1, where we present the sample paths of the estimated parameter for the Small cap data series[6], as well as autocorrelation plots for a lag up to 100, and kernel density estimate of the posterior parameter output. We do not have any significant autocorrelation for any of the parameters, which is a consistent result with Golightly and Wilkinson (2006a), who show a significant reduction in sample autocorrelations of the Simulation Filter as compared to the Gibbs sampler.

In order to examine whether the proposed multivariate diffusion replicates certain dynamic properties of the data, we simulate a very long series (of length 100 000) from the univariate NIG diffusion model for log prices X_{it} (1.3.18) and parameters corresponding to the Large cap series in Table 1.5.1, and examine the implied properties of their increments[7]. A stylized fact of asset returns is the persistence in autocorrelation in squared returns in contrast to the lack of autocorrelation in the original return series (except for possibly the first lag). If we examine the autocorrelation patterns in the data and the long simulated series, we find that this property is actually captured by the model, as displayed on Figure 1.5.2.

This finding is not surprising, if we consider the fact that the Euler discretization of a univariate diffusion of the generalized hyperbolic family can be considered as a special case of a nonlinear ARCH model (Tse et al., 2004), and thus it can be expected to exhibit volatility clustering and long memory properties. The same behavior is preserved in the multivariate specification as well.

Another important aspect of our analysis is the fit of each of the univariate diffusions to the empirical distribution of the data, as they will provide the inputs for the probability integral transform in the copula construction. Figure 1.5.3 illustrates the close replication of the stationary distribution by the considered marginal processes.

Note that the construction of a stationary diffusion leaves us with the freedom to choose either the volatility or the drift specification. Models that were proposed in literature treat either one or the other as a constant, that significantly facilitates estimation but leaves open the question as to whether such a simplified model would reproduce the dynamic

[6] Results for the Mid and Large caps series are qualitatively the same and we do not report them for brevity.

[7] Similar results are obtained for any of the univariate data series considered.

Figure 1.5.1: MCMC estimation output: Small caps, NIG diffusion

The figure displays the output from the MCMC estimation of the NIG diffusion for Small caps. The top figure represents the sample paths of the parameters from 100000 replications; the bottom left – the autocorrelation functions for the sample paths of each parameter, and the bottom right figure – the kernel density estimate of the parameter posterior distributions.

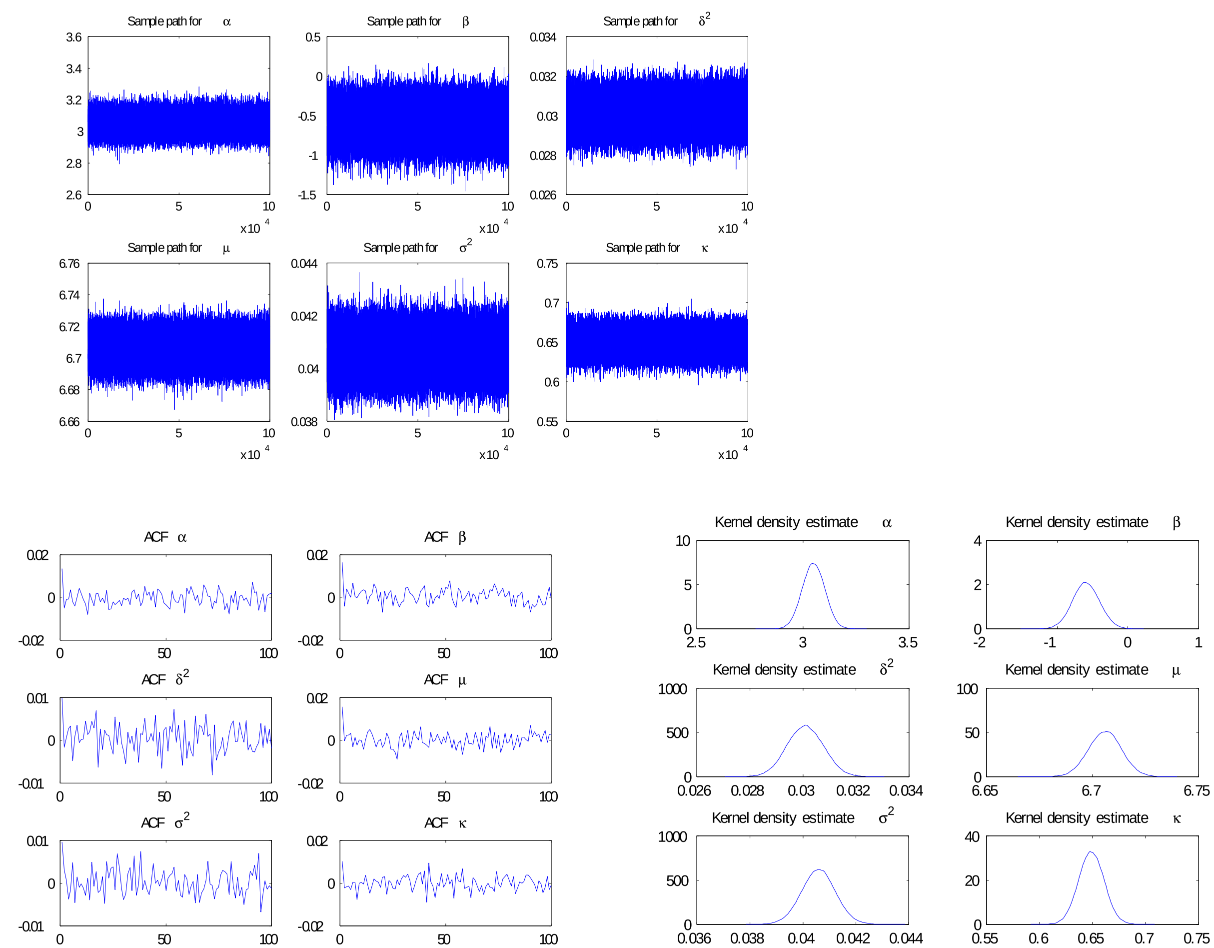


Figure 1.5.2: Autocorrelation plots for simulated and actual return series

Autocorrelation functions for the observed and simulated return series (Large cap). The top figures show autocorrelation in returns, the bottom – autocorrelation in squared returns. The simulated return series is obtained from the univariate diffusion with parameters for Large cap with length 100 000 using the Euler discretization.

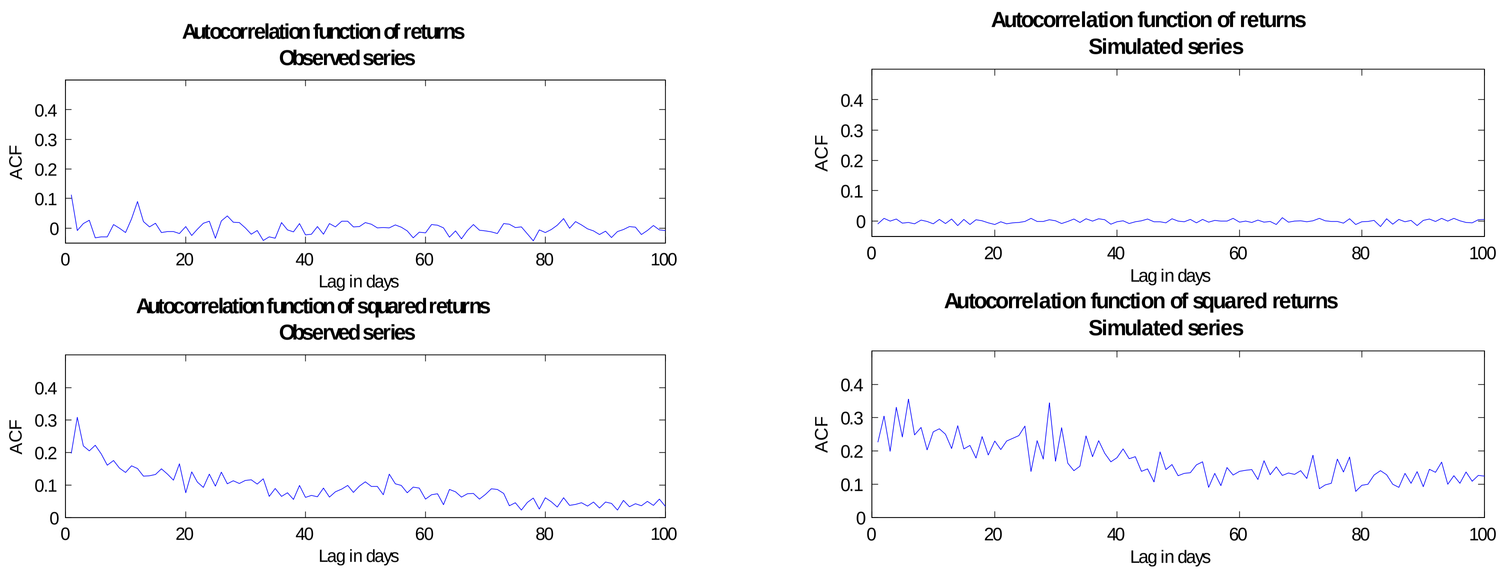


Figure 1.5.3: Fitting the marginal distributions

Empirical cumulative distribution functions plotted against the theoretical distribution of each of the three univariate models, with parameters given in Table 1.5.1.

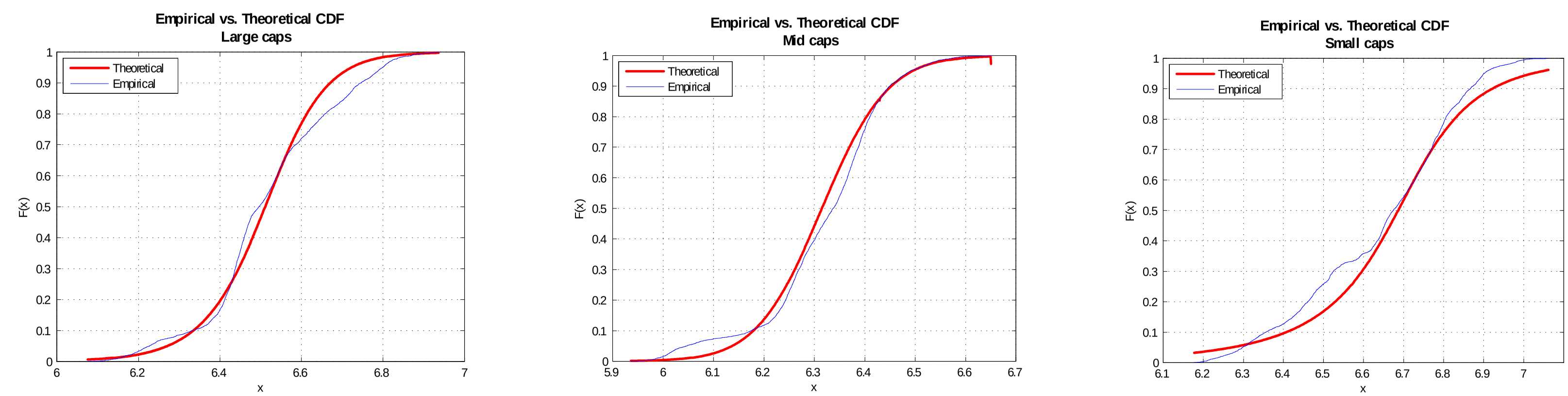
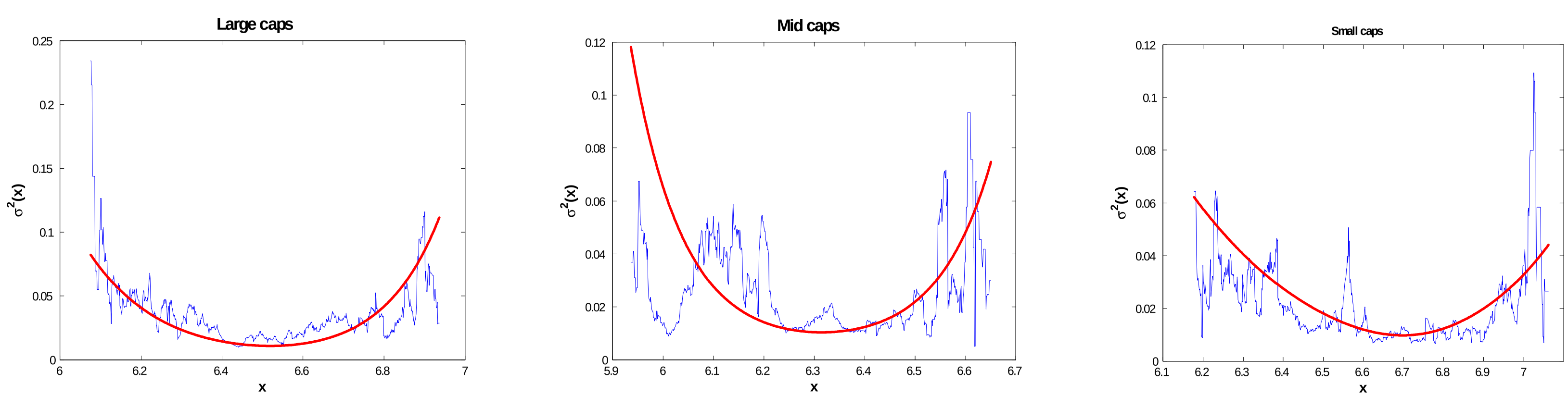


Figure 1.5.4: Fitting the volatility pattern

The parametric specification of the squared diffusion coefficients $\sigma_i \left[\widetilde{f}\left(X_{it}\right)\right]^{-\frac{1}{2}\kappa_i}, i = 1, ..., 3$ (parameters from Table 1.5.1) plotted against the nonparametric estimator of the diffusion term $V_n\left(x\right)$ for a choice of bandwidth parameter of $h = 0.01$ for each of the three univariate series.



properties of the data as well. As we estimate a general diffusion with nonconstant drift or volatility term through the parameters κ_i in (1.3.5), it would be of interest to examine how closely this chosen specification can account for the variability in the data. Following Bibby and Sorensen (1997), we contrast the parametric specification of the volatility term $\sigma_i \left[\widetilde{f}\left(X_{it}\right)\right]^{-\frac{1}{2}\kappa_i}$ against a nonparametric estimator of the squared diffusion coefficient, based on quadratic variation, as proposed in Florens-Zmirou (1993):

$$V_n\left(x\right) = \frac{\sum_{j=1}^{n} 1_{\left|X_{i,t_j}-x\right|<h}\left(X_{i,t_j} - X_{i,t_{j-1}}\right)^2}{\sum_{j=1}^{n} 1_{\left|X_{i,t_j}-x\right|<h}\left(t_j - t_{j-1}\right)}$$

with a bandwidth parameter h. Figure 1.5.4 displays the fit of the volatility specification for each of the univariate models. The U-shaped parametric volatility form (1.3.5) matches closely the non-parametric estimator. A constant volatility specification (achieved by setting κ_i to zero) would thus underestimate volatility in the cases when returns are in either tail of the distribution and fail to reproduce the empirical stylized fact that returns are highly volatile in extreme market downturns.

A check of the fit of the univariate models is done via the dynamic probability integral transform that uses the transitional probabilities of the discretized version of the diffusion between two consecutive observations with the Euler discretization scheme, as discussed in Section 1.3.1. For the model to be well specified, the series of uniform residuals should be $i.i.d.U(0,1)$. The residuals could then be analyzed using quantile plots, as illustrated on Figure 1.5.5. A formal test could be conducted using the statistic $stat = -2\sum_{i=1}^{n}\log U_i \sim \chi^2_{2n}$, following Bibby and Sorensen (1997). For 3997 observations, the test statistic for the Small caps is 7.8677e+003, for the Mid cap it is 7.8797e+003, and for the Large cap it is 8.1278e+003, none of which gives reasons to reject the correct model specification.

Finally, we proceed to a simulation study in order to validate the proposed MCMC estimation scheme. We simulate a sample from the univariate NIG diffusion with parameters corresponding to the estimates for the Large cap series (Table 1.5.1). The simulated series corresponds to 5 years of data and is simulated using Euler discretization. We then run the Simulation Filter for 100 000 iterations at each time step and with $m = 5$ or 15 latent data points between observations. Figure 1.5.6 summarizes the results.

The upper panel of Figure 1.5.6 plots the evolution of the sequential parameter estimates

Figure 1.5.5: A formal check of the univariate diffusion models

Quantile plots and autocorrelation plots of the uniform residuals for each of the univariate diffusion models.

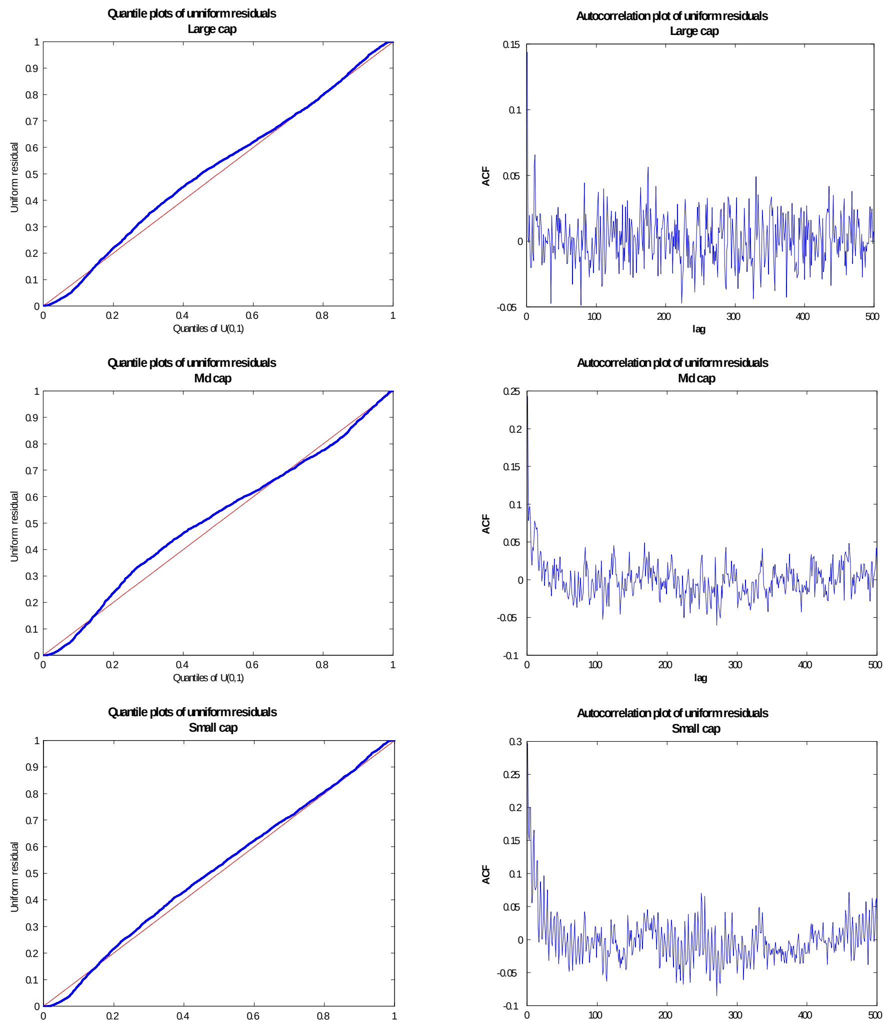
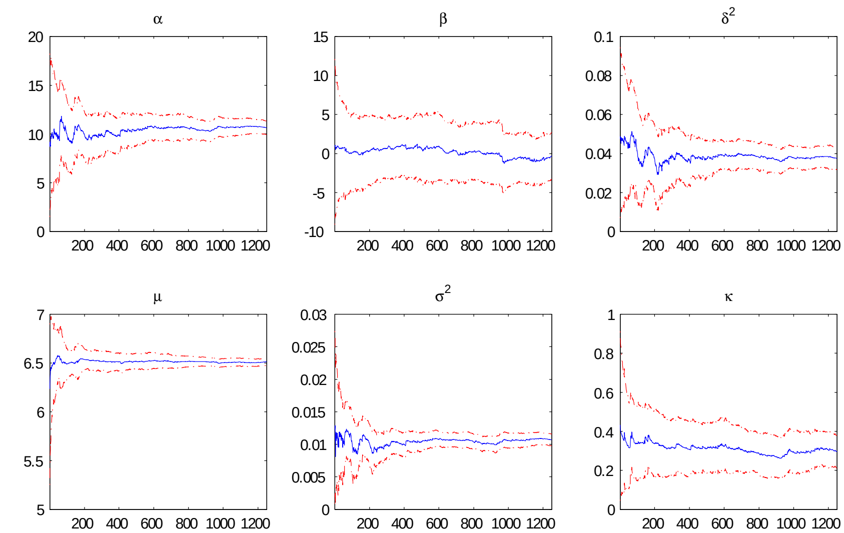
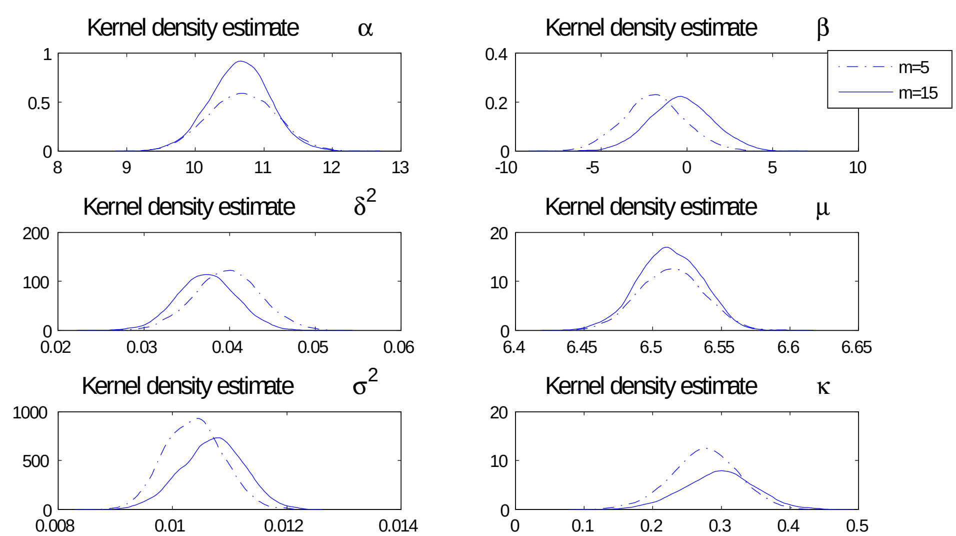


Figure 1.5.6: MCMC estimation output: simulated series

The figure displays the output from the MCMC estimation of the NIG diffusion for a simulated series with parameters taken from the estimates of the NIG diffusion for Large caps. The top figure represents the sequential parameter plots from 100000 replications with $m = 15$ latent data points between observations; the line in the middle is the smoothed posterior mean, and the two dashed lines represent the 95th and the 5th quantiles from the posterior parameter distribution. The bottom figure plots the kernel density estimate of the parameter posterior distributions for $m = 5$ and $m = 15$.





across time: the smoothed posterior mean, as well as the 5th and the 95th quantile of the parameter posterior. The lower panel shows the kernel density estimates of the parameter posteriors after running the Simulation Filter through the whole period for the two discretization cases of $m = 5$ or 15. The estimated parameters α, β, δ, and μ, corresponding to the stationary NIG distribution, are fairly close to the true ones for both choices of number of latent data points m.

1.5.2 The importance of modeling asymmetric tail dependence: a bivariate diffusion example

Having obtained estimates of the univariate marginal distributions for each data series, we now turn to estimating the model parameters that pertain to the dependence structure. The bivariate quantile plots for all three couples of data on Figure 1.2.1 have shown a substantial degree of quantile 'near' tail dependence that does not fade away as we approach the tails of the distribution, especially the left one. That is, in periods of extreme market downturns stocks continue being dependent - a feature that could possibly wipe out any diversification benefits of an all-stock portfolio. As well, the non-parametric test of exceedence correlations symmetry with exceedence levels chosen close to the tails, whose results are shown in Table 1.2.1, rejects symmetry for all couples of data, except the Large cap - Mid cap couple, for which in both tails exceedence correlations are high. In what follows, we verify whether a multivariate copula diffusion model could reproduce these properties of the data.

A good candidate for the purpose of modeling an asymmetric tail behavior is the bivariate Symmetrized Joe-Clayton copula, discussed in previous sections. It has two parameters, each one directly linked to the upper or lower tail dependence coefficient. So before we estimate a bivariate diffusion model based on this copula function, let us first look at the levels of tail dependence that could be achieved through it. In order to do so, we need to obtain the levels of its parameters, implied by the data, so we first estimate the copula parameters from the unconditional distribution of each couple of the state variables X, pertaining to each of the CRSP size indices. We apply the Canonical Maximum Likelihood estimation method which consists in first transforming the data into uniform variables using the empirical distribution, that is without imposing any parametric restrictions on the univariate marginals, and then estimating the copula parameters θ with MLE:

$$\widehat{\theta} = \arg\max_{\theta} \sum_{t=1}^{T} \ln c\left(\widehat{F}_i\left(x_i\right), \widehat{F}_j\left(x_j\right); \theta\right), \quad i,j = 1,2$$

where $\widehat{F}_i\left(x_i\right)$ is the empirical CDF of x_i, and $c\left(\cdot\right)$ is the chosen parametric copula function. We estimate the copula parameters for two choices of copulas - the tail independent Gaussian and the asymmetric tail dependent SJC copula. Then for each dependence function we trace quantile plots (Figure 1.5.7), where the levels of quantile dependence are obtained using (1.2.10), which are then contrasted against the quantile plots for the data itself.

The coefficients of upper and lower tail dependence for the Large cap - Mid cap couple are both high, which corresponds to the symmetric tail behavior in terms of exceedence correlations that we reported in Table 1.2.1. However, the upper tail coefficients for the other two couples of data are low, especially for the Large cap - Small cap couple, where $\tau^U = 0$, while the lower tail dependence coefficients are significantly higher, confirming the evidence of asymmetric tail behavior. Quantile dependence plots, implied by the so estimated coefficients, confirm the finding of higher dependence as we go further in the left tail. The quantile dependence plots for the SJC copula are closer to the data, while those corresponding to a Gaussian copula deviate from it, especially in the left tail, where Gaussian dependence fades away for decreasing quantile levels, while SJC copula-implied dependence maintains higher level, closer to the data.

We now turn to the estimation of a bivariate diffusion whose stationary distribution has a dependence structure governed by the asymmetric tail SJC copula. Using the Simulation MCMC filter, we estimate the bivariate model for all three couples of data, while keeping fixed the univariate marginal distribution parameters at their estimated values from the previous section. Results are reported in Table 1.5.2.

Note that the estimates of the upper and lower tail dependence parameters for the diffusion models are fairly close to the values obtained for the unconditional distribution, estimated using the Canonical Maximum Likelihood with uniform variates from the empirical distribution (Figure 1.5.7). Again upper tail dependence coefficients are lower than their corresponding lower tail dependence counterparts for all couples of data. Upper tail dependence is still relatively high for the Large cap - Mid cap couple.

The obtained parameter estimates are then used to simulate long series from each of the

Figure 1.5.7: Quantile dependence plots for varying copula specifications

Plots of quantile dependence for all three couples of de-trended log-prices of the three CRSP indices formed on the basis of size deciles for the period 1986-2005 (small-cap (deciles 1-3), mid-cap (deciles 4-7), and large-cap (deciles 8-10)). Quantile dependence plots of the data are contrasted by quantile dependence that comes from two alternative parametric copula specifications: no-tail dependent Gaussian copula and asymmetric dependent Symmetrized Joe-Clayton (SJC) copula. Fitted parameters for the Gaussian copula are $\rho = 0.6969$ (Large cap vs. Mid cap), $\rho = 0.2754$ (Large cap vs. Small cap), and $\rho = 0.7593$ (Small cap vs. Mid cap). Fitted parameters (upper (τ^U) and lower (τ^L) tail dependence) for the SJC copula are $\tau^U = 0.4473, \tau^L = 0.6372$ (Large cap vs. Mid cap), $\tau^U = 0, \tau^L = 0.3230$ (Large cap vs. Small cap), and $\tau^U = 0.1811, \tau^L = 0.7109$ (Small cap vs. Mid cap).

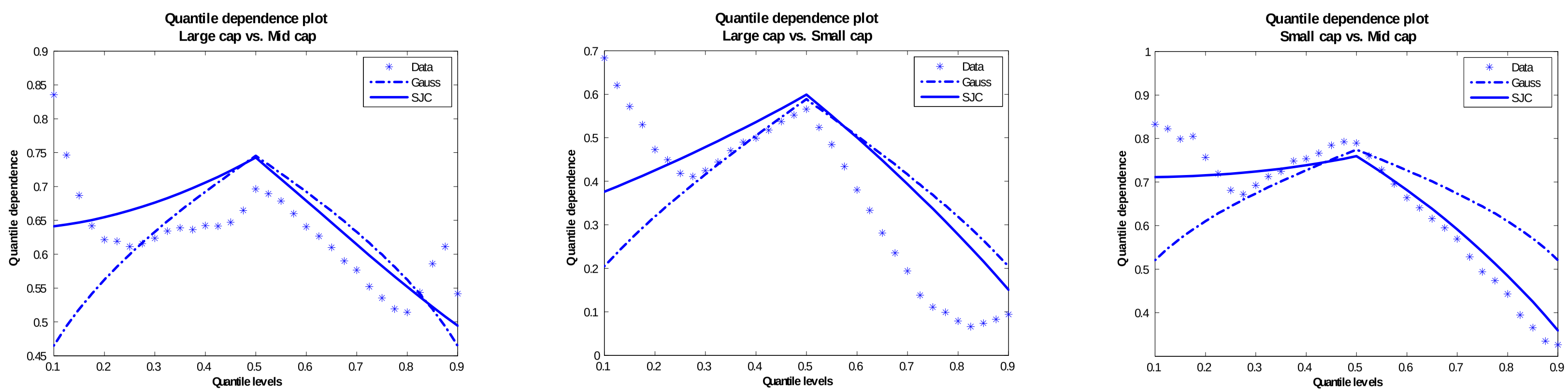


Table 1.5.2: Parameter estimates for a bivariate Symmetrised Joe-Clayton copula diffusion

The table summarizes the posterior parameter estimates from the MCMC output. Parameters τ^U and τ^L refer to the upper and lower tail dependence of the bivariate Symmetrised Joe-Clayton copula, while parameter ρ is the conditional correlation parameter of the bivariate diffusion. The rest of the parameters, pertaining to the marginals, are not estimated and are kept fixed at their corresponding values from Table 1.5.1. Monte Carlo standard errors are reported in parenthesis (multiplied by a factor of 1000) (obtained using the batch-mean approach). SIF refers to the simulation inefficiency factor for each parameter (its integrated autocorrelation time).

	Large cap - Mid cap	Large cap - Small cap	Small cap - Mid cap
τ^U	0.4171	0.0484	0.1835
(MC s.e.)	(0.1568)	(0.1209)	(0.1934)
(SIF)	(0.2890)	(1.0160)	(0.5710)
τ^L	0.6724	0.2700	0.6602
(MC s.e.)	(0.1585)	(0.2359)	(0.0608)
(SIF)	(1.0040)	(1.0279)	(1.1880)
ρ	0.5968	0.6514	0.6682
(MC s.e.)	(0.0107)	(0.0049)	(0.0050)
(SIF)	(1.4428)	(0.9354)	(2.0860)

three SJC copula diffusions. Further, we calculate the levels of quantile dependence for each bivariate series using (1.2.10). From each bundle of simulated series and their corresponding levels of quantile dependence, we then determine the obtainable degrees of dependence for each quantile level in bands between the 5th and the 95th percentile. Thus, for each quantile level we show the degrees of quantile dependence that can be reached in 90% of the cases with a SJC copula diffusion. Results are presented on Figure 1.5.8.

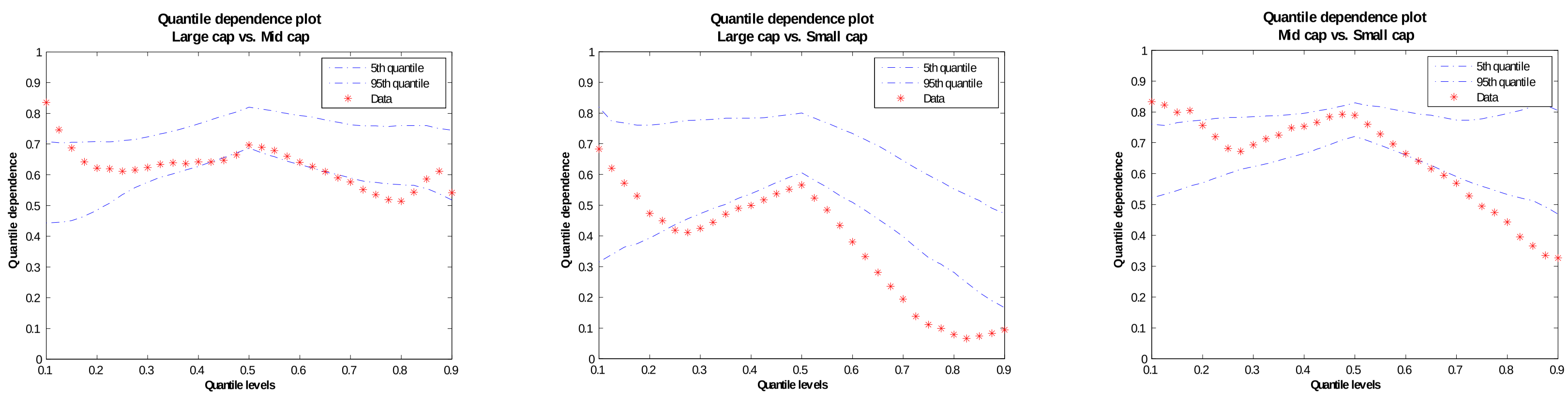
For the case of the Large cap - Mid cap couple, quantile dependence implied by the data generally falls within the bounds reachable under the estimated parameters for the SJC copula, with the exception of the extreme left tail, which would require an even higher left tail dependence parameter in order to accommodate the dependence found in the data. For the other two couples, the parameters for the SJC diffusion can reasonably well replicate the quantile dependence for the left tail.

1.5.3 A generalization to higher dimensions

Even though the SJC copula is intuitively appealing as its parameters are directly linked to the coefficients of upper and lower tail dependence, it could not be easily generalized to a

Figure 1.5.8: Quantile dependence plots for simulated bi-variate SJC diffusions

Plots of quantile dependence for all three couples of de-trended log-prices of the three CRSP indices formed on the basis of size deciles for the period 1986-2005 (small-cap (deciles 1-3), mid-cap (deciles 4-7), and large-cap (deciles 8-10)). Quantile dependence plots of the data are contrasted by quantile dependence that comes from simulated series of bivariate Symmetrized Joe-Clayton (SJC) copula diffusions with parameters taken from Table 3. 950 simulated series of length 50 000, using the Euler discretization, are used to determine the quantile dependence levels, obtainable between the 5th and the 95th percentile bands for a SJC diffusion.



higher dimension. Copula functions that can be extended in a straightforward manner to dimensions higher than 2 are the Elliptic copulas, whose form is given in (1.3.22) for the n-variate Gaussian copula and in (1.3.23) for the n-variate Student's t copula. However, they imply either no tail dependence or symmetric tail dependence (governed by the degrees of freedom parameter), while we have seen that data generally asks for a copula function that can accommodate asymmetric dependence between extreme realizations.

That is why we turn to copula functions in the Archimedean family that allow an extension to higher dimensions without imposing symmetry. One such candidate is the Gumbel copula and its survival counterpart, that model either upper or lower tail dependence. The most straightforward way that these copulas be generalized to n dimensions is to allow for the same parameter α to govern the dependence structure for all n random variables, as in (1.3.25) for the Gumbel copula and in (1.3.26) for the survival Gumbel.

However, we have seen that for the three couples of CRSP indices that we have considered the degrees of upper or lower tail dependence vary substantially between the couples (e.g. the Large cap - Mid cap couple are both upper and lower tail dependent, while the lower tail dependence for the Large cap - Small cap couple is close to zero). Thus, imposing the same dependence parameter across all variables could be seriously misleading. A remedy to this problem is the nested copula construction for the Archimedean family that we have discussed in the previous sections. It nests the Archimedean generator functions with different parameters and can thus impose different degrees of dependence for the random variables (the highest dependence being achievable for the most deeply nested couple). The three-variate nested Archimedean copula, expressed in terms of the copula generator and its inverse is given by (1.3.27). Thus, we may pick up the size decile couple that has the highest dependence and model it as the most deeply nested couple. The generating function for this couple will then be φ_1 with a dependence parameter α_1. Thus we obtain the first copula, $C(u_1, u_2; \alpha_1) = \varphi_1^{-1}(\varphi_1(u_1) + \varphi_1(u_2))$. We then couple it with the third remaining data series using a second generating function φ_2 with a dependence parameter α_2 that implies lower dependence than α_1 and obtain the nested copula $C(u_1, u_2, u_3; \alpha_1, \alpha_2) = \varphi_2^{-1}(\varphi_2(C(u_1, u_2; \alpha_1)) + \varphi_2(u_3))$. This subsequent nesting of generating functions requires that they are quite parsimonious in nature in order to keep the resulting copula function tractable, and the Gumbel copula that we use is a good candidate for that. In our application we use either of the size decile couples as the most deeply nested one, although the most fitted couple for that is the Large cap - Mid cap one, as it implies high dependence in both tails.

When using the extreme value Gumbel copula, there is also some concern that we are implicitly imposing asymptotic dependence between extreme realizations of the random variables (see Poon et al., 2004). In order to allow for asymptotic independence, we consider instead the mixture copula function as defined in (1.3.28) which combines the two extreme value Gumbel copulas with the tail independent Gaussian one. If the estimate of the weight for the Gaussian copula goes close to 1, then our series is asymptotically independent. Otherwise there is some degree of dependence in either of the tails, depending on the weighting of the Gumbel copula or its survival counterpart. As well, in order to allow for richer parametrization of the dependence structure, we consider a mixture copula of the two extreme value ones with the Student's t as in (1.3.29). In this case we always have asymptotic dependence, unless the degrees of freedom parameter goes to infinity, and the importance of dependence in each tail is again determined by the weight of the corresponding extreme value copulas.

Estimation results for the multivariate diffusion with a Gaussian dependence structure is given in the first column of Table 1.5.3. Then we add the three alternative cases of a diffusion with tail dependence as implied by the nested mixture copula ((1.3.28) with nested Gumbel and Survival Gumbel), and finally we consider the most parsimonious specification where there is only one parameter that determines upper tail dependence, and one for lower tail dependence ((1.3.28) with non-nested Gumbel and Survival Gumbel copulas).

First, note that for both the Gumbel and the Survival Gumbel copulas, the parameters corresponding to the most deeply nested couple (α_1^G and $\overline{\alpha}_1^G$ respectively) have lower values than the corresponding parameters for the second generating function (α_2^G and $\overline{\alpha}_2^G$), as they have to respect the condition that assures that the obtained nested Gumbel function is indeed a copula.

The relatively high and symmetric lower and upper tail dependence coefficients for the Mid-Large cap couple that we found earlier are indeed reflected in the estimation results for the nested Gaussian-Gumbel-Survival Gumbel diffusion for the case where it is most deeply nested in the copula specification (Table 1.5.3, second column). The two parameters

Table 1.5.3: Parameter estimates for the dependences structure (tri-variate diffusion, Gaussian underlying)

Estimation results for the trivariate diffusions using the Gaussian copula, the nested Gaussian-Gumbel-Survival Gumbel (Ga-G-SG) mixture copula (the most deeply nested couple is given in parenthesis), the nonnested Gaussian-Gumbel-Survival Gumbel (Ga-G-SG) mixture copula. Monte Carlo standard errors (multiplied by a factor of 1000), and Simulation Inefficiency Factors (SIF) are given in parenthesis. The first three parameters (R_{12}, R_{13}, R_{23}) correspond to the off-diagonal entries of the correlation matrix R_{Ga} for the Gaussian copula. The parameters α_1^G and α_2^G are the dependence parameters for the nested Gumbel copula, and the parameters $\overline{\alpha}_1^G$ and $\overline{\alpha}_2^G$ are the dependence parameters for the nested Survival Gumbel copula. For the nonnested case, the relevant parameters are α_1^G for the Gumbel copula and $\overline{\alpha}_1^G$ for the Survival Gumbel copula. ω^G and ϖ^G are the corresponding weights for the Gumbel and the survival Gumbel copula for the mixture model. The parameters ρ_{12}, ρ_{13} , and ρ_{23} are the off-diagonal entries of the correlation matrix in the diffusion specification. Results are obtained for 50000 Monte Carlo replications with a thinning factor of 5 with 10 latent data points simulated between each pair of observations.

	Gaussian	Ga-G-SG (Large-Mid cap)	Ga-G-SG (Large-Small cap)	Ga-G-SG (Small-Mid cap)	Ga-G-SG (nonnested)
R_{12}	0.5671	0.5347	0.4636	0.6634	0.5758
MC s.e.	0.3701	0.3326	0.7224	0.6114	0.3537
SIF	0.8621	1.0437	2.3408	0.6891	0.9540
R_{13}	0.2723	0.5179	0.7443	0.3907	0.2571
MC s.e.	0.7875	0.4191	0.6868	0.4915	0.5131
SIF	0.7359	0.7188	2.7202	0.8441	0.7251
R_{23}	0.5207	0.4152	0.6110	0.3085	0.4698
MC s.e.	0.4399	0.3302	0.6260	0.5521	1.3536
SIF	0.9162	1.6992	0.8855	1.2236	1.5260
α_1^G	-	0.2972	0.3358	0.3318	0.4494
MC s.e.	-	0.3546	0.8711	0.3463	0.3541
SIF	-	0.5754	1.5005	1.5945	1.2328
α_2^G	-	0.6335	0.6238	0.7235	-
MC s.e.	-	0.1928	0.4072	0.7235	-
SIF	-	0.9156	2.6644	1.0750	-

Table 1.5.3: Parameter estimates for the dependences structure (tri-variate diffusion, Gaussian underlying) (cont.)

	Gaussian	Ga-G-SG (Large-Mid cap)	Ga-G-SG (Large-Small cap)	Ga-G-SG (Small-Mid cap)	Ga-G-SG (nonnested)
$\overline{\alpha}_1^G$	-	0.3618	0.1993	0.2613	0.4354
MC s.e.	-	0.1998	0.3006	0.5387	1.0229
SIF	-	0.2375	1.8385	1.3787	1.6558
$\overline{\alpha}_2^G$	-	0.6544	0.6415	0.6107	-
MC s.e.	-	0.4667	0.3408	0.7371	-
SIF	-	0.8040	1.0323	1.1141	-
ω^G	-	0.3321	0.2107	0.2431	0.3832
MC s.e.	-	1.0111	0.5641	0.3543	0.7265
SIF	-	2.0983	1.3229	0.8943	1.0348
$\overline{\omega}^G$	-	0.2853	0.2752	0.2531	0.2324
MC s.e.	-	0.3789	0.2519	0.6596	0.3619
SIF	-	1.4739	0.7950	1.2326	2.1457
ρ_{12}	0.7894	0.7917	0.7795	0.6935	0.8287
MC s.e.	0.0195	0.0086	0.0080	0.0095	0.0104
SIF	1.2371	0.2271	0.7370	0.6121	1.2730
ρ_{13}	0.5078	0.5089	0.5185	0.4685	0.5499
MC s.e.	0.0189	0.0229	0.0291	0.0236	0.0105
SIF	0.8625	0.9588	1.5205	0.5720	0.6771
ρ_{23}	0.7162	0.7158	0.6760	0.5676	0.7366
MC s.e.	0.0209	0.0067	0.0102	0.0159	0.0137
SIF	0.8581	0.7418	1.1571	0.9287	1.1969

that determine upper and lower tail dependence for this couple, α_1^G and $\overline{\alpha}_1^G$ respectively, are almost equal, pointing to tail symmetry. As well, given the fact that for the Gumbel copula the tail dependence coefficient can be determined using $\tau_G^U = 2 - 2^{\alpha}$, while for the lower tail dependence implied by the Survival Gumbel we have that $\tau_G^L = 2 - 2^{\overline{\alpha}}$, then for this particular couple we have $\tau_G^U = 0.7712$ and $\tau_G^L = 0.7150$. These values are higher than what we obtained under the alternative bivariate diffusion with dependence modeled following a SJC copula, but it is still not surprising as now in the mixture specification these extreme value copulas are weighted with the tail independent Gaussian copula, so the resulting tail dependence should be lower.

Further, for the two alternative cases for which Large-Small or Mid-Small are the most deeply nested couples the lower tail dependence parameter $\overline{\alpha}_1^G$ is lower than the upper tail dependence parameter α_1^G, indicating higher dependence in the left tail, again confirming the previously found evidence.

When we consider the multivariate diffusion with extreme dependence modeled with the non-nested version of the Gumbel and Survival Gumbel copulas, we find almost symmetric tail dependence (the values for the parameters α_1^G and $\overline{\alpha}_1^G$ imply tail dependence coefficients of $\tau_G^U = 0.6345$ and $\tau_G^L = 0.6477$). As in this case there is only one parameter governing dependence in either the left or the right tail across all data series, extreme dependence for some couples may be over/underestimated.

We then repeat the same estimation experiment, but with a Student's t copula used to model the dependence structure of the multivariate diffusion. Results for it are reported in the first column of Table 1.5.4. We proceed as before by adding the two extreme value copulas in their nested specification as in (1.3.29). The second column of Table 1.5.4 contains the results for this case when the Mid-Large cap couple is taken to be the most deeply nested one. Finally we consider a non-nested version, but it has a different form from that of the mixture diffusion with a Gaussian underlying. As the Student's t copula can model symmetric upper and lower tail dependence, which is different for all alternative couples considered, as it is determined by the degrees of freedom parameter γ, as well as the correlation matrix, then, in order to add tail asymmetry, we need only to consider one of the tails and add an extreme value copula that accounts for dependence in it. We chose to model separately the left tail and thus we add to the t-copula a Survival Gumbel that has lower tail dependence. Results for this specification are reported in the third column.

As with the case when we had a Gaussian underlying copula in the mixture model, here again the parameters, driving upper and lower tail dependence for the most deeply nested couple (the Large-Mid cap one) are very close, and imply tail coefficients of $\tau_G^U = 0.7870$ and $\tau_G^L = 0.7917$. However, unlike the Gaussian case, the Gumbel copula claims almost half of the weight in the mixture specification, so it has the major role in determining upper tail dependence. For the nonnested version of the t-mixture copula, adding only the lower tail dependent Survival Gumbel copula increases its weight, but leaves the dependence parameter almost unchanged.

1.5.4 Model selection through Bayes factors

The multivariate diffusion models considered above imply different dependence structures through their stationary distributions. Bayes factors provide us with a guideline of how to select a model among the alternatives. So far we have seen that the mixture model with either a Gaussian or a t-copula, combined with the nested version of the extreme value Archimedean copulas provide the richest specification in terms of tail dependence modeling. In what follows we will verify whether either one of these two models will be indeed selected on the basis of the Bayes factor criterion.

We compute the log of Bayes factors, following (1.4.10) as $\log\left(p\left(Y \mid \mathcal{M}_b\right)\right) - \log\left(p\left(Y \mid \mathcal{M}_j\right)\right)$. As a benchmark model ($\mathcal{M}_b$) we take either the Gaussian or the Student's t - extreme value nested mixture copula diffusion (with the Large-Mid cap couple being the most deeply nested one).The alternatives considered ($\mathcal{M}_j$) are the tail independent Gaussian diffusion, the symmetric tail dependent t-copula diffusion, or any of the non-nested specifications considered. Results are provided in Table 1.5.5.

The Bayes factor selection criterion suggests that each of the benchmark models will be preferred to the alternatives. Results point even more strongly in favour of the extreme value nested mixture copulas when the nonnested models are taken as alternatives. This suggests that the highly parsimonious dependence structure, implied by the nonnested copulas, is detrimental to the models, at least for the purposes of selection through Bayes factors.

Further, when we compare the two benchmark models, Bayes factors point in favour of the Student's t mixture copula, with a value for the log of the Bayes factor of 9.06 when the

Table 1.5.4: Parameter estimates for the dependences structure (tri-variate diffusion, Student's t underlying)

Estimation results for the trivariate diffusions using the Student's t copula, the Student's t – nonnested Survival Gumbel mixture copula, and the Student's t – nested Gumbel - Survival Gumbel mixture copula (the most deeply nested couple is given in parenthesis). Monte Carlo standard errors (multiplied by a factor of 1000), and Simulation Inefficiency Factors (SIF) are given in parenthesis. The first three parameters (R_{12}, R_{13}, R_{23}) correspond to the off-diagonal entries of the correlation matrix R_T for the Student's t copula. The parameters α_1^G and α_2^G are the dependence parameters for the nested Gumbel copula, and the parameters $\overline{\alpha}_1^G$ and $\overline{\alpha}_2^G$ are the dependence parameters for the nested Survival Gumbel copula. For the nonnested case, the relevant parameters are α_1^G for the Gumbel copula and $\bar{\alpha}_1^G$ for the Survival Gumbel copula. ω^G and $\bar{\omega}^G$ are the corresponding weights for the Gumbel and the survival Gumbel copula for the mixture model. ν is the degrees of freedom parameter for the Student's t copula. The parameters ρ_{12}, ρ_{13} , and ρ_{23} are the off-diagonal entries of the correlation matrix in the diffusion specification. Results are obtained for 50000 Monte Carlo replications with a thinning factor of 5 with 10 latent data points simulated between each pair of observations.

	t	t-G-SG (Large - Mid cap)	t-SG (nonnested)
R_{12}	0.4408	0.2574	0.5266
MC s.e.	0.5433	1.4015	0.6040
SIF	1.3619	0.7629	1.3392
R_{13}	0.5273	0.2362	0.4154
MC s.e.	0.6911	0.9873	0.6353
SIF	0.9564	1.0469	0.8209
R_{23}	0.3334	0.3161	0.4461
MC s.e.	0.5146	0.5147	0.9027
SIF	1.1373	1.1320	0.9049
α_1^G	-	0.2786	-
MC s.e.	-	0.2191	-
SIF	-	0.5660	-
α_2^G	-	0.6570	-
MC s.e.	-	0.5395	-
SIF	-	1.0512	-

Table 1.5.4: Parameter estimates for the dependences structure (tri-variate diffusion, Student's t underlying) (cont.)

	t	t-G-SG (Large - Mid cap)	t-SG (nonnested)
$\overline{\alpha}_1^G$	-	0.2730	0.3434
MC s.e.	-	0.2961	0.5440
SIF	-	0.6114	0.7326
$\overline{\alpha}_2^G$	-	0.6660	-
MC s.e.	-	0.5939	-
SIF	-	1.3265	-
ω^G	-	0.5118	-
MC s.e.	-	0.4382	-
SIF	-	0.7870	-
$\overline{\omega}^G$	-	0.1529	0.2829
MC s.e.	-	0.2248	0.9130
SIF	-	1.4495	1.9105
ν	5.4774	3.9575	4.8266
MC s.e.	4.8170	2.4907	5.8874
SIF	0.8904	0.7732	0.9437
ρ_{12}	0.8184	0.7837	0.8166
MC s.e.	0.0074	0.0223	0.0171
SIF	0.3969	1.1166	1.2428
ρ_{13}	0.5113	0.4922	0.5522
MC s.e.	0.0286	0.0296	0.0085
SIF	1.5033	0.9770	0.6370
ρ_{23}	0.7165	0.7045	0.7372
MC s.e.	0.0085	0.0129	0.0092
SIF	0.3875	0.8620	0.6073

Table 1.5.5: Bayes factors

Log Bayes factors for tri-variate diffusions with dependence modeled using alternative copula functions. Benchmark models (M_b) are those involving the mixed copula diffusions with an Elliptic copula and the nested version of the extreme value Gumbel - Survival Gumbel copulas (Large-Mid cap being the most deeply nested couple). Two choices for the Elliptic copula are considered: the Gaussian one (Gauss-G-SG), and the Student's t one (t-G-SG). The four alternative diffusions ($M_j, j = 1, ..., 4$) are a Gaussian, a Student's t (t), and two nonnested versions of the mixture copula diffusion: the Gaussian-Gumbel-Survival Gumbel (Gauss-G-SG (nonnested)) and the Student's t - Survival Gumbel (t-SG (nonnested)).

	Gaussian	Student's t	Gauss-G-SG (nonnested)	t-SG (nonnested)
Gauss-G-SG (Large - Mid cap)				
Bayes factors	206.52	208.67	464.89	386.32
t-G-SG (Large - Mid cap)				
Bayes factors	215.58	217.73	473.95	395.02

latter is taken as the benchmark $\mathcal{M}_b$. But still this is far from the significantly higher values of the factors when the other alternative models are considered. This is not surprising, as the two nested mixture models are close in the way they treat the dependence structure, while the model with the Student's t underlying copula provides a more versatile way to account for dependence between extreme realizations.

1.6 Discussion and concluding remarks

In this chapter we introduce a multivariate diffusion model for stock prices based on copula functions that is able to reproduce a number of stylized facts for both the univariate return series and the dependence structure. It extends the univariate stationary diffusion modeling based on the Generalized Hyperbolic family of distributions that has proved successful in replicating dynamic return characteristics as a slowly decaying autocorrelation function of squared returns (or volatility clustering effect, as alternatively modeled under stochastic volatility or an ARCH process), or static properties like thick tails and excess kurtosis. Seeking to reproduce increased dependence when there are extreme market downturns, we extend the copula-GARCH approach to a continuous-time diffusion framework where the stationary distribution of the process is modeled using a copula function that can account for tail dependence. As well, it is achieved without including jumps in the stock price process, as in Das and Uppal (2004) or Liu et al. (2003). Such a process may prove useful for dynamic portfolio allocation applications, as tractable portfolio solutions can be obtained in this continuous time framework under market completeness.

There are a number of ways in which the model can be extended. There is overwhelming empirical evidence that the correlation of asset returns changes dynamically through time. Popular discrete time approaches include the GARCH-DCC model of Engle (2002), while in continuous time a promising alternative is the Wischart process of Bru (1991). Our model specification imposes constant conditional correlation for asset returns, that we have assumed for simplicity, but that can be extended to a more general model where correlation is modeled as either a function of the state variables of the model itself, or rendered stochastic by being represented as a function of exogenous factors. There is empirical evidence that the dynamics of asset return correlations are linked to the phase of the business cycle and tend to increase in periods of recession (e.g. Ledoit et al., 2003; Erb et al., 1994). As well, Longin and Solnik (1995) find that correlations for international stock market indices increase during hectic periods of high volatility.

Another possible extension concerns the dependence structure of the assets, modeled through a copula function. The present specification assumes that the parameters governing dependence are fixed. A number of studies have addressed time variation in dependence through a dynamic copula approach. In the case of modeling asymmetric dependence between exchange rates, Patton (2004) finds significant implications of the time variation in the copula dependence parameters, while Goorbergh et al. (2003) find substantial pricing differences for multivariate options when a dynamic copula model is used contrary to one with a fixed dependence structure, especially for market conditions marked with increased volatility. In our setup, time variation in the dependence parameter could be achieved by modeling it as a function of exogenous factors that are stochastically time varying themselves and that have a potential of explaining increased dependence in extreme down markets.

Chapter 2

Dependence Modeling of Joint Extremes via Copulas: A Dynamic Portfolio Allocation Perspective

2.1 Introduction

Modeling the dependence between asset returns is the corner stone for portfolio allocation decisions or risk management in general. Failing to account for specific features of the dependence structure of the data may lead to an improper assessment of the risk exposure and thus to suboptimal portfolio decisions. An already established stylized fact of asset returns is the co-movement asymmetry present in their dependence structure in that assets tend to be more correlated during bear markets than during bull markets. Using Extreme Value Theory, Longin and Solnik (2001) provide evidence of the dependence asymmetries present in several major stock market indices. Poon et al. (2004) confirm the evidence of stronger left tail dependence, but stress on the importance of considering parametric models that allow for asymptotic independence and thus avoid the risk of overestimating the probability of joint occurrence of tail events. There are also studies that provide a theoretical justification to this empirical fact: in a rational expectations equilibrium model Ribiero and Veronesi (2002) obtain endogenously excess stock return comovements during market downturns as a result of increased uncertainty about the state of the economy.

In this chapter we adopt a multivariate diffusion process for stock prices that is able to accommodate the above mentioned stylized facts. It has a predetermined stationary density that we model using copula functions that incorporate possibly asymmetric dependence in the upper or the lower tail of the distribution. The copula diffusion construction with a pre-specified stationary density relies on a result in Chen et al. (2002) and allows us to obtain increased tail dependence when markets suffer from extreme downturns. As the copula function underlying the stationary distribution of the process captures the needed dependence structure, we do not have to revert to the inclusion of jumps in prices and volatility as in Liu et al. (2003), or to systemic jumps common for all assets, as in Das and Uppal (2004), in order to replicate this stylized dependence feature.

From a portfolio allocation perspective, analyzing the extremal behavior that could be achieved through a multivariate diffusion by a copula construction is worthwhile, as tractable portfolio allocation rules and hedging behavior could be obtained by applying the martingale solution technique in a complete market setting. While copulas are widely studied in the context of multivariate option pricing and in credit risk modeling (mainly pricing of multiname credit derivatives), there are few applications for portfolio choice, all existing studies being focused on the unconditional portfolio behavior. As copulas can be designed to allow for a fairly general (possibly time varying) dependence structure, independently of the marginals, it is of interest to examine the influence of such a modeling approach on the dynamic hedging component of a portfolio, which constitutes the main motivation for the present chapter. Numerous studies have pointed out the adverse effect of high dependence among assets during market downturns on diversification benefits, but to our knowledge none has yet addressed the issue of isolating dynamic hedging demands that arise when the data generating process of asset returns incorporates this extremal dependence from the mean-variance component or hedging demands that arise from stochastically varying interest rates.

Unconditional portfolio selection under the influence of higher moments and dependence asymmetries has been studied by Chunhachinda et al. (1997) and Prakash et al. (2003) via polynomial goal programming. However, no extension for moments higher than the third one can be obtained using that approach. Jondeau and Rockinger (2005) overcome this problem by studying asset allocation by using a Taylor series approximation of expected utility as a function of higher moments. A study relying on copula specification for the dependence structure of asset returns is that of Patton (2004); however no extension to conditional allocation is offered. Das and Uppal (2004) examine dynamic portfolio choice in the presence of dependence asymmetries in the form of asymmetric conditional correlation by building a jump diffusion model with perfectly correlated (systemic) jumps across assets. An alternative specification aimed at capturing the same stylized behavior is that of Ang and Bekaert (2002) who review asset allocation under asymmetric response of correlation to returns in a regime-switching model. Surprisingly, in this context hedging demands (defined as the difference between a one-period ahead and a multi-horizon portfolio allocation) are found to be negligible, so that an investor loses little by acting myopically. However, these results are obtained under the CRRA assumption. An alternative preference specification could potentially increase the effect of extreme dependence asymmetries on dynamic portfolio choice (as for example a general utility specification like the HARA utility that models intolerance towards wealth levels under a certain boundary).

Having in mind the mixed evidence on the intertemporal hedging component of portfolio choice, conditional on either utility or distributional assumptions, we investigate its importance by generalizing the dependence framework beyond correlation modeling, which is a relevant dependence measure only in an elliptic distributions context. We also extend the utility specification beyond the CRRA case, and also consider the more general HARA utility. We rely on a simulation-based technique for portfolio selection, following Detemple et al. (2003) , where the need to accommodate a fairly general data-generating process is coupled by the need to overcome the curse of dimensionality of a large-scale problem. The solution methodology uses an extension of the Ocone and Karatzas (1991) formula under the complete markets assumption to obtain explicit expressions for the optimal portfolio and its hedging components for a general multivariate diffusion specification. The optimal investment strategy is represented as the sum of a myopic (mean-variance) component and two dynamic terms that represent hedges against changes in the short-term interest rate and the market price of risk. The explicit solutions for those terms involve expectations of random variables and their Malliavin derivatives that can be simulated using a standard discretization scheme. The approach remains tractable in large-scale problems, which matches one of the merits of dependence modeling via copula functions.

We use the above solution methodology to address the following questions. First, we isolate the intertemporal hedging demands that arise from a data generating process based on a Gumbel-Gaussian copula mix, that allows for asymptotic independence or dependence (possibly asymmetric) through varying weights in the mixture copula. We find substantial hedging demands, that increase with the investment horizon and decrease with the agent's degree of relative risk aversion. Those dynamic hedges are then compared to the ones obtained under the nested case of a Gaussian copula that serves as a tail independence benchmark model. An agent that uses the latter data generating process would then allocate more wealth to the risky assets as compared to an investor who is conscious of the existence of tail dependence. The extreme value Gumbel copula generates less risky asset demand than its symmetric tail counterpart – the Student's t copula.

Second, we study the economic significance of taking into account extreme tail dependence by quantifying it through the certainty equivalent cost (or the utility cost of behaving suboptimally).

Third, we check the robustness of our results to the choice of the utility function. We consider the benchmark CRRA case, as well as the more general HARA utility specification, that can be modeled to allow intolerance towards wealth shortfalls.

The reminder of the chapter is organized as follows. In section 2.2 we present the model, based on the asymmetric tail copula diffusion process. We further address the portfolio allocation problem, and the solution methodology in a complete market setting. In section 2.3 we study the importance of modeling extreme value dependence for dynamic portfolio selection. Section 2.4 concludes.

2.2 The model and the complete market portfolio solution

The evidence of increased dependence during market downturns than during market upturns has been shown to have a considerable impact on unconditional portfolio allocation when short sales are allowed and dependence is modeled within the copula framework (Patton, 2004). Ang and Bekaert (2002) also find significant costs of ignoring this dependence structure for the alternative way to replicate it through a regime-switching model that links high correlation with high volatility in the presence of a conditionally risk-free asset. Despite of that, they find insignificant intertemporal hedging demands, so that an investor would not lose much if he behaves myopically and solves just a one-period problem. However, in a recent paper Buraschi et al. (2007) show considerable hedging demands induced by time-varying correlation, their effect on total portfolio weights being the strongest in periods of market downturns. This would suggest that a model accounting for extremal dependence would be able to incite significant intertemporal demands. In order to isolate the effect of the spatial dependence structure of a process from that of time-changing correlation, we consider a process with a constant conditional correlation, but whose stationary distribution allows for possibly asymmetric dependence in the extremes. In what follows, we solve for the intertemporal hedging demands induced by such a model and study the implications of this particular tail behavior on optimal portfolio allocation.

We consider an investor with CRRA or the more general HARA utility over terminal wealth, allocating it between 1 riskless and 3 risky assets for a finite horizon T. The model could be extended to include any number of assets without rendering it intractable because of the flexibility offered by copula functions.

2.2.1 The general complete market setup

In a general complete market setup we assume that uncertainty is driven by a d-dimensional standard Brownian motion and that the price of the risky asset can be expressed as [1]:

$$S_{it} = \exp\left(\phi_i(t) + X_{it}\right), i = 1, ..., d \tag{2.2.1}$$

for some deterministic function of time $\phi_i(t)$, which we assume to be linear in t, $\phi_i(t) = k_i t$ with a linear trend parameter k_i, and where

$$dX_t = \mu(X_t)\, dt + \Lambda(X_t)\, dW_t \tag{2.2.2}$$

Thus, applying Itô's lemma we obtain for the price process for $i = 1, ..., d$:

$$\begin{aligned} dS_{it} &= S_{it}\mu_i^S(\ln S_{it} - k_i t)\, dt + S_{it}\sum_{j=1}^{d}\Lambda_{ij}(\ln S_{it} - k_i t)\, dW_{jt} \\ \text{where } \mu_i^S(X_t) &= \mu_i(X_t) + k_i + \frac{1}{2}\sum_{j=1}^{d}\overline{\sigma}_{ij}(X_t)^2 \end{aligned}$$

where $\overline{\sigma}_{ij}$ are entries of the matrix Λ in the diffusion term of the process for the de-trended log-price X. As pointed out in Bibby and Sorensen (1997), there is empirical evidence that the increments of the process for the log-price are nearly uncorrelated but not independent, which motivates the specification in (2.2.1). It is chosen as the most straightforward generalization of the Black Scholes model. The exact parametrization of the drift and the diffusion term will be discussed in the subsequent section, where we present a method to

[1] Following the parametrization of Bibby and Sorensen (1997) and Rydberg (1999)

construct a diffusion with a pre-specified stationary distribution.

The short rate r_t is the $(d+1)$'th state variable in the model and its dynamics are given by:

$$dr_t = \mu_t^r\left(r_t\right)dt + \sigma_t^r\left(r_t\right)dW_t \tag{2.2.3}$$

i.e. it depends on the same Brownian motion that drives the uncertainty for the log-price process. The money market account B_t follows:

$$dB_t = B_t r_t dt \tag{2.2.4}$$

The vector of state variables is then given by $Y = (X^\intercal r)$and satisfies the SDE:

$$dY_t = \mu_t^Y\left(Y_t, t\right)dt + \sigma_t^Y\left(Y_t, t\right)dW_t \tag{2.2.5}$$

where $\mu_t^Y\left(Y_t, t\right) = \begin{pmatrix} \mu\left(X_t\right)^\intercal & \mu_t^r\left(r_t\right) \end{pmatrix}^\intercal$, and $\sigma_t^Y\left(Y_t, t\right)$ is obtained by stacking the corresponding volatility terms from the SDEs for X_t and r_t.

As the market is assumed to be complete (given an invertible Λ matrix), we can define the market price of risk as:

$$\theta\left(X_t, r_t\right) = \theta\left(Y_t\right) = \theta_t = \Lambda\left(X_t\right)^{-1}\left[\mu^S\left(X_t\right) - r_t 1\right] \tag{2.2.6}$$

assumed to be continuously differentiable and satisfying the Novikov condition:

$$E\left[\exp\left(\frac{1}{2}\int_0^T \theta_s^\intercal \theta_s ds\right)\right] < \infty.$$

The associated state price density can be expressed as:

$$\xi_t \equiv \exp\left\{-\int_0^t r_s ds - \int_0^t \theta_s^\intercal dW_s - \frac{1}{2}\int_0^T \theta_s^\intercal \theta_s ds\right\} \tag{2.2.7}$$

and it satisfies the stochastic differential equation:

$$d\xi_t = -\xi_t r_t dt - \xi_t \theta_t^\intercal dW_t \tag{2.2.8}$$

2.2.2 The multivariate copula diffusion model

The process that we use for the state variables X, governing stock prices, is able to replicate certain univariate properties of asset returns as a leptokurtic distribution with respect to the normal density, volatility clustering and semi-heavy tails. The correct modeling of the tail behavior is particularly important as it is the impact of dependence between tail realizations on optimal portfolio decisions that we aim to study. Also, the model allows for a parsimonious treatment of the dependence structure and nests an array of dependence features, ranging from asymptotic dependence to tail independence. It is achieved through a flexible construction using copula functions that allow us to separate the impact of the properties of the marginal distributions from that of the dependence structure.

The construction of the multivariate diffusion for X that we discuss below follows the lines of Chapter 1 and relies on a result in Chen et al. (2002) and exploits the relationship that exists between the invariant density, the drift and the diffusion term for the process in (2.2.2):

$$\begin{aligned} \mu_j (x_1, ..., x_n) &= \frac{1}{2q(x_1, ..., x_n)} \sum_{i=1}^{d} \frac{\partial (v_{ij}(x_i, x_j) q(x_1, ..., x_n))}{\partial x_i} \\ \Sigma &= \Lambda\Lambda^\intercal \quad \text{with entries } v_{ij}(x_i, x_j) \end{aligned} \tag{2.2.9}$$

where Λ is a lower triangular matrix, q is a strictly positive continuously differentiable multivariate density function, which is indeed the stationary density of the process, and Σ is a continuously differentiable positive definite matrix. Thus, in order to specify the process for X, we need to determine its invariant density and propose a certain form for its diffusion term $\Lambda(X)$.

The invariant density q of the n-variate diffusion is obtained using the copula decomposition formula following Sklar's theorem that builds a multivariate distribution with density q using a dependence (copula) function $\widetilde{c}$ and marginal densities $f^i, i = 1, ..., n$, or rather functions $\widetilde{f^i} \propto f^i$ that are proportional to them, as we do not need the normalizing constant:

$$q(x_1, ..., x_n) \equiv \widetilde{c}(x_1, ..., x_n; \theta^c) \prod_{i=1}^{n} \widetilde{f^i}\left(x_i; \theta^{i,M}\right) \tag{2.2.10}$$

where $\widetilde{c}(x_1, ..., x_n; \theta^c) = c\left(F^1(x_1), ..., F^n(x_n); \theta^c\right)$. The copula function c is defined on the probability integral transforms F^i corresponding to each univariate series, so that it contains information on the dependence structure of the multivariate distribution regardless of the individual marginal specifications, as the latter are distributed as $Uniform\,(0,1)$. Thus, the parameters θ^c, pertaining to the copula, can be considered as driving dependence between the random variables $x_1, ..., x_n$. On the other hand, the univariate properties of each series are determined by the distributional assumptions on f^i and its corresponding parameters $\theta^{i,M}$.

For the marginal series we chose the Generalized Hyperbolic (GH) family of distributions, introduced by Barndorff-Nielsen (1977) and further used in a number a studies for modeling stochastic processes for stock prices (e.g. Eberlein and Keller, 1995; Bibby and Sorensen, 1997; Prause, 1999; Rydberg, 1999). It allows us to address univariate static properties of stock returns as departures from normality through high kurtosis or tails thicker than those implied by a Gaussian distribution, as well as dynamic features as persistence in autocorrelation for the squared increments of log prices, similar to stochastic volatility or GARCH models. From the perspective of the portfolio allocation application that we consider, correct accounting for univariate properties of the data is important, as it allows us to determine any demands that arise beyond those that could be attributed to the sensitivity of the investment in the risky assets to higher moments (as studied in Jondeau and Rockinger (2005) or Cvitanic et al. (2008)). Thus, we are interested in the portfolio implications of increased tail dependence that is ignited by the dependence structure, regardless of the marginals.

The form and properties of the GH family of distributions, as well as its different subclasses are discussed in the Appendix. One important property they have is the semi-heavy tails, expressed as:

$$\lim_{x\to\pm\infty} f_{GH}\left(x;\lambda,\alpha,\beta,\delta,\mu\right) \sim |x|^{\lambda-1}\exp\left\{\left(\mp\alpha+\beta\right)x\right\} \tag{2.2.11}$$

(Prause, 1999; Barndorff-Nielsen and Blaesid, 1981).

Thus, the class can easily accommodate any tail behavior ranging from power to exponential decline, and can account for tail asymmetries.

The dependence structure is entirely modeled by the copula $c(\cdot)$ and its corresponding parameters θ^c. As we aim at determining the impact of tail dependence on optimal portfolio demands, we consider several parametric families of copulas that allow for different degrees of dependence between extreme realizations. Before reviewing the alternative choices for the copula, recall that lower (λ_L) or upper (λ_U) tail dependence coefficients have the following representations in terms of the copula C:

$$\tau_T^U = \lim_{u \to 1} \frac{1 - 2u + C\left(u, u\right)}{1 - u}$$

$$\tau_T^L = \lim_{u \to 0} \frac{C\left(u, u\right)}{u}$$

First, we model dependence for the diffusion process for the state variables X using Elliptic copulas. A member of this family, the Gaussian copula, defines our benchmark dependence structure. Its tail coefficients are both zero, indicating no asymptotic dependence between the state variables. The second member of the Elliptic class of copulas that we consider is the Student's t copula, which allows for tail dependence, however symmetric, through its additional degrees of freedom parameter ν. The tail dependence coefficient for the t-copula is given by $\tau_T^U = \tau_T^L = 2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{1-\rho}/\sqrt{1+\rho}\right)$, where ρ is an off-diagonal element of the correlation matrix and $t_\nu^{-1}\left(u\right)$ is the inverse of the univariate CDF of the Student's t distribution. Tail dependence decreases for increasing levels of the degrees of freedom parameter and eventually goes to zero when $\nu \to \infty$, i.e. when the t-copula converges to the Gaussian one.

Next we consider the Archimedean family of copulas, and more specifically the extreme value Gumbel copula that can model upper tail dependence through its dependence parameter α^G, rendering $\tau_{Gumbel}^U = 2 - 2^{\alpha}$, while $\tau_{Gumbel}^L = 0$, and its survival counterpart for which the roles of upper and lower tail dependence switch places. Combining those copulas by assigning weights to each one of them renders a dependence function that has asymmetric upper and lower tail dependence, determined by the corresponding Gumbel (α^G) and Survival Gumbel ($\overline{\alpha}^G$) parameters. In order to take into account the possibility that the state variables do not exhibit asymptotic dependence, we add to the above mixture copula the Gaussian one, so that we obtain:

$$
\begin{aligned}
& C_m^{Ga}\left(u; R_{Ga}, \alpha^G, \overline{\alpha}^G, \omega, \overline{\omega}\right) && (2.2.12) \\
= \quad & \omega C^G\left(u; \alpha^G\right) + \overline{\omega}\overline{C}^G\left(u; \overline{\alpha}^G\right) + (1 - \omega - \overline{\omega})\, C^{Ga}\left(u; R_{Ga}\right)
\end{aligned}
$$

where C^G refers to the Gumbel copula, $\overline{C}^G$ - to the Survival Gumbel, and C^{Ga} - to the Gaussian, and the parameters $\{\omega, \overline{\omega}\}$ are their corresponding weights. Thus, our benchmark tail independent model is obtained by setting the weights ω and $\overline{\omega}$ to zero, while any weight parameter different from zero would entail possibly asymmetric upper (lower) tail dependence. In order to obtain our symmetric tail benchmark, we alternatively build a mixture dependence function using the Student's t copula instead:

$$
\begin{aligned}
& C_m^{t}\left(u; R_T, \upsilon, \alpha^G, \overline{\alpha}^G, \omega, \overline{\omega}\right) && (2.2.13) \\
= \quad & \omega C^G\left(u; \alpha^G\right) + \overline{\omega}\overline{C}^G\left(u; \overline{\alpha}^G\right) + (1 - \omega - \overline{\omega})\, C^{T}\left(u; R_T, \upsilon\right)
\end{aligned}
$$

where C^T refers to the t-copula.

For the above mixture copulas we consider the nested version of the Gumbel copula, as described in the Appendix. It allows for different dependence parameters between consecutively nested couples of variables, and thus permits a more general treatment of the dependence structure than the usual n-variate Gumbel copula which imposes the same parameter across all variables. The latter (non-nested) specification has a more parsimonious nature, but potentially restraints the achievable degrees of tail dependence. As the construction of a nested Archimedean copula is not so straightforward in higher dimensions and in order to investigate the portfolio implications of assuming a homogenous dependence structure across assets, we also consider the above mixture copulas for non-nested versions of the Gumbel and the Survival Gumbel copulas:

$$
\begin{aligned}
& C_m^{Ga^*}\left(u; R_{Ga}, \alpha_*, \overline{\alpha}_*, \omega, \overline{\omega}\right) && (2.2.14) \\
= \quad & \omega C_*^G\left(u; \alpha_*\right) + \overline{\omega}\overline{C}_*^G\left(u; \overline{\alpha}_*\right) + (1 - \omega - \overline{\omega})\, C^{Ga}\left(u; R_{Ga}\right)
\end{aligned}
$$

when using the Gaussian copula, or:

$$C_m^{t^*}(u;R_T,\upsilon,\overline{\alpha}_*,\overline{\omega})=\overline{\omega}\overline{C}_*^G(u;\overline{\alpha}_*)+(1-\overline{\omega})\,C^T(u;R_T,\upsilon) \tag{2.2.15}$$

for the Student's t copula, where α_* is the dependence parameter for the nonnested Gumbel copula that determines upper tail dependence, and $\overline{\alpha}_*$ is the parameter of the nonnested Survival Gumbel copula that determines lower tail dependence. Note that in the last case we have used only the Survival Gumbel dependence function. It is the most parsimonious mixture that allows for asymmetric behavior in the tails, as the t-copula already models both upper and lower tail dependence, while the Survival Gumbel parameter adds asymmetry to the structure by adding additional weight for the dependence in the left tail. This is indeed the stylized fact of stock returns that we seek to reproduce: increased dependence when markets jointly decline.

The form of the copula functions used above is given in more detail in the appendix.

Finally, the only term that is left to be determined in (2.2.9) is the specification of the diffusion term of the process. For it we chose a constant conditional correlation specification, given by:

$$\begin{aligned} v_{ij}(x_i,x_j) &= \rho_{ij}\sigma_i^X(x_i)\,\sigma_j^X(x_j) \\ \sigma_i^X(x_i) &= \sigma_i\left[\widetilde{f}^i(x_i)\right]^{-\frac{1}{2}\kappa_i} \end{aligned} \tag{2.2.16}$$

which extends the univariate specification of Bibby and Sorensen (2003) to the case of a multivariate diffusion, where $\left(\sigma_i^X\right)^2>0$ and $\kappa_i\in[0,1]\,,i=1,...,d$. The function $\widetilde{f}^i(x_i)\propto f^i(x_i)$, i.e. it is proportional to the ith univariate marginal distribution, chosen to belong to the GH family. Note that for the sake of simplicity we have assumed a constant conditional correlation specification through the time invariant parameter ρ_{ij}. This setup could be further extended by modeling the correlation parameter as a function of stochastic state variables, but as we are interested in the dependence achievable through the unconditional distribution of the process, we restrain from considering this more general case.

In what follows, we will briefly present the martingale solution technique for the portfolio allocation problem at hand that gives rise to the solution for the optimal terminal wealth (Cox and Huang, 1989) and the financing portfolio (Ocone and Karatzas, 1991), as well as the Monte Carlo solution method for finding optimal portfolio shares proposed by Detemple et al. (2003).

2.2.3 The investor's problem and the optimal portfolio policy

Considering the case of no intermediate consumption, the evolution of wealth equation is given by:

$$d\omega_t = r_t \omega_t dt + \omega_t \alpha_t^{\intercal} \left[\left(\mu_t^S - r_t 1 \right) dt + \Lambda_t dW_t \right], \quad \omega_0 = \overline{\omega} \tag{2.2.17}$$

where ω_t denotes the wealth at time t and α_t - the amount of wealth invested in the risky assets. Working under the assumption of time-separable von Neumann-Morgenstern preferences, the investor's problem of optimally allocating terminal wealth ω_T for an investment horizon T between 1 riskless and d risky assets is as follows:

$$\max_{\omega_T} U(\omega_T) \equiv E\left[u(\omega_T) \right] \tag{2.2.18}$$

conditional on the dynamic budget constraint given by (2.2.17), and the nonnegativity of wealth constraint $\omega_t \geq 0$, where $u(\cdot)$ is a strictly increasing and concave utility function that satisfies the Inada conditions $\lim_{x\to\infty} u'(x) = 0$ and $\lim_{x\to 0} u'(x) < \infty$.

The equivalent static optimization problem, as shown in Cox and Huang (1989), is reduced to maximizing expected utility of terminal wealth, subject to a static budget constraint:

$$E\left[\xi_T \omega_T \right] \leq \overline{\omega} \tag{2.2.19}$$

and the non-negativity of wealth constraint. After forming the Lagrangian for this static constrained optimization problem, the first order conditions for optimality are expressed as:

$$\begin{aligned} u'(\omega_T) &= y \xi_T \\ E\left[\xi_T \omega_T \right] &\leq \overline{\omega} \end{aligned}$$

where y is the Lagrange multiplier or the shadow price for the budget constraint. Letting $I(\cdot)$ denote the inverse of the marginal utility function, it follows from Cox and Huang (1989) that the optimal terminal wealth is given explicitly by $\omega_T^* = \max\left(I\left(y\xi_T\right), 0\right)$ and y satisfies the static budget constraint $E\left[\xi_T \max\left(I\left(y\xi_T\right), 0\right)\right] = \overline{\omega}$. Thus the optimal expression for terminal wealth leads us to the optimal wealth at time $t < T$:

$$\xi_t \omega_t^* = E_t\left[\xi_T \max\left(I\left(y\xi_T\right), 0\right)\right] \tag{2.2.20}$$

In order to find the optimal portfolio policy that generates this optimal wealth process, Ocone and Karatzas (1991) use the Clark-Ocone formula which allows expressing optimal portfolio shares as expectations of the state variables and their Malliavin derivatives. According to the Clark-Ocone formula, any random variable X can be decomposed into an expectation part and a volatility part that involves Malliavin derivatives: $X = E[X] + \int_0^T E_t\left[\mathcal{D}_t X\right] dW_t$. In fact, this formula identifies the integrand in the Martingale Representation theorem. Using this result, Ocone and Karatzas (1991) proceed to explicitly determining the optimal portfolio policy α_t^* by considering the discounted wealth process $\xi_t \omega_t^*$. On one hand, using Itô's lemma on (2.2.17) and (2.2.8), the volatility of the process is given by $-\xi_t \omega_t^* \theta_t^\intercal + \xi_t \omega_t^* \alpha_t^\intercal \Lambda_t$. On the other hand, an application of the Clark-Ocone formula states that the volatility of $\xi_t \omega_t^*$ is given by its Malliavin derivative $\mathcal{D}_t\left(\xi_t \omega_t^*\right)$. Equating both terms leaves us with the following explicit expression for the optimal portfolio:

$$\alpha_t^* = \left(\Lambda_t^\intercal\right)^{-1} \theta_t + \left(\xi_t \omega_t^*\right)^{-1} \left(\Lambda_t^\intercal\right)^{-1} \left(\mathcal{D}_t\left(\xi_t \omega_t^*\right)\right)^\intercal \tag{2.2.21}$$

where $\xi_t \omega_t^*$ is given by (2.2.20). Thus, to solve for the optimal portfolio, we need to evaluate the expression involving the Malliavin derivative of the discounted wealth process. In order to do so, we need to revert to the chain rule of Malliavin calculus[2]:

$$\begin{aligned} \mathcal{D}_t\left(\xi_t \omega_t^*\right) &= \mathcal{D}_t\left(E_t\left[\xi_T \max\left(I\left(y\xi_T\right), 0\right)\right]\right) \\ &= E_t\left[\mathcal{D}_t\left(\xi_T I\left(y\xi_T\right)^+\right)\right] \\ &= E_t\left[\left[I\left(y\xi_T\right)^+ + y\xi_T \frac{\partial I\left(y\xi_T\right)}{\partial\left(y\xi_T\right)} 1_{I(y\xi_T)>0}\right] \mathcal{D}_t \xi_s\right] \end{aligned} \tag{2.2.22}$$

[2] For a random variable X and a differentiable function $\varphi(X)$, the Malliavin derivative of X is given by $\mathcal{D}_t \varphi\left(X_s\right) = \frac{\partial \varphi(X)}{\partial X} \mathcal{D}_t X_s$.

The Malliavin derivative of a stochastic process satisfying a SDE given by $Y_t = Y_0 + \int_0^t \mu\left(Y_s\right) ds + \int_0^t \sigma\left(Y_s\right) dW_s$ satisfies $\mathcal{D}_t Y_s = \int_t^s \frac{\partial \mu(Y_s)}{\partial Y}\left(\mathcal{D}_t Y_v\right) dv + \int_t^s \frac{\partial \sigma(Y_s)}{\partial Y}\left(\mathcal{D}_t Y_v\right) dW_s$.

where we have used the fact that the conditional expectation operator and the Malliavin derivative operator commute. Applying further the chain rule on $\mathcal{D}_t\xi_s$ and using the SDE satisfied by the state price density process (2.2.8), we obtain:

$$\mathcal{D}_t\xi_s = -\xi_s \left[\int_t^s \left(\mathcal{D}_t r_v + \theta_v^{\intercal} \left(\mathcal{D}_t\theta_v\right)\right) dv + \int_t^s dW_v^{\intercal} \cdot \left(\mathcal{D}_t\theta_v\right) + \theta_t^{\intercal} \right] \tag{2.2.23}$$

As the short rate r_t and the market price of risk θ_t processes depend in turn on the state variables Y_t, we can develop further the above expression for the Malliavin derivative of the state price density process, by realizing that:

$$\begin{aligned} \mathcal{D}_t r_v &= \frac{\partial r\left(Y_s\right)}{\partial Y}\mathcal{D}_t Y_s \\ \mathcal{D}_t \theta_v &= \frac{\partial \theta\left(Y_s\right)}{\partial Y}\mathcal{D}_t Y_s \end{aligned} \tag{2.2.24}$$

and that the Malliavin derivatives of the state variables satisfy:

$$d\left(\mathcal{D}_t Y_s\right) = \frac{\partial \mu^Y\left(Y_s\right)}{\partial Y}\left(\mathcal{D}_t Y_s\right) ds + \sum_{j=1}^{d+1} \frac{\partial \sigma_{\cdot j}^Y\left(Y_s\right)}{\partial Y} dW_{js}\left(\mathcal{D}_t Y_s\right) \tag{2.2.25}$$

where $\sigma_{\cdot j}^Y\left(Y_s\right)$ is the j^{th} column of the volatility term for Y_t. Thus, the solution for the optimal portfolio weights is reduced to the computation of conditional expectations of state variables and their Malliavin derivatives. Realizing that $I\left(y\xi_T\right) + y\xi_T \frac{\partial I\left(y\xi_T\right)}{\partial\left(y\xi_T\right)} = \omega_T\left(1 - \frac{1}{R\left(\omega_T\right)}\right)$, where $R\left(\omega_T\right)$ is the coefficient of relative risk aversion, leads us to the explicit expressions for the optimal portfolio weights given by Theorem 1 in Detemple et al. (2003) as represented below, whose contribution to the Ocone and Karatzas formula lies in the realization that Malliavin derivatives satisfy stochastic differential equations and can thus be simulated using Monte Carlo methods and standard discretization schemes like the Euler scheme. The optimal portfolio rules are decomposed into a mean-variance term and two hedging expressions, that are given in terms of conditional expectations involving the Malliavin derivatives of the interest rate process and the market price of risk process:

$$
\begin{aligned}
\alpha_t &= \alpha_t^{MV} + \alpha_t^{IRH} + \alpha_t^{MPRH} \qquad (2.2.26)\\
\alpha_t^{MV} &= \left(\Lambda_t \left(X_t\right)^{\intercal}\right)^{-1} \frac{1}{R\left(\omega_t\right)} \theta\left(Y_t\right) E_t \left[\xi_{t,T} \frac{\omega_T}{\omega_t} \frac{R\left(\omega_t\right)}{R\left(\omega_T\right)} I_{\omega_T>0}\right]\\
\alpha_t^{IRH} &= -\left(\Lambda_t \left(X_t\right)^{\intercal}\right)^{-1}\\
&\quad E_t \left[\xi_{t,T} \frac{\omega_T}{\omega_t} \left(1 - R\left(\omega_T\right)^{-1}\right) I_{\omega_T>0} \int_t^T \mathcal{D}_t r_s ds\right]\\
&= -\left(\Lambda_t \left(X_t\right)^{\intercal}\right)^{-1} a\left(X_t, r_t\right)\\
\alpha_t^{MPRH} &= -\left(\Lambda_t \left(X_t\right)^{\intercal}\right)^{-1}\\
&\quad E_t \left[\xi_{t,T} \frac{\omega_T}{\omega_t} \left(1 - R\left(\omega_T\right)^{-1}\right) I_{\omega_T>0} \int_t^T \left(dW_s + \theta_s ds\right)^{\intercal} \mathcal{D}_t \theta_s\right]\\
&= -\left(\Lambda_t \left(X_t\right)^{\intercal}\right)^{-1} b\left(X_t, r_t\right)
\end{aligned}
$$

Thus, the mean-variance component α_t^{MV} gives the portfolio allocation for a single-period investor or one with a log-utility function, while the other two terms reflect the behavior of an investor who hedges against future changes in the short rate (the α_t^{IRH} term) and the market price of risk (the α_t^{MPRH} term), as the Malliavin derivatives measure the sensitivity of the state variables to innovations in the Brownian motions that drive uncertainty. If we have a constant opportunity set or a log-utility investor with unit relative risk aversion, the hedging terms disappear and the portfolio is entirely determined by mean-variance optimization. Alternatively, if relative risk aversion tends to infinity, the mean-variance component will tend to zero and the portfolio will be entirely given by the intertemporal hedging terms in the limit.

As we are interested in the effect of extreme value dependence in the state variable process on optimal allocation rules, this ability to split the portfolio terms into a mean-variance term and an intertemporal hedging term is particularly appealing, as we could then test whether future changes in the opportunity set driven by this particular form of dependence have an effect on the hedging demand in the following two perspectives: whether it can induce a substantial hedging demand as a part of the total allocation, and whether hedging demands actually diminish the total portfolio allocation in the risky assets and shift it to the riskless money market account when extremal dependence is present, as compared to a case with no extremal dependence in the stationary distribution of the state variables.

Further, as the proposed state variable process is fairly general and includes substantial non-linearities due to the copula functions and the form of the marginal distributions, the ability to obtain portfolio shares through a simulation-based technique is crucial.

2.2.4 Implementation through Monte Carlo simulations

The Monte Carlo simulation technique implemented in Detemple et al. (2003) proceeds as follows. The state variables and their Malliavin derivatives form a joint system $(Y_s, \mathcal{D}_t Y_s)$, to which we add the relative state prices $\xi_{t,s} = \frac{\xi_s}{\xi_t}$, as well as the two integrals in the hedging terms:

$$\begin{aligned} H_{t,s}^{IR} &= \int_t^s \mathcal{D}_t r_s ds = \int_t^s \frac{\partial r(Y_v)}{\partial Y} \mathcal{D}_t Y_v dv \\ H_{t,s}^{MPR} &= \int_t^s (dW_v + \theta_v ds)^\intercal \mathcal{D}_t \theta_v = \int_t^s (dW_v + \theta_v ds)^\intercal \frac{\partial \theta(Y_v)}{\partial Y} \mathcal{D}_t Y_v dv \end{aligned}$$

As all these terms solve stochastic differential equations, they can be simulated using a standard discretization scheme, so that we obtain a set of MC estimates

$$\left(Y_s^i, \mathcal{D}_t Y_s^i, \xi_{t,s}^i, H_{t,s}^{IR,i}, H_{t,s}^{MPR,i}\right)_{i=1}^{MC}$$

where MC is the number of Monte Carlo paths that are being simulated.

The hedging terms depend further on the particular choice of a utility function. We take into consideration two utility function specifications: the Constant Relative Risk Aversion (CRRA) one and the Hyperbolic Absolute Risk Aversion (HARA) utility function, of which CRRA is a special case. We choose CRRA because of the considerable simplicity it introduces in the hedging term expressions, while the HARA utility not only introduces more generality in the portfolio problem, but also allows for a more pronounced effect of the extremal dependence structure on portfolio hedging demands through the intolerance towards wealth being below a certain threshold that it implies. The HARA utility function is given by:

$$u(x) = \frac{1}{1-R}(x+B)^{1-R} \tag{2.2.27}$$

where R and B are exogenous constants. The special CRRA case is obtained by setting $B = 0$. The coefficient of relative risk aversion is given by $R(x) = \frac{R}{x+B}x$, which is simply equal to R in the case of CRRA. When $B < 0$ the utility function displays intolerance towards wealth falling below the threshold $-B$.

For the benchmark CRRA case, the portfolio weights simplify considerably. The inverse of the marginal utility function is given by $I(z) = z^{-\frac{1}{R}}$, so that optimal terminal wealth is $\omega_T = (y\xi_T)^{-1/R}$, and the constant y is given by $y = \left(\frac{E_0\left[\xi_T^{1-1/R}\right]}{\overline{\omega}}\right)^R$.The mean-variance term is also simplified to $\alpha_t^{MV} = \left(\Lambda_t(X_t)^\intercal\right)^{-1}\frac{1}{R}\theta(Y_t)$, while the two hedging terms have the following expressions, independent of wealth:

$$
\begin{aligned}
\alpha_t^{IRH} &= -\left(\Lambda_t(X_t)^\intercal\right)^{-1}\left(1-\frac{1}{R}\right)E_t\left[\frac{\xi_{t,T}^{(1-1/R)}}{E_t\left[\xi_{t,T}^{(1-1/R)}\right]}\int_t^T \mathcal{D}_t r_s ds\right] \qquad (2.2.28)\\
\alpha_t^{MPRH} &= -\left(\Lambda_t(X_t)^\intercal\right)^{-1}\left(1-\frac{1}{R}\right)\\
&\quad E_t\left[\frac{\xi_{t,T}^{(1-1/R)}}{E_t\left[\xi_{t,T}^{(1-1/R)}\right]}\int_t^T (dW_s+\theta_s ds)^\intercal \mathcal{D}_t\theta_s\right] \qquad (2.2.29)
\end{aligned}
$$

The conditional expectations could then be estimated by averaging over the MC terminal values of the simulated paths. The rate of convergence of these estimated values to the true ones depends on the number of Monte Carlo paths and is of order $1/\sqrt{MC}$.

The solution in the case of HARA utility follows the same lines with the exception that now optimal wealth enters the portfolio terms. What is more, in this case the non-negativity of wealth constraint may become binding for the case of $B > 0$.

A key to improving efficiency of the simulation method of Detemple et al. (2003) is the transformation of the state variable process to one with unit volatility, which allows eliminating the stochastic integral from the Malliavin derivatives. Even though this method is quite appealing for univariate diffusions, its generalization to a multivariate one is not so straightforward. The diffusions of the GH family are indeed impossible to transform in closed form, unless the κ coefficient in their diffusion term is set to 0. The same is true for the multivariate construction as well, unless we consider the simple construction of a diffusion of the gradient field type with a constant volatility coefficient.

Thus we need to resort to other methods that achieve variance reduction in the Monte

Carlo simulations. The use of low discrepancy points is one such possibility. The low discrepancy sequences are formed of selected deterministic points and have the property of spanning the whole region of interest. Their advantage over random points in financial applications have been established in numerous studies, among which Joy et al. (1996), Boyle and Imai (2002). We have chosen the Sobol low discrepancy sequence for the Monte Carlo simulations for finding optimal portfolio shares.

2.2.5 The short rate process

The short rate does not enter the specification for the adjusted log price X, and for simplicity we assume that it follows a Vasicek process:

$$dr_t = \kappa^r \left(\theta^r - r_t\right) dt + \sigma^r dW_t \tag{2.2.30}$$

This allows for an analytic expression for the Malliavin derivative of the short rate, as (2.2.25) can be explicitly solved, as all stochastic terms disappear:

$$\mathcal{D}_t r_s = \sigma^r \exp\left(-\kappa^r \left(s - t\right)\right) \tag{2.2.31}$$

The estimated parameters for the short rate are $\kappa^r = 0.2001$, $\theta^r = 0.0293$, $\sigma^r = -0.0069$[3].

2.2.6 Induced hedging demands

We could further split the hedge terms into induced demands that arise from hedging fluctuations in the sources of risk to which all the other state variables are exposed, and demands that are related to the source of risk, specific only to the particular asset. Due to the fact that Λ, whose terms are defined in (2.2.16), is a lower d-dimensional triangular matrix, the asset d is the only one that is exposed to W_d risk. Thus, its demand is governed only by the need to hedge against fluctuations in W_d, the other sources of risk being hedged away by the rest of the assets. However, the asset $d-1$ is no longer the only one exposed to W_{d-1} risk, as the asset d shares this exposure as well. Thus, the position in asset d will induce hedging demands in $d-1$, and so forth. To see that more clearly, let $\lambda_{(i,j)}$ represent the

[3] Data for the 3-month Treasury bill rates is from the H.15 Federal Reserve Statistical Release.

terms of the inverse of the Λ matrix. Then the interest rate and the market price of risk hedging terms for asset $d-1$ can be decomposed into:

$$\begin{aligned} \alpha_{t,(d-1)}^{IRH} &= -\lambda_{(d-1,d-1)} a_{(d-1)}\left(X_t, r_t\right) - \lambda_{(d,d-1)} a_{(d)}\left(X_t, r_t\right) \\ \alpha_{t,(d-1)}^{MPRH} &= -\lambda_{(d-1,d-1)} b_{(d-1)}\left(X_t, r_t\right) - \lambda_{(d,d-1)} b_{(d)}\left(X_t, r_t\right) \end{aligned} \quad (2.2.32)$$

The last term in those expressions thus refers to the induced hedging demands for asset $d-1$.

2.3 The importance of modeling extreme value dependence

Having established the solution technique and its implementation through Monte Carlo simulations, we proceed to the core of our study that is establishing the effect of extreme value dependence in the state variable process on the optimal portfolio policy. To this end, we consider a benchmark model for which the dependence tends to zero as we go further in the tails of the stationary distribution (the Gaussian diffusion model), and five models that allow for tail dependence: one symmetric (the Student's t diffusion) and four asymmetric (the extreme value mixture diffusion of Gaussian, Gumbel and survival Gumbel copulas in the stationary distribution as well as the mixture of Student's t with Gumbel and survival Gumbel copulas, in their nested and non-nested forms).

The empirical application of the portfolio solution described above relies on data from the daily CRSP database. More specifically, we consider US stock capitalization decile indices for the period 1990-2005. These indices represent yearly rebalanced portfolios based on market capitalization. The stock universe includes stocks listed on NYSE, AMEX, and NASDAQ. All ten capitalization decile indices were grouped in three sub-categories: small-cap (deciles 1-3), mid-cap (deciles 4-7), and large-cap (deciles 8-10). This dataset has been used in Ang and Chen (2002) to study the exceedence correlation patterns of the market and stock portfolios, as well as in Patton (2004) for the portfolio implications of this form of dependence in an unconditional context.

The above construction of a stationary diffusion with a prespecified stationary density (2.2.9)-(2.2.16) poses a serious estimation problem, as its conditional density is not explicitly known. Thus, as with the implementation of the solution for the optimal portfolio, we rely on the standard Euler discretization scheme with data augmentation, i.e. introducing latent data points between each pair of observations. This technique has been used in Pedersen (1995) for simulated maximum likelihood estimation of diffusions, or in Elerian et al. (2001), Roberts and Strammer (2001), or Eraker (2001) for MCMC analysis. The estimation scheme we apply in the present setup relies on an MCMC estimation algorithm with data augmentation following the sequential inference procedure of Golightly and Wilkinson (2006a). Details of the algorithm are presented in the appendix. We use a two-step estimation procedure which allows us to choose the appropriate marginal distribution for each data series. We estimate a NIG stationary distribution for all series, except the Mid caps, for which the more general GH construction appears to be appropriate (a NIG diffusion for the Mid caps is rejected on the basis of the uniform residuals obtained by the probability integral transform). Table 2.3.1 summarizes the estimation results for the parameters specific to each univariate series, while Table 2.3.2 gives the estimated parameters that describe the dependence structure for the diffusion specifications we consider.

As we are interested rather in the portfolio implications of tail dependence, we propose two experiments. The first one consists in calculating portfolio shares along realized paths of asset prices. Considering the state variable processes (the short rate and the de-trended log-returns) as given by their realized paths, we simulate the Malliavin derivatives and the hedging terms along these paths for the whole estimation horizon, and then compute optimal allocations while keeping the investment horizon fixed at its terminal value. Thus we can analyze the differences in the hedging demands obtained under the alternative models over a very long investment horizon, where the effect of the stationary distribution would indeed be the most visible. We repeat this exercise with a rolling-window horizon instead of a fixed one in order to evaluate the impact of a long horizon on the hedging demands.

The second experiment is a simulation study for varying investment horizons and degrees of risk aversion, while keeping the starting point fixed this time, in which all state variables and hedging terms are simulated ahead. By looking at the profile of the optimal portfolio policy for each horizon over different levels of risk aversion, we are able to determine to which extent hedging demands are sensitive to the level of risk aversion in the utility function or to the choice of utility function (CRRA or the more general HARA utility).

Table 2.3.1: Parameter estimates for the univariate series

The table summarizes the posterior parameter estimates from the MCMC output. Monte Carlo standard errors are reported in parenthesis (multiplied by a factor of 1000) (obtained using the batch-mean approach). SIF refers to the simulation inefficiency factor for each parameter (its integrated autocorrelation time).

	Small cap	*Mid cap*	*Large cap*
α	3.0502	18.7839	10.6904
(MC s.e.)	(0.1616)	(0.5220)	(0.2193)
(SIF)	(0.0938)	(0.6694)	(0.6912)
β	-0.5911	0.4476	-1.5737
(MC s.e.)	(0.6329)	(2.9453)	(1.5404)
(SIF)	(0.1104)	(1.5392)	(1.7637)
δ^2	0.0301	0.0721	0.0410
(MC s.e.)	(0.0024)	(0.0011)	(0.0031)
(SIF)	(0.1219)	(1.0535)	(1.8122)
μ	6.7059	6.3101	6.5360
(MC s.e.)	(0.0249)	(0.0129)	(0.0102)
(SIF)	(0.1038)	(0.5407)	(0.4991)
σ^2	0.0406	0.0400	0.0082
(MC s.e.)	(0.0022)	(0.0030)	(0.0006)
(SIF)	(0.1142)	(1.4686)	(1.2930)
κ	0.6490	0.4670	0.5102
(MC s.e.)	(0.0373)	(0.0235)	(0.0850)
(SIF)	(0.0955)	(1.4322)	(1.7551)
λ	0.5	-1.4295	0.5
(MC s.e.)	-	(0.0519)	-
(SIF)	-	(1.1704)	-

Table 2.3.2: Parameter estimates for the dependences structure

Estimation results for the trivariate diffusions using the Gaussian copula, the nested Gaussian-Gumbel-Survival Gumbel mixture copula (the most deeply nested couple is given in parenthesis), the nonnested Gaussian-Gumbel-Survival Gumbel mixture copula, the Student's t copula, the Student's t – nonnested Survival Gumbel mixture copula, and the Student's t – nested Gumbel - Survival Gumbel mixture copula. Monte Carlo standard errors (multiplied by a factor of 1000), and Simulation Inefficiency Factors (SIF) are given in parenthesis. The first three parameters (R_{12}, R_{13}, R_{23}) correspond to the off-diagonal entries of the correlation matrix R_{Ga} for the Gaussian copula or the correlation matrix R_T for the Student's t copula. The parameters α_1^G and α_2^G are the dependence parameters for the nested Gumbel copula, and the parameters $\overline{\alpha}_1^G$ and $\overline{\alpha}_2^G$ are the dependence parameters for the nested Survival Gumbel copula. For the nonnested cases, the relevant parameters are α_1^G for the Gumbel copula and $\overline{\alpha}_1^G$ for the Survival Gumbel copula. ω^G and $\overline{\omega}^G$ are the corresponding weights for the Gumbel and the survival Gumbel copula for the mixture model. ν is the degrees of freedom parameter for the Student's t copula. The parameters ρ_{12}, ρ_{13}, and ρ_{23} are the off-diagonal entries of the correlation matrix in the diffusion specification (3.11). Results are obtained for 50000 Monte Carlo replications with a thinning factor of 5 with 10 latent data points simulated between each pair of observations.

	Gaussian	Gauss-G-SG (Large Mid cap)	Gauss-G-SG (nonnested)	t	t-G-SG (Large-Mid cap)	t-SG (nonnested)
R_{12}	0.5671	0.5347	0.5758	0.4408	0.2574	0.5266
MC s.e.	0.3701	0.3326	0.3537	0.5433	1.4015	0.6040
SIF	0.8621	1.0437	0.9540	1.3619	0.7629	1.3392
R_{13}	0.2723	0.5179	0.2571	0.5273	0.2362	0.4154
MC s.e.	0.7875	0.4191	0.5131	0.6911	0.9873	0.6353
SIF	0.7359	0.7188	0.7251	0.9564	1.0469	0.8209
R_{23}	0.5207	0.4152	0.4698	0.3334	0.3161	0.4461
MC s.e.	0.4399	0.3302	1.3536	0.5146	0.5147	0.9027
SIF	0.9162	1.6992	1.5260	1.1373	1.1320	0.9049
α_1^G	-	0.2972	0.4494	-	0.2786	-
MC s.e.	-	0.3546	0.3541	-	0.2191	-
SIF	-	0.5754	1.2328	-	0.5660	-
α_2^G	-	0.6335	-	-	0.6570	-
MC s.e.	-	0.1928	-	-	0.5395	-
SIF	-	0.9156	-	-	1.0512	-

Table 2.3.2: Parameter estimates for the dependences structure (cont.)

	Gaussian	Gauss-G-SG (Large Mid cap)	Gauss-G-SG (nonnested)	t	t-G-SG (Large-Mid cap)	t-SG (nonnested)
$\overline{\alpha}_1^G$	-	0.3618	0.4354	-	0.2730	0.3434
MC s.e.	-	0.1998	1.0229	-	0.2961	0.5440
SIF	-	0.2375	1.6558	-	0.6114	0.7326
$\overline{\alpha}_2^G$	-	0.6544	-	-	0.6660	-
MC s.e.	-	0.4667	-	-	0.5939	-
SIF	-	0.8040	-	-	1.3265	-
ω^G	-	0.3321	0.3832	-	0.5118	-
MC s.e.	-	1.0111	0.7265	-	0.4382	-
SIF	-	2.0983	1.0348	-	0.7870	-
$\overline{\omega}^G$	-	0.2853	0.2324	-	0.1529	0.2829
MC s.e.	-	0.3789	0.3619	-	0.2248	0.9130
SIF	-	1.4739	2.1457	-	1.4495	1.9105
ν	-	-	-	5.4774	3.9575	4.8266
MC s.e.	-	-	-	4.8170	2.4907	5.8874
SIF	-	-	-	0.8904	0.7732	0.9437
ρ_{12}	0.7894	0.7917	0.8287	0.8184	0.7837	0.8166
MC s.e.	0.0195	0.0086	0.0104	0.0074	0.0223	0.0171
SIF	1.2371	0.2271	1.2730	0.3969	1.1166	1.2428
ρ_{13}	0.5078	0.5089	0.5499	0.5113	0.4922	0.5522
MC s.e.	0.0189	0.0229	0.0105	0.0286	0.0296	0.0085
SIF	0.8625	0.9588	0.6771	1.5033	0.9770	0.6370
ρ_{23}	0.7162	0.7158	0.7366	0.7165	0.7045	0.7372
MC s.e.	0.0209	0.0067	0.0137	0.0085	0.0129	0.0092
SIF	0.8581	0.7418	1.1969	0.3875	0.8620	0.6073

2.3.1 Portfolio allocation along realized paths (fixed horizon)

This first experiment aims at determining the effect of extremal tail dependence on portfolio choice along the realized trajectories of the state variables. Keeping the horizon fixed, we obtain optimal portfolio weights by simulating the remaining elements of the system $\left(Y_s^i, \mathcal{D}_t Y_s^i, \xi_{t,s}^i, H_{t,s}^{IR,i}, H_{t,s}^{MPR,i}\right)_{i=1}^{MC}$. As the vector of market prices of risk is unobservable, we filter it from the data by simulating additional data points between each pair of observations, while keeping the parameters fixed at their posterior means, and then integrating out the latent data points over the simulated Monte Carlo trajectories. The optimal portfolio shares are obtained for a CRRA investor with levels of relative risk aversion of 5 and 10, and are recorded weekly. Table 2.3.3 reports the three components of the optimal portfolio of the investor: the intertemporal hedging terms (against stochastic changes in the market price of risk and the interest rate) and the mean-variance term for varying investment horizons and for all of the alternative diffusions considered. Summary statistics for the optimal portfolio shares for each individual asset are given in Table 2.3.4.

For all of the selected horizons, the extreme value mixture diffusions lead to lower market price of risk hedging demands for the risky assets, thus shifting the portfolio allocation to the riskless asset when the possibility of increased correlation during extreme down markets is accounted for. Those demands increase in absolute terms with the coefficient of relative risk aversion. The mean-variance and the interest rate hedging terms do not show so much disparity between the alternative specifications, so the differences in the total risky demand are driven primarily by the need to hedge changes in the market prices of risk for the different data generating processes. This is clearly seen from Figure 2.3.1, which traces the different decompositions of portfolio terms for the whole investment horizon under the assumption of Gaussian – extreme value copula diffusion (they display a similar pattern for all of the alternative processes considered).

The market price of risk hedges show considerable variations along the sample path, while the interest rate hedges are stable and decline steadily as the horizon decreases, due to the fixed maturity effect (Figure 2.3.1). The market price of risk hedge terms switch signs throughout the period, and determine to a great extent the variations of the total portfolio holdings in the risky assets, as displayed on the upper right panel of Figure 2.3.1, which contrasts the mean-variance component to the total risky asset demand. The two

Table 2.3.3: Portfolio hedging terms along realized paths of the state variables (CRRA investor)

Market timing: portfolio shares and hedging components along realized trajectories of asset prices and the interest rate for the whole 15 year estimation horizon. Gaussian (Ga), Gaussian-Gumbel mix (Large and Mid caps being the most deeply nested couple, Ga-G-SG), nonnested Gaussian-Gumbel mix (nonnested Ga-G-SG), Student's t (T), Student's t – Gumbel mix (Large and Mid caps being the most deeply nested couple, T-G-SG), and nonnested Student's t – Survival Gumbel (T-nonnested SG) diffusions, CRRA investor.

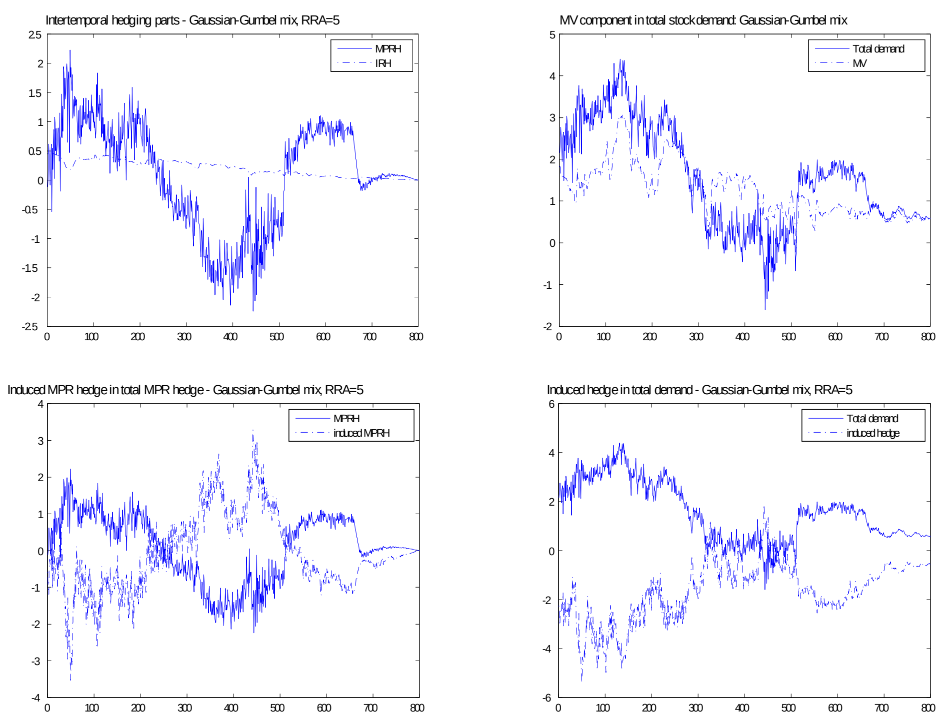
Horizon	Relative risk aversion $R = 5$						Relative risk aversion $R = 10$					
(years)	Ga-G-SG	T	Ga	Nonnested Ga-G-SG	T-G-SG	Nonnested T-SG	Ga-G-SG	T	Ga	Nonnested Ga-G-SG	T-G-SG	Nonnested T-SG
Market price of risk hedge												
14	0.6150	1.6649	1.7138	0.4644	0.1289	2.5075	0.6581	1.7676	1.8897	0.4418	0.1253	2.7037
12	0.9215	1.3911	0.9860	1.0300	-0.1559	1.8179	1.0571	1.4682	1.0494	1.1969	-0.1981	2.1337
10	-0.6778	0.2402	-0.0003	0.0415	-0.7918	0.2804	-0.8938	0.3298	-0.0798	0.0098	-1.0030	0.3077
8	-1.6192	-0.8384	-0.7031	-1.5765	-1.6580	0.0352	-2.2082	-1.1583	-1.0038	-2.3284	-2.1638	-0.0549
6	-1.3031	-0.3589	0.0916	-0.7041	-0.7553	-0.1816	-2.0187	-0.4423	0.2611	-0.9949	-1.2770	-0.1370
4	0.9671	0.8060	1.0838	1.1186	0.8790	0.5221	1.6489	1.2828	1.6307	1.8764	1.6023	0.8897
2	-0.0134	0.3954	0.2622	0.2457	0.2129	0.3342	-0.1001	0.5521	0.3287	0.3551	0.2722	0.6132
Interest rate hedge												
14	0.3777	0.3767	0.3773	0.3624	0.3829	0.3569	0.4353	0.4335	0.4352	0.4202	0.4409	0.4172
12	0.3233	0.3177	0.3241	0.3083	0.3247	0.2934	0.3931	0.3884	0.3937	0.3773	0.3965	0.3683
10	0.3217	0.3152	0.3180	0.3067	0.3128	0.2781	0.4149	0.4099	0.4127	0.4005	0.4109	0.3813
8	0.2468	0.2447	0.2576	0.2379	0.2332	0.2117	0.3458	0.3444	0.3533	0.3390	0.3364	0.3197
6	0.1267	0.1325	0.1375	0.1269	0.1190	0.1095	0.2013	0.2053	0.2097	0.1989	0.1963	0.1848
4	0.0647	0.0688	0.0767	0.0657	0.0595	0.0574	0.1159	0.1189	0.1263	0.1135	0.1123	0.1062
2	0.0383	0.0433	0.0453	0.0355	0.0331	0.0290	0.0807	0.0856	0.0878	0.0766	0.0754	0.0693
Mean-variance term												
14	1.7098	1.9427	1.9027	1.3938	1.8115	0.8510	0.8758	0.9937	0.9753	0.7184	0.9269	0.4422
12	1.7205	1.7983	1.7854	2.0215	1.5983	2.0899	0.9297	0.9772	0.9640	1.0996	0.8674	1.1660
10	1.6490	1.5676	1.6616	1.7058	1.5539	1.2943	0.9453	0.9059	0.9583	0.9899	0.9072	0.7889
8	1.6096	1.4219	1.5017	1.4357	1.5983	1.3058	1.0023	0.8894	0.9154	0.9094	1.0244	0.8765
6	0.6582	0.4871	0.7481	0.6013	0.4198	0.4253	0.4648	0.3355	0.5070	0.4187	0.3079	0.3192
4	0.8303	0.7831	0.9647	0.7233	0.7896	0.6128	0.6607	0.6011	0.7058	0.5556	0.6621	0.5037
2	0.6325	0.7652	0.7250	0.5963	0.5578	0.5437	0.5922	0.6722	0.6249	0.5720	0.5655	0.5768

Table 2.3.4: Portfolio hedging terms along realized paths of the state variables: summary statistics
Summary statistics of the optimal portfolio shares and the intertemporal hedging components along realized trajectories of asset prices and the interest rate for the whole 15 year estimation horizon. Gaussian (Ga), Gaussian-Gumbel mix (Large and Mid caps being the most deeply nested couple, Ga-G-SG) and Student's t (T) diffusions, CRRA investor

	$RRA = 5$				$RRA = 10$			
	Large caps	Mid caps	Small caps	Sum risky assets	Large caps	Mid caps	Small caps	Sum risky assets
Gaussian								
Max	0.2488	2.8125	2.1973	3.1028	0.3966	3.5136	2.1276	3.6109
75% quantile	-0.3558	1.4868	0.3613	1.1334	-0.5391	1.9672	0.4341	1.5575
Median	-0.6367	1.1019	0.1258	0.3882	-0.8358	1.4933	0.1508	0.4586
25% quantile	-0.8907	0.5230	-0.2191	-0.0564	-1.0825	0.7340	-0.3508	-0.1224
Min	-1.5364	-0.4043	-1.3360	-1.1610	-2.0323	-0.7152	-1.9642	-1.7680
Gaussian – Gumbel – Survival Gumbel								
Max	-0.0004	5.5814	0.1646	2.2215	0.0854	6.2170	0.2516	2.3343
75% quantile	-0.4158	3.0379	-0.1451	0.8131	-0.6514	3.5943	-0.2587	1.1015
Median	-0.8004	1.9596	-0.9229	0.1005	-1.0518	2.7644	-1.2379	0.1334
25% quantile	-1.2788	0.9521	-1.7777	-0.5435	-1.5080	1.4465	-2.2372	-0.7413
Min	-2.4428	-0.4193	-3.4611	-2.2433	-2.8499	-0.8255	-4.6594	-3.4652
Students's t								
Max	0.2529	8.3955	0.3195	3.3301	0.5339	9.5037	0.5407	3.6088
75% quantile	0.0089	4.6416	0.0122	1.0719	0.0040	5.2310	0.0096	1.4244
Median	-0.7103	2.1024	-0.8540	0.4323	-1.0272	2.7634	-1.0882	0.6022
25% quantile	-2.3891	0.6970	-1.4890	-0.1877	-2.6797	1.0862	-1.8337	-0.2601
Min	-4.6624	0.0014	-3.3162	-1.3549	-5.2222	0.0015	-4.5054	-1.9645

Figure 2.3.1: Portfolio hedging terms along realized paths of the state variables

The figures represent the evolution of the hedging and mean-variance terms along realized trajectories of the state variables for the whole 15 year estimation horizon for the nested Gaussian-Gumbel-Survival Gumbel diffusion (Large and Mid caps being the most deeply nested couple). On the horizontal axes time increments are weekly. The top left figure plots the intertemporal hedging terms (MPRH stands for market price of risk hedge, IRH stands for interest rate hedge). The top right figure presents the mean-variance (MV) component as a part of the total asset demand. The two bottom figures plot the induced hedging demands: the left one plots the sum of the induced MPRH terms vs. the total MPRH, while the one on the right plots the sum of the total induced hedging demand vs. the total demand for the risky assets.



lower panels display the evolution of the hedging demands throughout the period that are induced by the positions in the rest of the assets. Induced hedges are considerable in magnitude and are opposite in sign with respect to the total hedging demand.

An investor who uses the extreme value mixture diffusion as a data generating process consistently underinvests in the risky assets as compared to an investor who believes that log-prices are driven by a tail independent Gaussian process. Thus disregarding the effect of extreme dependence in the tails leads to increased portfolio holdings in the risky assets for the most part of the 15-year period we consider, as seen on Figure 2.3.2, which compares the two elliptical models with their extreme-value mixture counterparts. However, the Student's t model performs almost identically as the Gaussian, the intertemporal market price of risk hedging terms being virtually indistinguishable for most of the investment horizon. Figure 2.3.3 illustrates the impact of considering the more richly parametrized nested version of the Archimedean copulas. The Gaussian mixtures do not display a significant change in the hedging demands, while for the Student's t mixtures the nested version leads to substantially lower demand for the risky assets.

So far we have analyzed the behavior of the portfolio hedging terms for the entire estimation period. As it would be of greater interest to contrast periods of increased frequency of tail events to considerably calm periods, we look at three subsamples: the period of 1992-1995, characterized with low volatility and no tail events, and the periods of 1997-2000 and 2001-2005, during which there were several market crashes, and asset return volatility was substantially higher, as it can be seen from Figure 2.3.4, on which we have plotted the GARCH volatility estimates for each of the return series. The last two subperiods thus include the October 1997 crash caused by the economic crisis in Asia, and the bear market in 2002, related to the 'Internet bubble'.

For the intertemporal hedging terms we consider three competing diffusion specifications: the tail independent Gaussian benchmark, the symmetric tails Student's t diffusion and the specification based on the extreme-value copula mixture of Gaussian and Gumbel copulas in its nested version. The investment horizon is kept fixed at the end of each period, so that the hedging terms decline towards zero with the approach of the terminal date for each subperiod. Figure 2.3.5 displays the market price of risk hedges, as well as the mean-variance terms for the competing data generating processes, under the assumption of a

Figure 2.3.2: Portfolio hedging terms along realized paths of the state variables

Market price of risk hedge terms along realized trajectories of the state variables for the whole 15 year estimation horizon (left column) and for a rolling window horizon of 5 years (right column). On the horizontal axes time increments are weekly. Plotted are the market price of risk hedges of the Gaussian – extreme value mixture diffusion (Large and Mid caps being the most deeply nested couple, Ga-GSG) vs. the Gaussian diffusion (Ga), and of the Student's t – extreme value diffusion (T-G-SG) vs. the Student's t diffusion (T). In all cases we have a CRRA investor with relative risk aversion of 5.

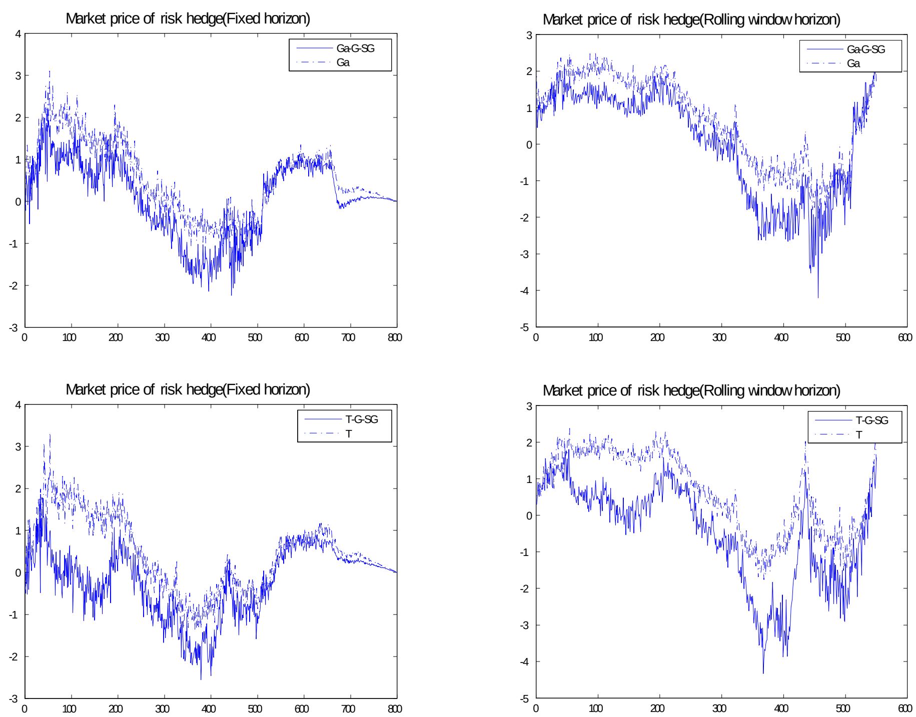


Figure 2.3.3: Market price of risk hedging terms along realized paths for the state variables: nested vs. nonnested Gaussian - Gumbel - Survival Gumbel copulas; nested Student's t Gumbel - Survival Gumbel vs. Student's t - nonnested Survival Gumbel

The sums of the market price of risk hedging terms for the three risky assets are reported for the whole 15 years estimation period for a rolling window horizon of five years. Hedging terms are reported every 10 days. Plotted are the hedging components for the nonnested Gaussian-Gumbel-Survival Gumbel specification against two nested alternatives (the most deeply nested couple for each case is given in parenthesis), as well as the hedging terms for the Student's t diffusion (t) against the extreme value mixture alternative represented by the Student's t – nonnested Survival Gumbel copula (t-SG). CRRA investor with a relative risk aversion parameter (RRA) of 5.

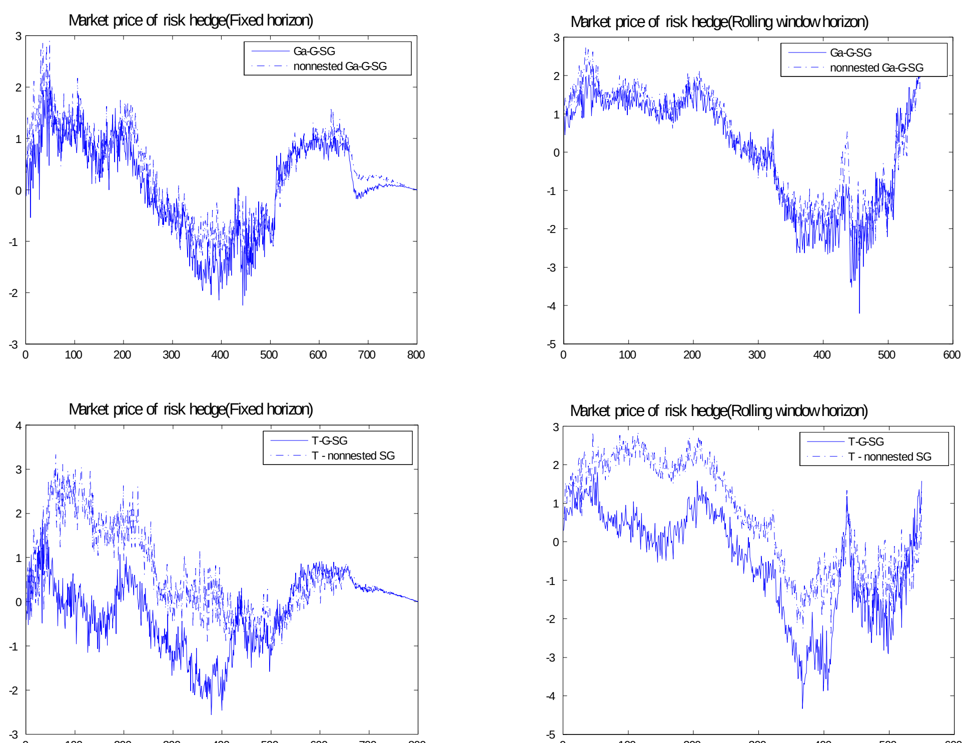
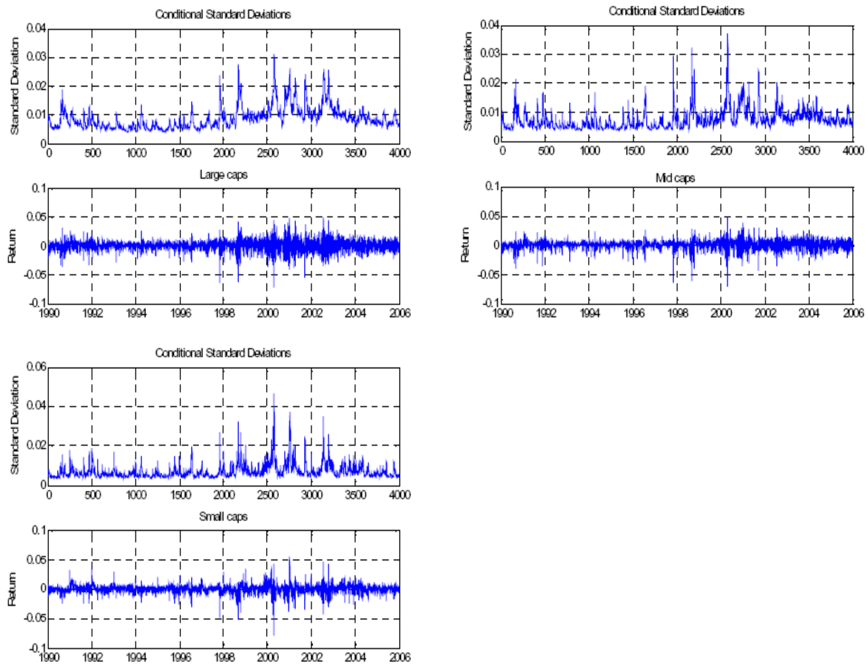


Figure 2.3.4: GARCH(1,1) volatility estimates for the three asset return series

Plotted are the GARCH(1,1) estimates of the conditional standard deviations of the three return series (Large, Mid and Small caps) for the whole estimation period 1990-2005.



CRRA investor with a coefficient of relative risk aversion equal to 5. We have conducted the same allocation experiment for a HARA investor as well, but the results are qualitatively similar to the CRRA case.

The two elliptical diffusions render similar hedging terms for the three subperiods, while the extreme-value copula reduces considerably the demand for the risky asset. That is true even for the first subperiod with no tail events and low volatility, even though in the second part of the period all hedging demands are very close to each other, which is not the case for the 2001-2005 period, marked with more extreme events. The mean-variance terms are almost identical across the alternative specifications, so the difference in risky asset demands comes almost exclusively from the market price of risk hedges.

In order to gather more insight into the impact of considering lower tail dependence during a 'bear' market as compared to a period with no extreme events, we look at the evolution of wealth generated by the portfolio allocation decisions for the alternative data generating processes for each of the three subperiods. Optimal wealth at time t is given by $\omega_t^* = E_t\left[\xi_{t,T} I\left(y^*\xi_T\right)\right]$, which for a CRRA investor reduces to $\omega_t^* = E_t\left[\xi_{t,T}\left(y^*\xi_T\right)^{-1/R}\right]$.

Figure 2.3.5: Intertemporal hedging terms and mean-variance demand for the three subperiods: 1992-1995, 1997-2000, 2001-2005

Plotted are Intertemporal Market Price of Risk hedge terms and Mean Variance terms for the three subsample periods (the first line corresponds to 1992-1992, the second – to 1997-2000, and the third – to 2001-2005). The investment horizon is kept fixed at the end of each period. The assumptions on investor's preferences are CRRA utility with a coefficient of relative risk aversion of 5. The data generating processes considered are Gaussian diffusion (Ga), Student's t diffusion (Student t), and nested Gaussian-Gumbel-Survival Gumbel diffusion with Big-Mid caps being the most deeply nested couple (Ga-G-SG).

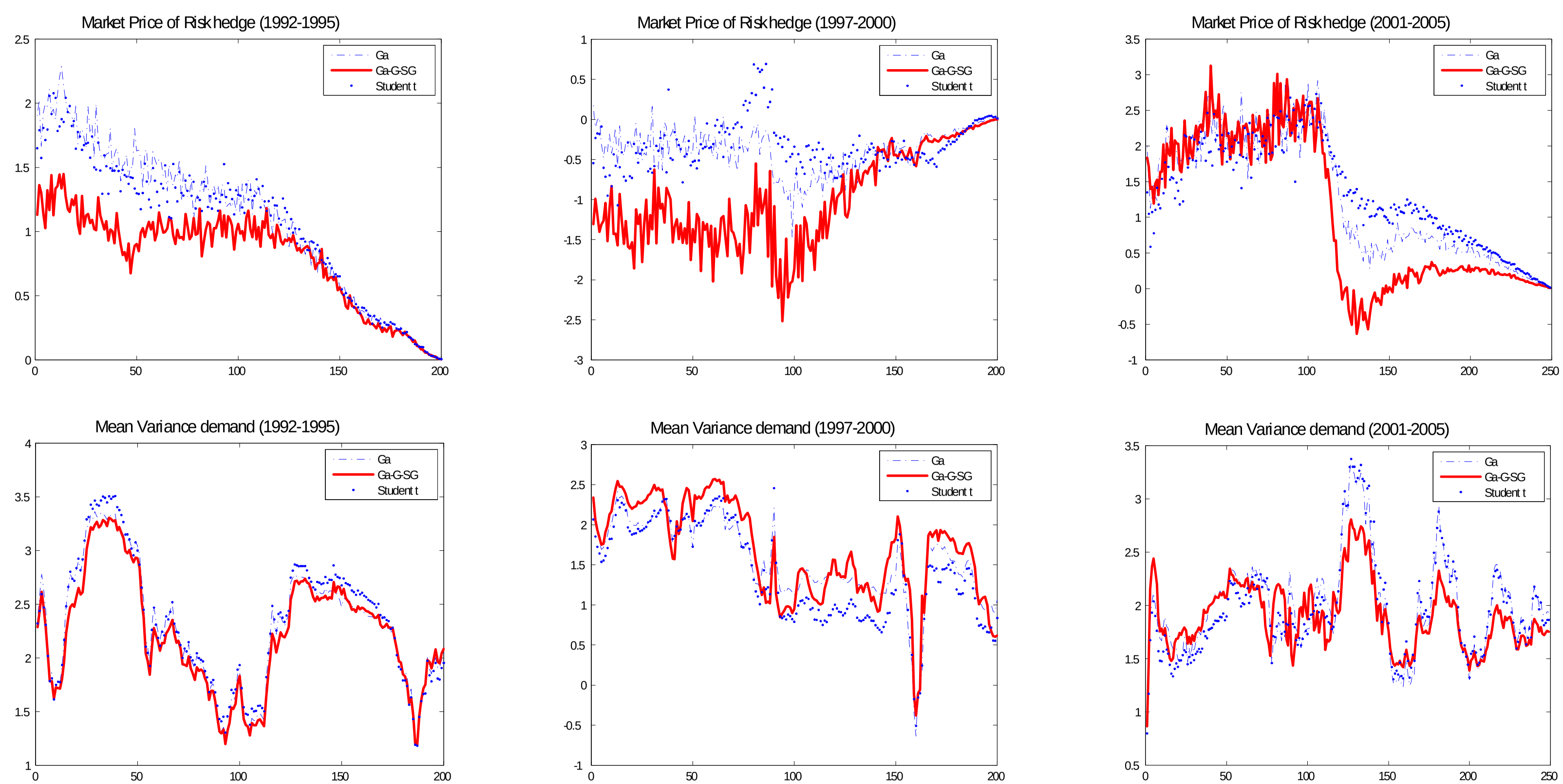


Figure 2.3.6: Evolution of wealth for the three subperiods: 1992-1995, 1997-2000, 2001-2005

Plotted is the wealth evolution for the three subsample periods (the first line corresponds to 1992-1992, the second – to 1997-2000, and the third – to 2001-2005). The investment horizon is kept fixed at the end of each period. The assumptions on investor's preferences are CRRA utility with a coefficient of relative risk aversion of 5. The data generating processes considered are Gaussian diffusion (Ga), Student's t diffusion (Student t), and nested Gaussian-Gumbel-Survival Gumbel diffusion with Big-Mid caps being the most deeply nested couple (Ga-G-SG).

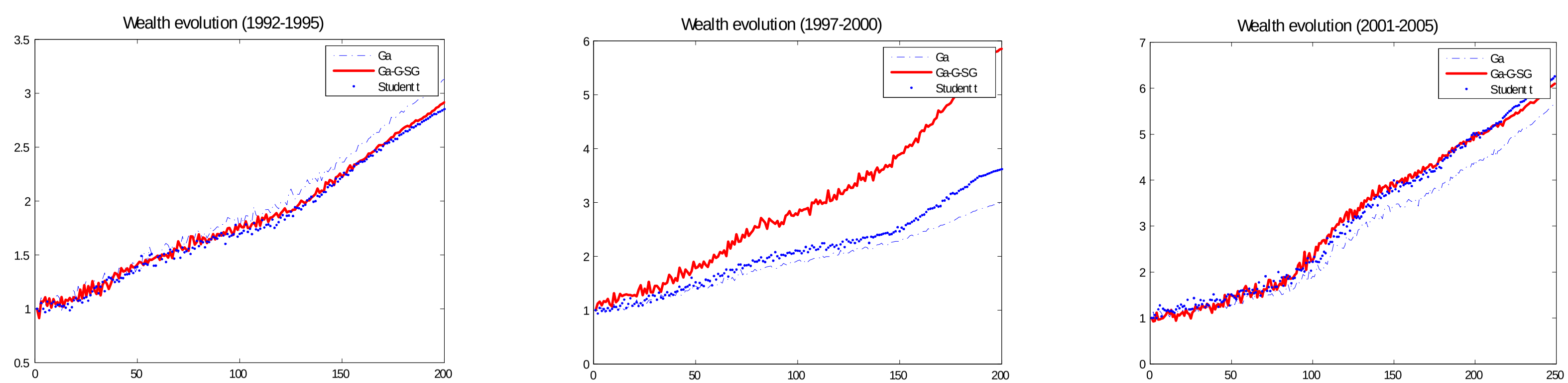


Figure 2.3.6 illustrates the optimal wealth for an investor with CRRA preferences for a coefficient of risk aversion of 5. For the 1992-1995 period all three diffusion specifications render similar wealth growth. Thus, even though the intertemporal hedging terms for the risky assets are lower when the investor takes into account the dependence between extreme low returns, it does not translate into a change in her wealth evolution for this calm period with no extreme events. To the contrary, the loss in terms of wealth for the subsequent period of 1997-2000 is substantial for both elliptic specifications. The costs of ignoring tail dependence will be further considered in more detail through the concept of the certainty equivalent cost in the next sections.

2.3.2 Portfolio allocation along realized paths (rolling window horizon)

In order to examine the effect of a varying horizon, we perform a second experiment along the realized trajectories of the state variables, in which the only difference with respect to the above mentioned exercise is the fact that we keep the horizon fixed at 5 years. Table 2.3.5 reports results for the hedging terms and the mean-variance terms of optimal portfolios, recorded each week, for periods of 14 to 6 years before the end of the sample horizon (not to be confused with the investment horizon of 5 years for each allocation decision).

We again confirm the previous finding that the extreme value diffusion model of the nested Archimedean mixture leads to considerably lower hedging demands for the risky assets as compared to the asymptotically independent Gaussian model. Overall, as seen from the right column of Figure 2.3.1, the pattern of the market price of risk hedges for a rolling window horizon remains similar to that of the fixed horizon case, and its volatility again determines to a great extent the total portfolio allocations in the risky assets.

2.3.3 Portfolio allocation along simulated paths

Having established the impact of modeling extreme value dependence on portfolio terms using actual data, we now turn to the case of simulating ahead the whole system of state variables, Malliavin derivatives and hedging terms $\left(Y_s^i, \mathcal{D}_t Y_s^i, \xi_{t,s}^i, H_{t,s}^{IR,i}, H_{t,s}^{MPR,i}\right)_{i=1}^{MC}$ for any of the alternative diffusion specifications for a CRRA and a HARA investor with levels of relative risk aversion ranging between 2 and 10, for a horizon of up to 3 years. For the HARA utility assumption, we consider the case of $B = -0.2$, and unit initial wealth, for

Table 2.3.5: Intertemporal hedging demands for rolling window horizon of five years

Market timing: portfolio shares and hedging components along realized trajectories of asset prices and the interest rate for the whole 15 year estimation period with a fixed horizon of 5 years (moving window horizon). Gaussian (Ga), Gaussian-Gumbel mix (Large and Mid caps being the most deeply nested couple, Ga-G-SG) and Student's t (T) diffusions, CRRA investor

Horizon	Relative risk aversion $R = 5$						Relative risk aversion $R = 10$					
(years)	Ga-G-SG	T	Ga	Ga-G-SG	T-G-SG	T-SG	Ga-G-SG	T	Ga	Ga-G-SG	T-G-SG	T-SG
				Nonnested		Nonnested				Nonnested		Nonnested
Market price of risk hedge												
14	1.3583	1.9149	1.9713	1.3773	0.6180	1.8346	1.3412	1.9262	1.9395	1.2905	0.6286	1.7972
12	1.0323	1.4639	1.5521	1.3703	0.5283	1.7410	0.9562	1.3854	1.4711	1.3489	0.6045	1.7212
10	0.1429	0.5494	0.8332	0.0374	-0.3418	0.4433	0.0252	0.4796	0.7466	-0.0515	-0.3650	0.3620
8	-2.5122	-1.1788	-0.9858	-1.4259	-2.8141	-1.4403	-2.8621	-1.3222	-1.0693	-1.4465	-3.1367	-1.7442
6	-1.2140	-0.7662	-1.0978	-1.3323	-1.5726	-0.4250	-1.2882	-0.8205	-1.2345	-1.4775	-1.5879	-0.4199
Interest rate hedge												
14	0.2650	0.2635	0.2651	0.2574	0.2681	0.2577	0.2981	0.2965	0.2983	0.2896	0.3016	0.2899
12	0.2601	0.2585	0.2604	0.2514	0.2636	0.2516	0.2927	0.2908	0.2929	0.2828	0.2965	0.2831
10	0.3054	0.3042	0.3056	0.2984	0.3080	0.2985	0.3436	0.3422	0.3438	0.3357	0.3465	0.3358
8	0.2985	0.2986	0.2985	0.2976	0.2988	0.2974	0.3358	0.3359	0.3359	0.3348	0.3362	0.3346
6	0.2240	0.2227	0.2239	0.2181	0.2268	0.2185	0.2519	0.2505	0.2519	0.2454	0.2552	0.2459
Mean-variance term												
14	1.7942	2.0332	1.9998	1.4809	1.8972	0.9190	0.8971	1.0166	0.9999	0.7405	0.9486	0.4595
12	2.0093	2.1239	2.0821	2.3927	1.8828	2.6020	1.0046	1.0620	1.0411	1.1963	0.9414	1.3010
10	2.1675	2.0943	2.2107	2.2977	2.1183	1.9235	1.0837	1.0471	1.1054	1.1489	1.0592	0.9617
8	2.4966	2.2255	2.2323	2.3043	2.6265	2.3531	1.2483	1.1128	1.1162	1.1522	1.3133	1.1766
6	1.3128	0.9241	1.3745	1.1664	0.9031	0.9581	0.6564	0.4620	0.6872	0.5832	0.4515	0.4790

which case the investor displays aversion towards wealth shortfalls below a certain threshold. Results are illustrated on Figure 2.3.7 and 2.3.8 for the intertemporal hedging terms and the considered diffusion specifications.

The interest rate hedge, very close across all specifications, is positive and increases with the horizon, as well as with the level of relative risk aversion. We find the opposite behavior for the market price of risk hedging term that decreases with the coefficient of relative risk aversion. For all horizons considered and for all levels of risk aversion the nested extreme value mixture diffusion both in the Gaussian and the Student's t cases induces lower intertemporal demand for the risky assets, compared with its two elliptical counterparts. This effect is accentuated in the case of a HARA investor, who has lower MPR hedging demands and higher IR hedging demands for all cases.

As in all of the above experiments, for either observed or simulated paths of the state variables, the differences in the total asset demand across alternative data generating processes are driven mainly by the Market price of risk hedging demands, we conduct a simple simulation study in order to reveal the sensitivity of the market price of risk hedging terms to changes in the parameters that describe the dependence structure. To this end, we solve for the optimal portfolio when the data generating process is the most parsimonious extreme-value diffusion – the nonnested Gaussian-Gumbel-Survival Gumbel, by simulating ahead the state variables and their Malliavin derivatives for a horizon of 6 and 12 months for changing values of the parameters that determine the weights of the extreme value copulas (ω^G and ω^{SG}). Thus for $\omega^G + \omega^{SG} = 0$ we obtain the Gaussian copula with no tail dependence, while for $\omega^G + \omega^{SG} = 1$ dependence is driven entirely by the extreme value Archimedean copulas, and asymptotic tail dependence is at its highest values. Table 2.3.6 summarizes the results of this comparative statics experiment.

When the part of the extreme-value copulas increases versus that of the Gaussian copula, the market price of risk hedging terms decrease, reflecting the higher risk of joint occurrence of tail events.

2.3.4 The cost of ignoring extremal dependence

Having thus obtained the optimal portfolio shares, we proceed to the assessment of the importance of the intertemporal hedging demands, induced by the asymmetric dependence

Figure 2.3.7: Intertemporal hedging terms for simulated paths of the state variables: Gaussian vs. Extreme value mixture diffusion

Intertemporal hedging terms (Market price of risk, Interest rate and Induced MPRH hedging terms) for the simulation experiment for a CRRA investor and a HARA (B=-0.2) investor with levels of relative risk aversion ranging between 2 and 10, for a horizon of 1 and 3 years. The hedging terms are plotted for 2 alternative models: Gaussian (Ga), and nested Gaussian – Gumbel – Survival Gumbel (Ga-G-SG, Large and Mid caps being the most deeply nested couple). Parameters for the alternative models are taken from Table 2.3.2.

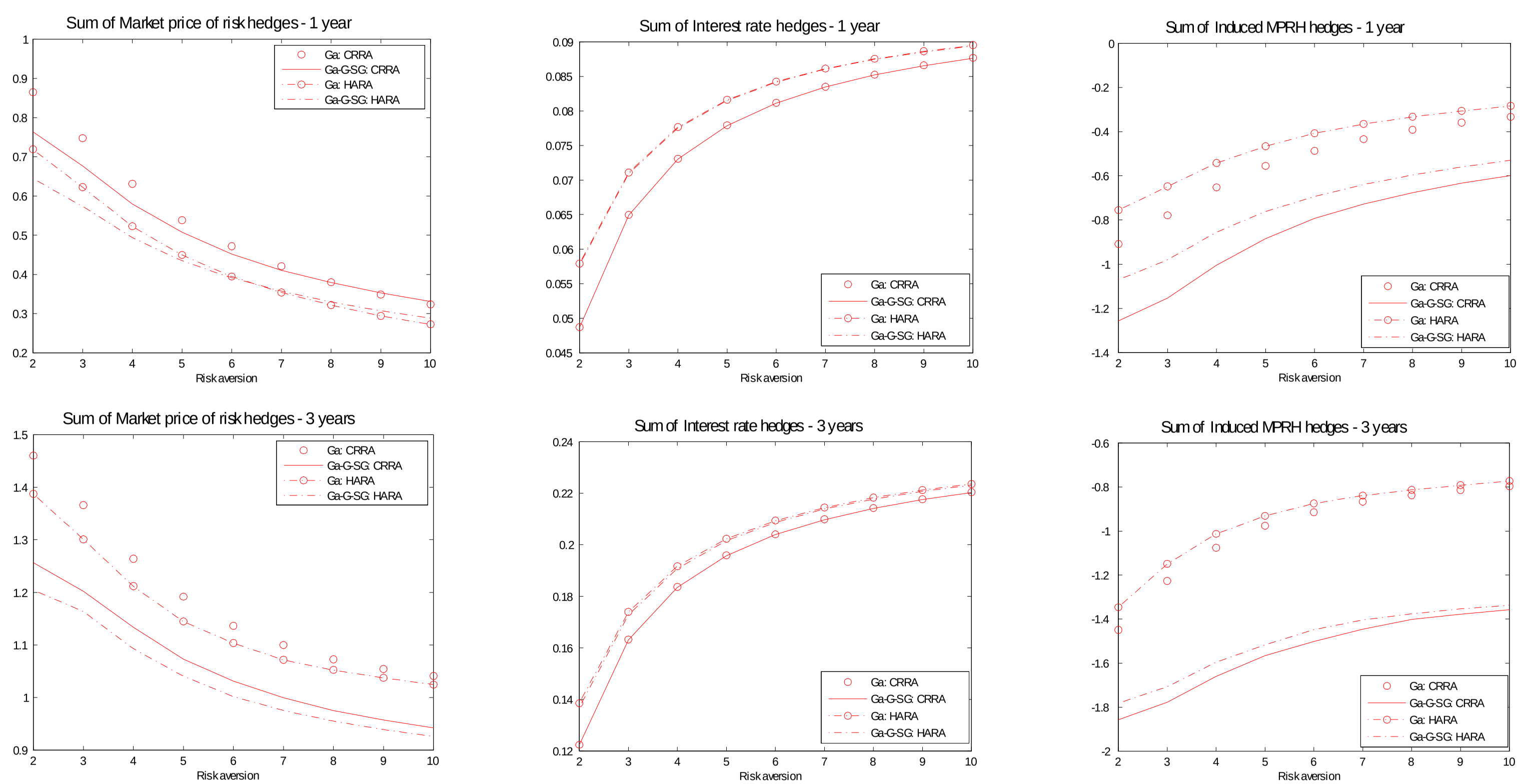


Figure 2.3.8: Intertemporal hedging terms for simulated paths of the state variables: Student's t vs. Extreme value mixture diffusion

Intertemporal hedging terms (Market price of risk, Interest rate and Induced MPRH hedging terms) for the simulation experiment for a CRRA investor and a HARA (B=-0.2) investor with levels of relative risk aversion ranging between 2 and 10, for a horizon of 1 and 3 years. The hedging terms are plotted for 2 alternative models: Student's t, and nested Student's t – Gumbel – Survival Gumbel (Large and Mid caps being the most deeply nested couple). Parameters for the alternative models are taken from Table 2.3.2.

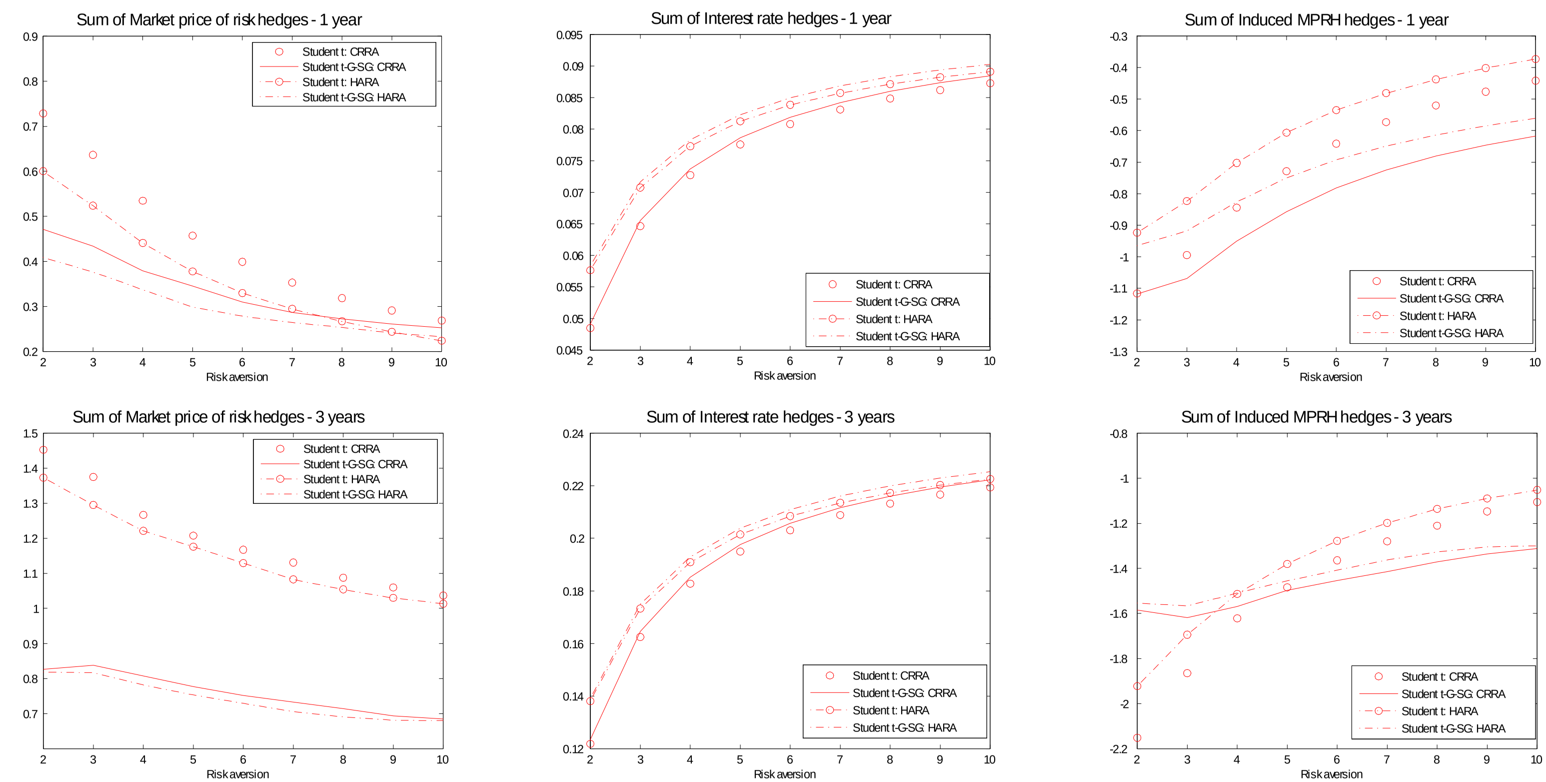


Table 2.3.6: Sensitivity of the Market price of risk hedging term to changes in the parameters describing the dependence structure

The table reports the sum of the Market price of risk hedging terms for varying parameters that describe the dependence structure for a fixed horizon of 6 and 12 months and a CRRA investor with coefficients of relative risk aversion of 2, 5, 10 and 20. The model considered is the nonnested Gaussian-Gumbel-Survival Gumbel model for changing values of the parameters that determine the weights of the mixture copula ($\omega^G, \overline{\omega}^G$).

	T = 6 months				T = 12 months			
$RRA:$	2	5	10	20	2	5	10	20
$\omega^G + \overline{\omega}^G$	(assuming equal weights for the Gumbel and the Survival Gumbel copulas)							
0	0.4604	0.3792	0.2979	0.2478	0.7668	0.6318	0.5039	0.4266
0.2	0.3865	0.3257	0.2642	0.2264	0.6609	0.5724	0.4829	0.4290
0.4	0.3098	0.2499	0.1974	0.1659	0.5556	0.4700	0.3923	0.3464
0.6	0.2411	0.1666	0.1144	0.0840	0.4443	0.3353	0.2580	0.2137
0.8	0.2296	0.1830	0.1488	0.1292	0.4379	0.3701	0.3162	0.2855
1	0.1915	0.1447	0.1126	0.0944	0.3851	0.3457	0.3124	0.2939

structure, in terms of the cost of ignoring these asymmetries. To this end, we propose to follow the approach, largely exploited in literature on portfolio allocation that uses as criterion the utility cost of ignoring the particular data structure in terms of the certainty equivalent (i.e. in the present case, the utility cost of using non-optimal weights that come from a data generating process that assumes tail independence, whereas data is characterized by asymmetries in the extremal dependence structure).

In order to find the additional wealth that is initially required by an investor to use a suboptimal allocation strategy for a given horizon, we have to solve the following equation for the value functions that have to be equal at the investment horizon:

$$E_0\left[u\left(\omega_T^* \mid \omega_0 = 1\right)\right] = E_0\left[u\left(\omega_T \mid \omega_0 = \overline{\omega}\right)\right] \tag{2.3.1}$$

where ω_T^* is the optimal terminal wealth obtained under the optimal data generating process, ω_T is the terminal wealth obtained by using an alternative (suboptimal) data generating process for the state variables, and $\overline{\omega}$ is the initial wealth required by the investor in order to form suboptimal portfolio shares. In the case of a CRRA utility, that is homogenous in initial wealth, the above equation simplifies to:

$$E_0\left[\left(\omega_T^*\right)^{1-R} \mid \omega_0 = 1\right] = E_0\left[\left(\omega_T\right)^{1-R} \mid \omega_0 = \overline{\omega}\right] \tag{2.3.2}$$

Table 2.3.7: The cost of ignoring extreme dependence as modeled by the extreme value mixture diffusion (Gaussian – Gumbel – Survival Gumbel)

The certainty equivalent costs (dollars) for using a suboptimal allocation strategy of assuming that state variables are driven by a Gaussian diffusion, a Student's t diffusion, or a nonnested Gaussian – Gumbel – Survival Gumbel diffusion while the true data generating process is the nested Gaussian – Gumbel – Survival Gumbel diffusion. Results for a CRRA and a HARA investor with levels of relative risk aversion of 5 and 10, for a horizon of 6 months to 3 years. For the HARA utility $B = -0.1$ or $B = 0.1$, and initial wealth is set to 1.

	HARA ($B=-0.1$)		CRRA		HARA ($B=0.1$)	
Horizon	$RRA=5$	$RRA=10$	$RRA=5$	$RRA=10$	$RRA=5$	$RRA=10$
vs. Gaussian						
6 months	0.0319	0.0198	0.0344	0.0210	0.0370	0.0223
1 year	0.0609	0.0393	0.0650	0.0414	0.0692	0.0434
2 years	0.1030	0.0672	0.1085	0.0698	0.1139	0.0725
3 years	0.1584	0.1129	0.1635	0.1153	0.1685	0.1178
vs. Student's t						
6 months	0.0209	0.0136	0.0224	0.0144	0.0240	0.0152
1 year	0.0414	0.0294	0.0438	0.0306	0.0462	0.0318
2 years	0.0803	0.0631	0.0830	0.0645	0.0857	0.0659
3 years	0.1365	0.1197	0.1384	0.1207	0.1404	0.1217
vs. nonnested Gaussian – Gumbel – Survival Gumbel						
6 months	0.0263	0.0154	0.0286	0.0166	0.0309	0.0177
1 year	0.0432	0.0241	0.0470	0.0259	0.0507	0.0278
2 years	0.0819	0.0525	0.0865	0.0547	0.0910	0.0569
3 years	0.1438	0.1100	0.1476	0.1118	0.1514	0.1137

The inverse of the marginal utility function for a CRRA investor has the form $I(z) = z^{-\frac{1}{R}}$, so that for $\overline{\omega}$ we obtain $\overline{\omega} = \left\{ E_0\left[(\xi_T^*)^{1-\frac{1}{R}} \right] / E_0\left[(\xi_T)^{1-\frac{1}{R}} \right] \right\}^{\frac{R}{1-R}}$.

We examine the certainty equivalent costs for two scenarios. First, the true data generating process for the de-trended log-prices is assumed to be the Gaussian – extreme value nested mixture diffusion (2.2.12) and we solve for the utility cost of using either one of the Elliptical diffusions instead or the alternative nonnested mixture model for a horizon of 6 months to 3 years. Results for a CRRA and a HARA investor are reported in Table 2.3.7.

We find significant certainty equivalent costs for choosing a suboptimal data generating process in this case, as high as 16 cents per dollar for the Gaussian copula and the longest horizon. The costs increase with the investment horizon, and there are no significant qualitative differences between a HARA and a CRRA investor. The two HARA cases considered with a positive or negative parameter B provide an upper and a lower limit for the certainty equivalent cost of the benchmark CRRA investor. One loses the most if using a Gaussian

Table 2.3.8: The cost of ignoring extreme dependence as modeled by the extreme value mixture diffusion (Student's t – Gumbel – Survival Gumbel)

The certainty equivalent costs (in dollars) for using a suboptimal allocation strategy of assuming that state variables are driven by a Gaussian diffusion, a Student's t diffusion, or a Student's t – nonnested Survival Gumbel diffusion while the true data generating process is the nested Student's t – Gumbel – Survival Gumbel diffusion. Results for a CRRA and a HARA investor with levels of relative risk aversion of 5 and 10, for a horizon of 6 months to 3 years. For the HARA utility $B = -0.1$ or $B = 0.1$, and initial wealth is set to 1.

	HARA ($B = -0.1$)		CRRA		HARA ($B = 0.1$)	
Horizon	$RRA = 5$	$RRA = 10$	$RRA = 5$	$RRA = 10$	$RRA = 5$	$RRA = 10$
vs. Gaussian						
6 months	0.0545	0.0344	0.0587	0.0364	0.0629	0.0385
1 year	0.0969	0.0609	0.1039	0.0642	0.1108	0.0676
2 years	0.1987	0.1445	0.2067	0.1483	0.2146	0.1521
3 years	0.3314	0.2621	0.3388	0.2655	0.3461	0.2689
vs. Student's t						
6 months	0.0397	0.0245	0.0429	0.0261	0.0460	0.0276
1 year	0.0690	0.0428	0.0741	0.0453	0.0792	0.0477
2 years	0.1576	0.1230	0.1627	0.1255	0.1679	0.1280
3 years	0.2691	0.2312	0.2732	0.2331	0.2773	0.2351
vs. Student's t – nonnested Survival Gumbel						
6 months	0.0182	0.0100	0.0199	0.0109	0.0549	0.0321
1 year	0.0200	0.0061	0.0227	0.0074	0.0254	0.0088
2 years	0.0293	0.0132	0.0317	0.0144	0.0341	0.0156
3 years	0.0557	0.0422	0.0572	0.0430	0.0587	0.0437

model instead of the true process when allocating a portfolio, and the costs are lower for the two alternative specifications, as they take into account the increased dependence for tail events.

For our second experiment, we generate data from the Student's t – extreme value nested mixture diffusion (2.2.13) and consider suboptimal allocations for its elliptical or nonnested counterparts. Certainty equivalent costs are twice as large, as compared to the previous experiment, and are the most important for the Gaussian data generating process. However, the certainty equivalent costs are not as substantial for nonnested version in this case, probably due to the richer parameter specification of the Student's t diffusion.

The above simulation experiments have the underlying assumption of the true data generating process being an Elliptic – Extreme value mixture, and we simulate the state variables following the assumed process. We could alternatively look at the data itself and compute the certainty equivalent costs along realized paths of the state variables, similarly to the portfolio allocation experiments. We thus treat the paths of the state variables as observed and we assume alternative data generating processes when simulating the state price density. As its evolution depends only on the market prices of risk and the short rate evolution, we do not need to simulate Malliavin derivatives of the state variables in this case. We look at a rolling window horizon of 6 months to 3 years, in order to match the horizons in the simulation experiment, and consider the two Elliptical and the two Elliptic – Extreme value mixture nested diffusions as data generating processes. Figure 2.3.9 displays the paths of the certainty equivalent costs for the 3 rolling window horizons for the following three cases: (i) Gaussian diffusion while the copula underlying the true data generating process is the nested Gaussian – Gumbel – Survival Gumbel; (ii) Student's t diffusion while the copula of the true data generating process is the nested Student's t – Gumbel – Survival Gumbel; and (iii) Gaussian diffusion while data is assumed to come from a Student's t copula.

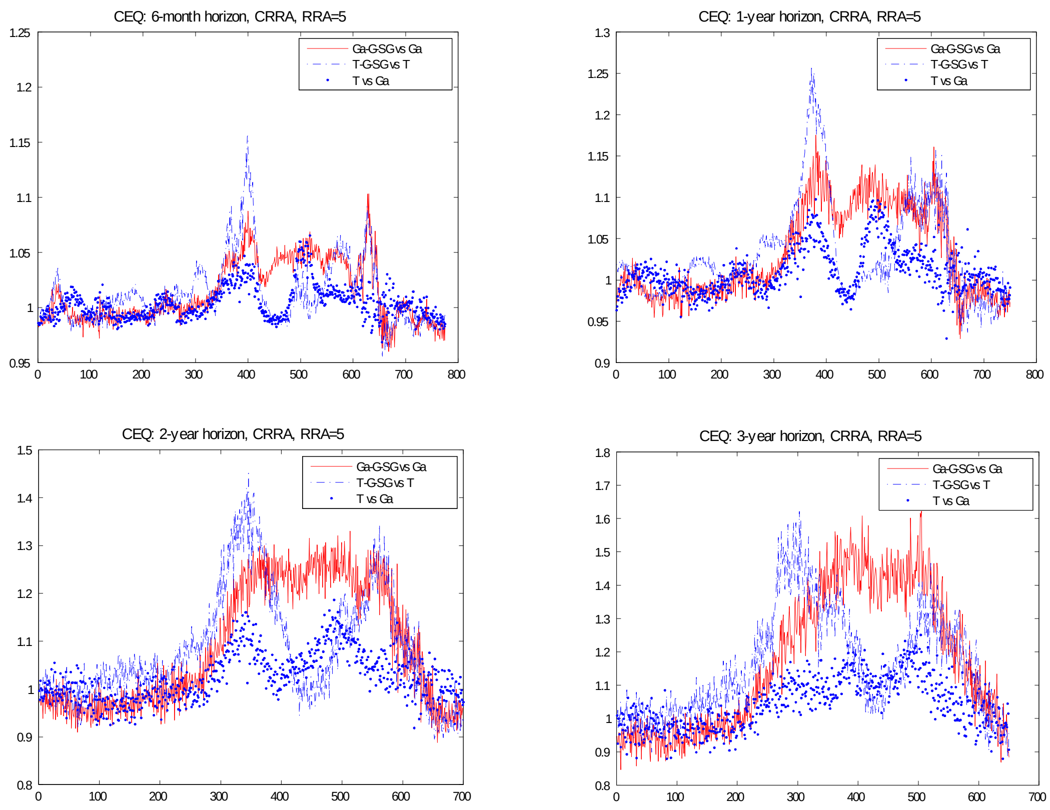
A value of 1 on the vertical axis would mean no certainty equivalent cost, while any value above it translates into a cost of the corresponding value minus 1, in cents per dollar. Alternatively, any value below 1 points to a gain, instead of loss, of using the alternative data generating process. For the first half of the sample and for all horizons the investor loses nothing by choosing a tail independent data generating process, while her costs are quite substantial for the second more volatile period, characterized by several market crashes. The costs increase with the investment horizon and are more pronounced for the case (i), for which they remain at a high level for the second half of the period. The costs for the Student's t case (ii) are higher than those in (i) on several occasions, but often drop to zero, as both diffusions have a certain degree of tail dependence, and only the tail asymmetry in the Extreme value diffusion would drive the differences between the two processes. The investor looses the less in case (iii), where both competing diffusions are in the Elliptic class.

2.4 Conclusion

In this chapter we analyze the importance of considering dependence between extreme realizations of stock market returns on intertemporal portfolio choice. In order to achieve this, we address two problems: develop a parametric model that replicates the extremal dependence found in the data, and apply a solution methodology for portfolio allocation that allows us to isolate the intertemporal hedging terms induced by the particular data

Figure 2.3.9: The cost of ignoring extreme dependence as modeled by the extreme value mixture diffusion: along realized paths of state variables for a rolling 6-month, 1,2, and 3-year horizon

The figures display the certainty equivalent cost (CEQ) of ignoring extreme value dependence across realized paths of the state variables for the whole estimation period and a rolling-window horizon of 6 months 1, 2, and 3 years. This is performed under the assumption of a CRRA investor with a coefficient of relative risk aversion of 5. The alternative data generating processes considered are the nested Gaussian-Gumbel-Survival Gumbel diffusion vs. the Gaussian diffusion, the Student's t nested Gumbel-Survival Gumbel diffusion vs. the Student's t diffusion, and the Student's t diffusion vs. the Gaussian diffusion. A value of 1 indicates no certainty equivalent cost of disregarding the benchmark model; any value above 1 points towards positive certainty equivalent cost in cents per dollar equal to the difference between the plotted value and 1; a value below 1 indicates loss, its magnitude being equal to the difference between the plotted value and 1.



generating process.

The idea of devising a model that is able to incorporate this specific dependence structure has been exploited in the discrete-time literature building upon copula models in a GARCH framework. However, no extension is provided to modeling the spatial dependence structure of a continuous-time stochastic process. To solve this problem, we develop a multivariate diffusion model with a prespecified stationary distribution, based on copula functions, that is able to reproduce the dependence structure of the data. It represents as well a multivariate generalization of the flexible class of univariate Generalized hyperbolic diffusion models of Rydberg (1999) and Bibby and Sorensen (2003), and is thus able to account for stylized features of asset returns as thick tails, skewness in the marginal distribution, and persistence in the autocorrelation of squared returns. The mixture model we propose nests the cases of asymptotic independence and tail dependence, thus covering a large spectrum of extremal dependence structures.

There is conflicting evidence however on the effect of asymmetric correlations on portfolio choice. Portfolio allocation problems in the copula framework have been addressed in literature only in the unconditional context (Patton, 2004), where the effect of dependence asymmetries is found to be substantial, while no extension is proposed for dynamic portfolio selection. In a multiperiod setting on the other hand, under alternative specifications that aim at capturing the same stylized behavior, dependence asymmetries are found to have no considerable effect on optimal portfolio shares (Ang and Bekaert, 2002), while dynamic hedging terms are not explicitly obtained. The model we develop has the advantage that it allows us to approach the dynamic portfolio selection problem within a complete market setup, and obtain explicit expressions for the dynamic hedging demands induced by the extreme value dependence data generating process. The solution methodology is flexible enough to allow for quite general specifications of the state variables and the utility function. We find that taking into account the dependence between realizations of tail events diminishes the intertemporal demand for the risky asset and induces substantial utility cost when this particular dependence structure is ignored.

However, these results are obtained under the assumption of constant conditional correlation. A recent study of Buraschi et al. (2007) has provided evidence of a substantial portfolio hedging component due to correlation risk, which is highly time-varying. Thus, relaxing the constant correlation assumption in our setting would allow us to isolate the effect of the time-varying conditional correlation from that of the particular stationary distribution chosen for the diffusion process. This is an extension that we pursue in the third Chapter of this thesis.

Chapter 3

Dynamic Correlation Hedging in Copula Models for Portfolio Selection

3.1 Introduction

An increasing body of literature is interested in modeling time variations in the conditional dependence of asset returns in terms of conditional covariances and correlations (Bollerslev et al. (1988) or Engle (2002) to cite a few). From a modeling perspective, popular choices for the time-varying correlation phenomenon are the Dynamic Conditional Correlation model of Engle (2002) in a discrete-time setting, or the continuous-time Wischart process, introduced by Bru (1991) that gives rise to an affine model and tractable portfolio allocation rules.

The main theme behind those models is the fact that the correlation structure of world equity markets is not constant over time, but is highly time varying. A number of studies have addressed this issue, as well as the driving factors behind this time variation. Based on data from the last 150 years, Goetzmann et al. (2005) find that correlations between equity returns vary substantially over time and achieve their highest levels during periods characterized by highly integrated financial markets. As well peaks in correlations and not only volatility can be attributed to major market crashes, as for example the Crash of 1929. Longin and Solnik (1995) study shifts in global equity markets correlation structure and reject the hypothesis of constant correlations among international stock markets. Moreover, they find evidence that correlations increase during highly volatile periods. Using Extreme Value Theory, Longin and Solnik (2001) find that international stock markets tend to be highly correlated during extreme market downturns than during extreme market upturns, establishing a pattern of asymmetric dependence. Further, Ang and Chen (2002) confirm this finding for the US market for correlations between stock returns and an aggregate market index. Another strand of literature connects the variability of stock return correlations to the phase of the business cycle. Ledoit et al. (2003) and Erb et al. (1994) show that correlations are time-varying and depend on the state of the economy, tending to be higher during periods of recession. Similar evidence is brought forward by Moskowitz (2003) who links time variation of volatilities and covariances to NBER recessions.

The above empirical findings find theoretical justification in Ribiero and Veronesi (2002) where in a Rational Expectations Equilibrium model time variations in correlations are obtained endogenously as a result of changes in agents' uncertainty about the state of the economy. Further, by relating recessionary periods to a higher level of uncertainty, excess co-movements across international stock markets are obtained during bad times when the global economy slows down.

The evidence of highly varying conditional correlations on equity markets has motivated us to propose a continuous time process for asset prices that incorporates the above mentioned stylized facts in two distinct ways. First, we allow for tail dependence between extreme realizations of asset returns by explicitly modeling the stationary distribution of the process using copula functions that incorporate dependence in the left or the right tail. This construction of a multivariate diffusion with a pre-specified stationary distribution relies on Chen et al. (2002) and it allows us to obtain higher dependence when markets experience downturns than during upward moves. However, this approach does not exploit the conditional correlation structure of the process. To this end, we further propose a specification for modeling correlation dynamics of the process using observed factors, including macroeconomic and market volatility factors. With those we aim at capturing the above mentioned features of asset returns, and namely the fact that correlations increase during extreme market downside moves, hectic periods and recessionary states of the economy.

This chapter further concentrates on the portfolio implications of those distributional assumptions. Staying within a complete market framework, we are able to undertake the standard portfolio solution methodology of Cox and Huang (1989), further developed by Ocone and Karatzas (1991) and Detemple et al. (2003), which allows us to obtain in closed form up to numerical integration the optimal portfolio components in terms of mean-variance demand and intertemporal hedging demands. For the case where we model conditional correlation as a function of observed factors, we are able to isolate the hedging demands for correlation risk, due to stochastic changes in the factors. We use the solution for the optimal portfolio allocation in order to address the following issues:

a) We test whether the implications of allowing for tail dependence through the stationary distribution and for dynamic conditional correlation on the optimal portfolio hedging demands are similar in magnitude and direction. As those distributional assumptions aim at replicating the same stylized behavior, it is interesting to see whether the portfolio effects will share this similarity. For an in-sample market timing exercise along realized paths of the state variables over a 20-year investment horizon and two risky funds, we find that allowing for dynamic conditional correlation generally drives up the intertemporal hedging demands, while allowing for tail dependence in the stationary distribution diminishes them. There is also a distinction in the portfolio composition between the risky funds: in the presence of dynamic conditional correlation the spread between the hedging demands for the two funds increases, while in the presence of tail dependence it decreases, bringing about smaller hedging components in absolute value for the two funds. Those effects become more important when increasing the investment horizon.

b) We further investigate the evolution of the correlation hedging demands implied by the observable factors. Using a factor to capture market-wide volatility and another one to account for macroeconomic conditions, we find that the total correlation demands due to those factors are generally negative throughout the period we consider. The impact of the macroeconomic factor is more significant and directs the behavior of the hedging demands.

c) We test whether results are sensitive to the particular choice of investment period. We consider two sub-periods that differ in the level of stock market volatility and macroeconomic conditions, and we consider an investor with investment horizon set at the end of each of these sub-periods. We find that for a relatively calm period with almost no extreme events towards its end the impact of tail dependence disappears once we allow for a data generating process that incorporates dynamics in the conditional correlation behavior. To the contrary, for a hectic period with declining macroeconomic conditions and a number of extreme events, especially towards its end, the importance of modeling tail dependence for the optimal hedging demand cannot be overwritten by allowing for dynamically varying correlations.

d) We further test the economic importance of considering dynamic conditional correlation or tail dependence using the concept of the certainty equivalent cost and find substantial utility loss due to disregarding either form of dependence, which increases with the investment horizon and for low levels of the agent's relative risk aversion. As well, we find substantial utility loss for disregarding dependence between extreme realizations, even when dynamic conditional correlation has already been accounted for, and vice versa. We also compare different dynamic conditional correlation specifications that take into account or not observable factors and we find that there is utility loss related to disregarding observable factors, especially factors related to macroeconomic conditions.

e) As well we study the sensitivity of the optimal hedging behavior for different levels of average correlation and find higher hedging demands for high correlation levels, when the impact of stochastic changes in conditional correlation on investor's utility is expected to be the highest. This finding is confirmed by the certainty equivalent cost of disregarding dynamic conditional correlation: the utility loss increases for higher levels of average correlation. Alternatively, we study the impact of disregarding tail dependence for varying levels of tail dependence coefficients in the data generating process and find that there are far more significant costs of disregarding dependence between extreme realizations when its level increases, even when dynamic conditional correlation is already taken into account.

The present study is closely related to the work of Buraschi et al. (2007) who solve for the optimal portfolio hedging behavior in the presence of correlation risk in a setting where both volatilities and correlations are stochastic, giving rise to separate demands for volatility and correlation risk. They model covariance dynamics using the analytically tractable Wischart process and study the portfolio impact of stylized facts of asset returns such as volatility and correlation persistence and leverage effects. However they work in an incomplete market setting which allows them to obtain closed-form portfolio solutions for only the CRRA investor. While in Buraschi et al. (2007) the correlation between the risky assets is stochastic and is driven by its independent risk source, the model of Liu (2007) allows for stochastic correlations that however are deterministic functions of return volatilities, which does not allow disentangling the portfolio effect of correlation from that of volatility. Under this model's assumptions, including quadratic returns, for which the four elements, describing the investment opportunity set (the short rate, the maximal squared Sharpe ratio, the hedging coefficient vector, and the unspanned covariance matrix), are all Markovian diffusions with quadratic drift and diffusion coefficients, it is again possible to obtain explicit dynamic portfolio solutions for an investor with CRRA utility. The portfolio problem can be solved under either complete markets (when utility is defined over consumption and terminal wealth) or incomplete markets (when utility is defined only over terminal wealth).

The portfolio solution methodology that we consider allows us to identify the intertemporal hedging demands that arise from the need to hedge against changes in the stochastic investment opportunity set, and separate them from the mean-variance component. As well, we can solve under general utility preferences, that are not constrained to the CRRA case. We consider a case when conditional correlation is modeled as a deterministic function of the state variables driving volatility, and alternatively as a function of observed state variables, linked to market-wide volatility and macroeconomic conditions. In the second case we are able to isolate the correlation hedging demands that appear due to the need to hedge against fluctuations in the observed factors.

The present study is also related to another strand of literature that studies the implications of asset co-movements on dynamic portfolio choice. Ang and Bekaert (2002) consider a regime-switching model of asset returns that accounts for asymmetries in their dependence structure by including a 'bear' regime with low expected returns, coupled with high volatilities and correlations, and a 'normal' regime with high expected returns, low volatilities and correlations. They find that the asymmetric correlation structure between the two regimes becomes important for an international investor only when she is allowed to trade in the risk-free asset. Only in this case there are any significant economic costs of disregarding regime switching. Liu et al. (2003) model event related jumps in prices and volatility in the double-jump framework, introduced by Duffie et al. (2000). The presence of event jumps renders the optimal portfolio holdings similar to those that could be obtained for an investor faced with short-selling and borrowing constraints. As well, event risk has a larger impact on the portfolio composition of investors with low levels of risk aversion. However, these results are obtained for a single risky asset portfolio. Das and Uppal (2004) consider the impact of systemic risk on dynamic portfolio choice by introducing a jump component in asset prices that is common for all assets. They work in a constant investment opportunity set and find that investors who ignore systemic risk would have larger holdings of the risky assets. As well, there is higher cost associated to ignoring systemic risk for investors with low levels of risk aversion and levered portfolios. In this setting there are portfolio effects due to higher moments that arise from the inclusion of jumps. Alternatively, Cvitanic et al. (2008) develop optimal allocation rules under higher moments when risky assets are driven by a time-changed diffusion of the Variance Gamma type, and find that ignoring skewness and kurtosis leads to overinvestment in the risky assets and a substantial wealth loss, especially for high volatility levels.

In this chapter we consider an alternative way to model asset co-movement asymmetries through the stationary distribution of the process for the state variables, driving the prices of the risky assets. We introduce an asymmetric dependence structure of the distribution explicitly by using copula functions that allow us to isolate the effect of the marginal distributions from that of the dependence structure itself. This allows us to model the above mentioned stylized facts without reverting to an incomplete market through the inclusion of jumps, which allows us to have a tractable portfolio solution for a general utility function specification. We chose between copula functions that incorporate dependence between extreme realizations of the state variables and copulas that imply no tail dependence and study the differences in the intertemporal hedging demands entailed by the alternative data generating processes.

The remainder of the chapter is organized as follows. Section 3.2 discusses several stylized facts of dynamic correlation and motivates the possibility to model it using observable factors. Section 3.3 describes the model, the solution to the portfolio choice problem, and the correlation hedging demands that appear due to observable factors driving correlation. Section 3.4 discusses the particular portfolio holdings for a bivariate application. In Section 3.5 we present numerical results used to gauge the importance of hedging demands that arise due to dynamic correlation or tail dependence. Section 3.6 concludes. Technical details are provided in the Appendix.

3.2 Dynamic correlation and exogenous factors

Established empirical findings point towards several stylized facts that characterize conditional correlation dynamics of asset returns. It tends to increase in periods of high market volatility, or in cases of extreme downside market moves. As well, it appears to be linked to the business cycle and is higher in recessionary states of the economy.

We approach the above mentioned facts in two methodologically distinct ways. First, we achieve increased correlation during market downturns through the stationary distribution of the multivariate diffusion of state variables that underlines the stock price process. With this 'static' approach we are able to achieve a certain degree of left tail dependence which translates into increased dependence for low levels of the state variables. Second, we allow for dynamic correlation of the state variables, driven by factors that are supposed to capture market volatility and the state of the business cycle. To this end, we choose the Chicago Board Options Exchange Volatility Index (VIX) which measures the implied volatility of S&P index options and thus incorporates market's expectations of near-term volatility. In order to incorporate the effect of the business cycle on the dynamics of correlation, we take the Chicago Fed National Activity Index (CFNAI) that synthesizes information on various macroeconomic factors in a single index. It is a monthly index that aggregates information on overall macroeconomic activity and inflation, as it is a weighted average of 85 indicators of national economic activity, ranging from production, employment, housing and consumption, income, sales, orders and inventories. The methodology behind the CFNAI is based on Stock and Watson (1999), who find a common factor behind various inflation indicators. The evolution of the VIX and of the CFNAI are given in Figure 3.2.1.

In order to appreciate the time variation in asset correlations, driven by the chosen indices, we estimate a DCC model with exogenous factors on the asset return series that will be used later in the portfolio application. Data used in this study consists in two stock market indices representing old economy stocks (S&P 500) and new economy stocks (NASDAQ) for the period 1986-2006. This relatively long period includes several market crashes among which the October 1987 crash in the beginning of the sample period, the Asian crisis that triggered the market crash in October 1997, as well as the Dot-com bubble crash in 2000-2002.

The DCC specification, as well as the estimated coefficients are given in Table 3.2.1, and

Figure 3.2.1: Evolution of the VIX index (upper panel) and of CFNAI index (bottom panel) for the period 1986 - 2006.

The VIX is quoted in terms of percentage points and the data is available at the daily frequency. The CFNAI is quoted monthly. A negative value of the CFNAI index indicates a below-average growth of the national economy, whereas a positive value of the index points towards an above-average growth. A zero value means that the economy grows at its historical average rate.

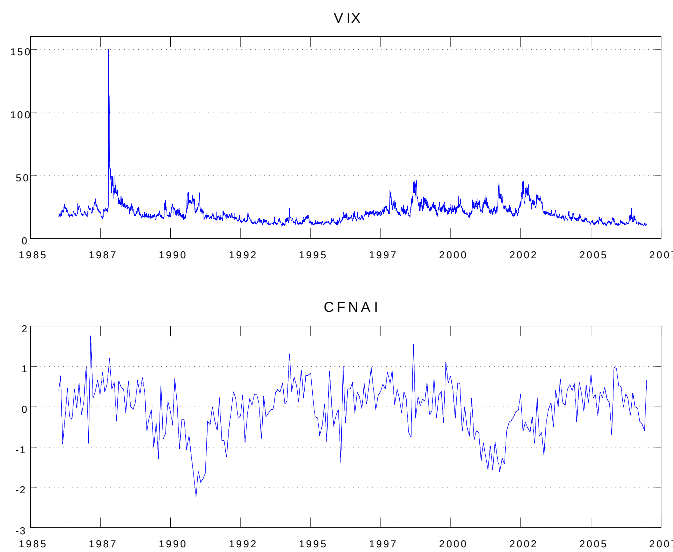


Table 3.2.1: Parameter estimates of a DCC model with exgenous factors for SP 500 and NASDAQ returns.

The model that we estimate is an extended version of the DCC model of Engle (2002) to include exogenous factors driving the conditional covariance and it has the following specification. Denote by y_t the $d \times 1$ vector of asset returns, and by F_t the $n \times 1$ vector of exogenous variables. Then for the conditional mean equation we have:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t \\ \varepsilon_t &= H_t^{1/2}\eta_t \text{ where } \eta_t \sim N(0,1) \text{ thus } \varepsilon_t \sim N(0, H_t) \end{aligned}$$

The conditional covariance matrix H_t can be expressed as $H_t = D_t R_t D_t = \left(\rho_{ij,t}\sqrt{h_{ii,t}h_{jj,t}}\right)$, where $\rho_{ij,t}$ are entries of the conditional correlation matrix and $h_{ii,t}$ are entries of the conditional covariance matrix. Further, $R_t = \widetilde{Q}_t^{-1} Q_t \widetilde{Q}_t^{-1}$, where $\widetilde{Q}_t^{-1} = diag\left(\sqrt{q_{ii,t}}\right)$. The dynamics of Q_t are given by:

$$Q_t = \overline{Q}\,(1-\alpha-\beta) + \alpha\widetilde{\varepsilon}_{t-1}\widetilde{\varepsilon}_{t-1} + \beta Q_{t-1} + Ip^{\intercal}F_{t-1}$$

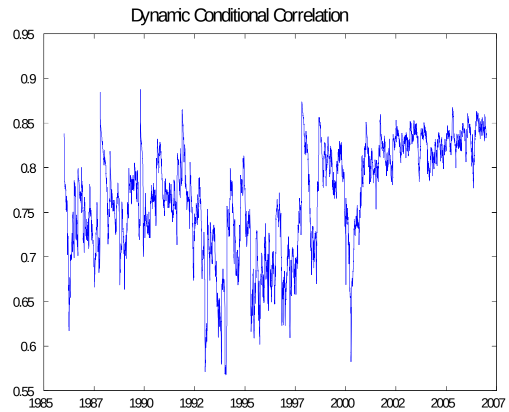
where $\widetilde{\varepsilon}_t \sim N(0, R_t)$, and it is a $d \times 1$ vector of standardized residuals $\widetilde{\varepsilon}_t = \frac{\varepsilon_t}{\sqrt{h_{ii,t}}}$, I is the identity matrix and p is an $n \times 1$ vector of parameters pertaining to the exogenous factors F_t.
In our case y_t denotes the returns of S&P 500 and NASDAQ, and F_t are the VIX and the CFNAI indices. Parameter estimates and their corresponding standard errors are given below.

	Parameters	**Standard errors** $\times$ 1000
α	0.0221	(0.1025)
β	0.9744	(0.0063)
$p_1\,(VIX)$	0.0008	(0.0000)
$p_2\,(CFNAI)$	-0.0001	(0.0081)

the correlation dynamics are plotted in Figure 3.2.2.

All the DCC parameters are significantly estimated which points towards a certain degree of persistence of correlation. Estimated correlation levels range between 0.55 and 0.90 and there can be seen a general tendency of increasing correlation over the years. There are some distinct spikes in conditional correlation, some of which can be linked to specific events (e.g. the late 1987 and 1997 market crashes). There is a distinct period of lower conditional correlations between 1992 and 1997, which is also characterized by low market volatility and a generally above average growth trend in the economy. The parameters for the exogenous factors that drive the time-varying conditional covariance have the expected signs: positive for the VIX and negative for the CFNAI, which translates into increasing conditional correlation during hectic periods and recessionary states of the economy.

Figure 3.2.2: Estimated dynamic conditional correlation for SP 500 and NASDAQ returns from a DCC model with exogenous factors



3.3 The investment problem

This section describes the problem faced by the investor in allocating her wealth between a set of risky assets and the money market account. It introduces the distributional and utility assumptions we impose and presents the general solution methodology using the Martingale technique following the portfolio decomposition formula of Detemple et al. (2003) and its implementation via Monte Carlo simulations. We consider the case where the investor maximizes expected utility of terminal wealth, so that we do not allow for intermediate consumption.

3.3.1 The economy

We define a filtered probability space $\left(\mathcal{F}_T^X, \{\mathcal{F}_T^X\}_{t=0}^T, P^Y\right)$over the investment horizon $[0, T]$ where $\mathcal{F}_T^Y$ is the filtration generated by state variables Y_t under the empirical probability measure P^Y. We consider a complete market setup with $d+1$ state variables $Y_{it}, i = 1, ...d$, where uncertainty is driven by $d+1$ Brownian motions $W_{it}, i = 1, ...d+1$. There are $d+2$ securities available for investment: d stocks, a long term pure discount bond, and the risk-free asset. The state variable vector Y_t consists of $d+1$ state variables X_t, each one affecting its corresponding stock price process, and a state variable Y_t^r that governs the dynamics of the short rate r_t, that is $Y_t = (Xt, Y_t^r)^\intercal$.

The investor has at her disposal the following three asset categories. First, she can invest in a risk-free money market account and its value at time t is given by:

$$B_0(t) = \exp\left\{\int_0^t r(s, Y_s^r)\, ds\right\} \tag{3.3.1}$$

As well, another tradeable asset in the portfolio is a default-free zero-coupon bond with a maturity T. Its price $B(t,T)$ at time t can be expressed as a conditional expectation under the equivalent martingale measure Q:

$$B(t,T) = E^Q\left[\exp\left\{-\int_t^T r(s, Y_s^r)\, ds\right\} |\mathcal{F}_t^Y\right] \tag{3.3.2}$$

The rest of the portfolio consists in a collection of stocks whose price process is modeled using the d state variables X_t:

$$S_i(t) = \exp(X_{it} + \varphi(t)) \quad , i = 1, ..., d \tag{3.3.3}$$

where $\varphi(t)$ is a deterministic function of time. This specification was chosen in order to be as close as possible to the Geometric Brownian motion underlying the Black-Scholes formula for option pricing: if the process for X_{it} is given by $X_{it} = X_{i0} + \sigma_i \int_0^t dW_{it}$, then we are exactly in the Black-Scholes setting where all the assets are independent from each other; if alternatively we apply a stochastic time transformation to the Brownian motion and define the process for X_{it} as $X_{it} = X_{i0} + \int_0^t \sigma(t, X_{it})\, dW_{it}$, then we obtain a simple generalization of the Geometric Brownian motion that already departs from the normality assumption. As it will be shown below, we will further introduce a drift to the process for the state variables X_t which will be consistent with a chosen stationary distribution for the process, as well as correlations between the Brownians that will be allowed to be stochastic. This will bring the model closer to the discrete-time alternative of a dynamic conditional correlation model, as the one introduced by Engle (2002).

3.3.2 The affine setup for the bond price

In what follows, we will restrict the framework for the bond price to the affine class, in that the short interest rate r_t will be an affine function of state variable Y_t^r. This will allow us to express the yield of the bond as an affine function of the state variable as well. Thus, we assume that the short rate can be expressed as:

$$r\left(t, Y_t^r\right) = \delta_0 + \delta_1 Y_t^r \tag{3.3.4}$$

The choice of a one-factor affine model for the short rate may be questionable as there is substantial empirical evidence concerning the shortcomings of affine models[1], and as well using only one factor to capture the dynamics of the term structure may be too restrictive. But as the specification for the bond is marginal for our portfolio application, we proceed with this simple specification which ensures tractable portfolio solutions. As well, Y_t^r has the simple interpretation as a state variable that models the dynamics of the interest rate risk factor which will further determine the hedging terms of the portfolio against changes in the stochastic interest rate.

Following the evidence of time-varying interest rate risk premia on the bond market (e.g. Chan et al., 1992), we allow the state variable Y_t^r to evolve over time according to a square-root process. Its dynamics under the objective measure P^Y are given by:

$$dY_t^r = \kappa_r\left(\theta^r - Y_t^r\right)dt + \sigma_r\sqrt{Y_t^r}\,dW_t^r \tag{3.3.5}$$

Following Dai and Singleton (2000), we assume a market price of risk of the form $\lambda\sqrt{Y_t^r}$, which ensures that the process for the state variable will be affine under the risk neutral measure as well. Then under the equivalent martingale measure the process will be:

$$dY_t^r = \overline{\kappa_r}\left(\overline{\theta^r} - Y_t^r\right)dt + \sigma_r\sqrt{Y_t^r}\,dW_t^{*r} \tag{3.3.6}$$

where $\overline{\kappa_r} = \kappa_r + \sigma_r\lambda$ and $\overline{\theta^r} = \kappa_r\theta^r / \left(\kappa_r + \sigma_r\lambda\right)$.

Given the affine term structure parametrization is admissible, we can obtain in closed form the price of the default-free bond:

$$B\left(t, T\right) = \exp\left\{a\left(T-t\right) + b\left(T-t\right)Y_t^r\right\} \tag{3.3.7}$$

where $a\left(\tau\right)$ and $b\left(\tau\right)$ solve the Ricatti equations:

[1] Backus et al. (1998) show that term premiums generated by affine models are too low compared to the observed data; Duffee (2002) finds that this class of models is not flexible enough to replicate temporal patterns in interest rates.

$$\begin{aligned}\frac{\partial a(\tau)}{\partial \tau} &= \overline{\theta^r}\,\overline{\kappa_r} b(\tau) - \delta_0 \\ \frac{\partial b(\tau)}{\partial \tau} &= -\overline{\kappa_r} b(\tau) + \frac{1}{2}\left(\sigma_r b(\tau)\right)^2 - 1\end{aligned}$$

Then the process for the bond price can be recovered from (3.3.6) and (3.3.7) and the specification of the market price of risk that we adopted. Thus, it can be shown that the bond price follows:

$$\begin{aligned}dB_t &= B_t\left[\mu^B\left(t, Y_t^r\right) dt + \sigma^B\left(t, Y_t^r\right) dW_t^r\right] \qquad (3.3.8)\\ \text{where } \mu^B\left(t, Y_t^r\right) &= r\left(t, Y_t^r\right) + b(\tau)\,\sigma_r \lambda Y \\ \text{and } \sigma^B\left(t, Y_t^r\right) &= b(\tau)\,\sigma_r \sqrt{Y_t^r}\end{aligned}$$

As a result of the CIR specification of the state variable Y_t^r, the market price of risk defined by $\Theta^B\left(t, Y_t^r\right) = \sigma^B\left(t, Y_t^r\right)^{-1}\left(\mu^B\left(t, Y_t^r\right) - r\left(t, Y_t^r\right)\right)$ is stochastic and is given by $\lambda\sqrt{Y_t^r}$. It should be noted that for the bond risk premium to be positive, the market price of risk and thus λ should be negative.

3.3.3 The copula diffusion for the stock price process with dynamic conditional correlation

In this section we will define the process for the state variables X_t that drive the stock prices. As we are interested in modeling the dependence between extreme realizations of returns, we will adopt the copula diffusion process, introduced in the first chapter and extend it to a dynamic conditional correlation specification. Thus, we introduce two channels for modeling extremal dependence: one through the properties of the stationary distribution of the process, and the second through the conditional correlation. We will explore two options for modeling the correlation dynamics. A first straightforward way to do so is to allow the conditional correlation to be time-varying by being specified as some known function of the state variables themselves. As there is evidence that correlation increases in volatile states and when returns are low, we propose to model correlation as a function of the volatility and the level of the state variables. Thus, the general form of the state variables X_t is given by:

$$\textit{Case A:} \qquad dX_t = \mu(X_t)\,dt + \Lambda(X_t)\,dW_t^X \tag{3.3.9}$$

where Λ is a lower triangular matrix, and W^X is a d-dimensional standard Brownian motion, independent of W^r. If we define a continuously differentiable positive definite matrix $\Sigma = \Lambda\Lambda^\intercal$, then its entries are given by $\nu_{ij}(X_t) = \Upsilon_{ij}(X_t)\,\sigma_i^X(X_t)\,\sigma_j^X(X_t)\,,\ \ i,j = 1,...,d$, where the conditional correlation coefficients $\Upsilon_{ij}(X_t)$ and the conditional volatility terms $\sigma_i(X_t)$ are functions of X_t and thus time varying.

The second way to model dynamic correlation that we explore is by rendering it stochastic in terms of a function of observable factors. Following the empirical evidence, that correlations increase in volatile periods and in bad states of the economy, we introduce two exogenous factors to account for that: the CBOE volatility index (VIX) and the Chicago Fed National Activity Index (CFNAI). Denoting these observable factors as F_t, we propose a second general specification for the state variable process X_t of the form:

$$\textit{Case B:} \qquad dX_t = \widetilde{\mu}(X_t, F_t)\,dt + \widetilde{\Lambda}(X_t, F_t)\,dW_t^X \tag{3.3.10}$$

where $\widetilde{\Lambda}$ is a lower triangular matrix, defined as a function of the state variables X_t, as well as the observable factors F_t. The entries of the continuously differentiable positive definite matrix $\widetilde{\Sigma} = \widetilde{\Lambda}\widetilde{\Lambda}'$ are given by $\widetilde{\nu}_{ij}(X_t, F_t) = \widetilde{\Upsilon}_{ij}(X_t, F_t)\,\sigma_i^X(X_t)\,\sigma_j^X(X_t)$, where the conditional correlation coefficient $\widetilde{\Upsilon}_{ij}(X_t, F_t)$ is stochastic in that it is modeled as a function of the observable factors F_t, as well as the state variables X_t. Note that in this second case we augment the state variable vector Y_t to include also the factors F_t: $Y_t = (X_t, F_t, Y_t^r)^\intercal$.

Using any of the above specifications for X_t and the fact that the stock price is defined following (3.3.3) , we can apply Itô's lemma in order to recover the stock price process:

$$\begin{aligned}
dS_{it} &= S_{it}\mu_i^S\left(\log S_{it}-\varphi\left(t\right)\right)dt \\
&\quad +S_{it}\sum_{j=1}^{d}\Lambda_{ij}^I\left(\log S_{it}-\varphi\left(t\right)\right)dW_{jt}^X
\end{aligned} \tag{3.3.11}$$

$$\begin{aligned}
\text{where } \mu_i^S\left(t,Y_t\right) &= \mu_i^I\left(Y_t\right)+\varphi'\left(t\right)+\frac{1}{2}\sum_{j=1}^{d}\overline{\sigma}_{ij}^2\left(Y_t\right) \quad ,I=1,2 \\
\mu_i^1\left(Y_t\right) &= \mu\left(X_t\right), \quad \mu_i^2\left(Y_t\right)=\widetilde{\mu}\left(X_t,F_t\right) \\
\text{and } \Lambda_{ij}^I\left(t,Y_t\right) \quad ,I &= 1,2 \text{ are entries of the corresponding matrix:} \\
\Lambda_{ij}^1\left(t,Y_t\right) &= \Lambda\left(X_t\right), \quad \Lambda_{ij}^2\left(t,Y_t\right)=\widetilde{\Lambda}\left(X_t,F_t\right)
\end{aligned}$$

It should be noted, that as we need to stay within the complete market setup, the number of sources of risk, generated by the Brownian motions, should be the same as the number of traded assets. Thus, when introducing the observable factors F in the stock price specification, we assume that their dynamics are governed by the same Brownian motions that drive the stock prices themselves.

As the market is complete and we have an invertible matrix $\Lambda^{(I)}$, we can define a market price of risk as $\Theta^S\left(t,Y_t^r\right)=\Lambda^{(I)}\left(t,Y_t\right)^{-1}\left(\mu^S\left(t,Y_t\right)-r\left(t,Y_t^r\right)\iota\right)$, where ι is a d-dimensional vector of ones.

Let us stack the drift and diffusion terms for the bond and the stocks so that to obtain:

$$\begin{aligned}
M\left(t,Y_t\right) &= \begin{pmatrix} \mu_i^S\left(t,Y_t\right) \\ \mu^B\left(t,Y_t\right) \end{pmatrix} \\
\Xi\left(t,Y_t\right) &= \begin{pmatrix} & & 0 \\ & \Lambda^{(I)}\left(t,Y_t\right) & \vdots \\ & & 0 \\ 0 & \ldots \quad 0 & \sigma^B\left(t,Y_t^r\right) \end{pmatrix}
\end{aligned}$$

Then the market price of risk for all the tradeable assets
$\Theta\left(t,Y_t\right)=\left(\Theta\left(t,Y_t\right)_1^S,...,\Theta\left(t,Y_t\right)_d^S,\Theta^B\left(t,Y_t\right)\right)$ is defined as:

$$\Theta\left(t,Y_t\right)=\Xi\left(t,Y_t\right)^{-1}\left(M\left(t,Y_t\right)-r\left(t,Y_t^r\right)\iota\right)$$

It is assumed to be continuously differentiable and satisfying the Novikov condition $E\left[\exp\left(\int_0^T \Theta(t,Y_t)^\intercal \Theta(t,Y_t)\,dt\right)\right] < \infty$. The market completeness implies the existence of a unique state price density ξ_t defined as

$$\begin{aligned}\xi_t &\equiv B_0(t)^{-1}\eta_t \exp\left\{-\int_0^t r(s,Y_s^r)\,ds\right\}\times\\ &\exp\left\{-\int_0^t \Theta(s,Y_s)^\intercal dW_s - \frac{1}{2}\int_0^t \Theta(s,Y_s)^\intercal \Theta(s,Y_s)\,ds\right\}\end{aligned}$$

where η_t is the Radon-Nykodym derivative, $\mathcal{F}_T^Y$-adapted. We can also define the conditional state price density that converts cash flows at time $v \geq t$ into cash flows at time t:

$$\begin{aligned}\xi_{t,v} &\equiv \xi_v/\xi_t \\ &= \exp\left\{\begin{array}{c} -\int_t^v r(s,Y_s^r)\,ds - \int_t^v \Theta(s,Y_s)^\intercal dW_s \\ -\frac{1}{2}\int_t^v \Theta(s,Y_s)^\intercal \Theta(s,Y_s)\,ds \end{array}\right\}\end{aligned} \tag{3.3.12}$$

Establishing the diffusion specification for the state variables X that drive the stock price dynamics

Having established two alternative ways to model the conditional correlation dynamics with the aim of answering the stylized fact that asset correlation increases in volatile periods when asset returns are low and the economy is in a downturn, we now turn to the other possibility of accommodating this stylized fact: through the stationary distribution of the state variables, as it has been already explored in the first chapter. Instead of focusing on the dynamics of a correlation measure (the correlation between state variables changing stochastically through time), in this chapter we have modeled the tail dependence (the asymptotic dependence between tail realizations of the state variables) in a 'static' sense. By imposing a certain stationary distribution on the state variables' process, one can obtain different degrees of tail dependence in the left or the right tail of the distribution. Thus, for low levels of the state variable, the tail dependence index may be high, while for high levels of the state variable it may be low, reproducing the stylized fact mentioned above.

For the sake of completeness, we will review the construction of a multivariate diffusion with a given invariant distribution, defined in terms of copula functions. It follows Chen et al. (2002) in exploiting the relationship that exists between the density of the stationary distribution, the drift and the diffusion term of the process defined in (3.3.9) or (3.3.10):

$$\mu_j = \frac{1}{2} q^{-1} \sum_{i=1}^{d} \frac{\partial \left(\nu_{ij} q \right)}{\partial x_i} \tag{3.3.13}$$

where μ and ν_{ij} denote either $\mu(X_t)$ and $\nu_{ij}(X_t)$ for *Case A* or $\widetilde{\mu}(X_t, F_t)$ and $\widetilde{\nu}_{ij}(X_t, F_t)$ for *Case B*, and q is a strictly positive continuously differentiable multivariate density function that is the stationary density of the Markov process for X. Thus the specification of the drift term μ depends on both the form of the invariant density (which will be modeled to determine the degree of asymmetric tail dependence of the state variables X, that is the 'static' representation of the stylized fact of co-movement asymmetries), and the form of the diffusion term Λ (which will be specified in a way to allow or not for dynamic conditional correlation, dependent or not on observable factors, that is the 'dynamic' representation of the same stylized fact).

In what follows we will establish the alternative assumptions on the form of both the invariant density and the volatility term.

The form of the invariant density. With the choice of the stationary distribution we seek to answer several questions concerning the behavior of asset returns. Our major concern is the ability to allow assets to be dependent when they move towards the tails of the distribution, especially for the left tail. This would ensure our model the ability to replicate the empirical fact that asset returns are increasingly dependent as they jointly move towards the lower quantiles of their distribution, that is during market downturns. As copula functions allow us the flexibility to impose different types of joint behavior on the variables while keeping the marginal distributions unchanged, we build the invariant density q based on the copula density representation following Sklar's theorem:

$$q(x_1, ..., x_d) \equiv \widetilde{c}(x_1, ..., x_d) \prod_{i=1}^{d} \widetilde{f}^i(x_i) \tag{3.3.14}$$

where $\widetilde{c}(x_1, ..., x_d) = c\left(F^1(x_1), ..., F^d(x_d)\right)$ is a copula density defined over the univariate CDFs $F^i(x_i)$, and $\widetilde{f}^i(x_i)$ are the corresponding non-normalized univariate densities. We choose the Normal Inverse Gaussian (NIG) distribution[2] to model the univariate behavior because of its proven ability to account for stylized facts of univariate asset return dynamics: autocorrelation of squared returns, semi-heavy tails, possibly asymmetric. Its tail behavior is richly parametrized, nesting tails that vary from an exponential to a power law. As well, NIG is one of the few members of the class of Generalized Hyperbolic (GH) distributions that is closed under convolution, that is if the distribution of log prices is modeled under a NIG law, then the distribution of the increments (asset returns) is also NIG. The univariate NIG diffusion is also an alternative to the widely used NIG Levy process (e.g. Eberlein and Keller, 1995; Prause, 1999) that allows for an infinite number of jumps in the price process, but that also imposes independence of the increments, which is not the case for its diffusion counterpart.

The most important feature of the copula density representation (3.3.14) is that it allows us to separate the effect of the marginal behavior from the implications of the dependence structure, modeled using a copula function. This is important for the portfolio application that we treat in this study, as it allows us to gauge the difference between the different ways to model asset dependence (and thus to reproduce or not the stylized fact of asymmetric asset co-movements) without the impact of the particular assumptions for the univariate stock price processes. Thus we could measure the impact of the 'static' representation of dependence, ranging from Gaussian (no extreme co-movements) to non-negative tail dependence (extreme co-movements, possibly asymmetric) on the optimal portfolio terms.

Let us first remind the definition of the coefficients of upper and lower tail dependence for couples of random variables X and Y: upper tail dependence is defined as the limit probability of the variable Y exceeding the upper quantile as we approach it, conditional upon the fact that the random variable X has exceeded that same quantile:

$$\lambda_U = \lim_{u \to 1} \Pr\left[Y > F_Y^{-1}(u) \,|X > F_X^{-1}(u)\right]$$

Alternatively, we define the coefficient of lower tail dependence as:

$$\lambda_L = \lim_{u \to 0} \Pr\left[Y \leq F_Y^{-1}(u) \,|X \leq F_X^{-1}(u)\right]$$

[2]See the appendix for details.

Both coefficients can be represented in terms of copula functions: $\lambda_U = \lim_{u \to 1} \frac{(1-2u+C(u,u))}{1-u}$ and $\lambda_L = \lim_{u \to 0} \frac{C(u,u)}{u}$. So different copulas will have different degrees of upper and lower tail dependence depending on their parametric specification. Thus, in order to allow for different degrees of tail dependence, we assume several copula specifications for c^3.

Case 1 *Gaussian copula* C^{Ga}*:* $\lambda_U = \lambda_L = 0$

In this case we allow for no dependence between tail realizations of the state variables. The parameter that governs dependence is the correlation coefficient ρ.

Case 2 *Student's t copula* C^t*:* $\lambda_U = \lambda_L = 2t_{\nu+1}\left(-\frac{\sqrt{\nu+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)$

where t_ν is the Student's t density for ν degrees of freedom. In this case the copula function allows for symmetric tail dependence, determined by the correlation parameter ρ and the degrees of freedom parameter ν.

Case 3 *A Gaussian - Symmetrized Joe-Clayton (SJC) mixture copula* C^{Ga-SJC}*:* $\lambda_U \neq \lambda_L$

The form of the mixture copula is given by:

$$C^{Ga-SJC} = \omega C^{SJC} + (1-\omega)\, C^{Ga}$$

where C^{Ga} stands for the Gaussian copula function and C^{SJC} - the Symmetrized Joe-Clayton copula, with a mixing parameter ω that determines the weights of each of the copulas. The symmetrized Joe-Clayton copula models separately upper and lower tail dependence and its form is particularly appealing, as the tail dependence coefficients are themselves the parameters of the copula function. It has been proposed by Patton (2004) as a symmetrized version of the Joe-Clayton copula, in order to overcome the drawback of the latter in that even when the coefficients of upper and lower tail dependence are equal to each other, there still exists some asymmetry in the copula, due to its functional form.

We consider a mixture specification with this copula and the tail independent Gaussian one in order to answer the concerns raised in Poon et al. (2004) that a copula specification whose coefficients explicitly allow for tail dependence may overestimate the dependence in the tail regions. Thus, by the mixture copula we let the data determine whether the

[3] See the appendix for details on the alternative specifications of the copula functions used in the paper.

dependence structure is closer to one imposing no tail dependence or to one that allows for it.

The cases considered above follow closely the ideas behind the copula diffusion introduced in the first chapter. In all of them dependence is modeled explicitly through the invariant density of the multivariate state variable process. In the following section we will extend this setup and will introduce dynamics in the modeling of dependence through the conditional correlation coefficient.

The conditional correlation dynamics. Before proceeding to the specification of the conditional correlation, we need to define the conditional volatility dynamics. Recall that the diffusion term of X was defined as a lower triangular matrix Λ and the entries of the variance-covariance matrix $\Sigma = \Lambda\Lambda^{\intercal}$ are given by $\nu_{ij}(X_t) = \Upsilon_{ij}(X_t)\,\sigma_i^X(X_t)\,\sigma_j^X(X_t)$. Borrowing the idea of Bibby and Sorensen (2003) for modeling the diffusion term of a univariate GH stationary process, we allow each $\sigma_i^X(X_t)$ to be a function of the state variables X_t:

$$\sigma_i^X(X_t) = \sigma_i\left[\widetilde{f}^i(x_i)\right]^{-\frac{1}{2}\kappa_i} \tag{3.3.15}$$

where $\widetilde{f}^i(x_i)$ is the non-normalized NIG density for X_i, and we have the following parameter restrictions: $\sigma_i > 0$ and $\kappa_i \in [0,1]$. By expressing the volatility term as the inverse of a power function of the density $\widetilde{f}$ we obtain the familiar U-shape for the volatility, typical for a stationary process. This specification is especially interesting, as it nests the constant conditional volatility as a special case, setting $\kappa_i = 0$. Thus, for the portfolio allocation application, we could easily isolate a volatility hedging component due to stochastic conditional volatility by opposing a model with $\kappa_i \neq 0$ to one that restricts the conditional volatility to be constant ($\kappa_i = 0$).

Earlier in this section we have discussed two possibilities of rendering the conditional correlation coefficient dynamic: through modeling it as a function of the state variables X or by allowing it to be influenced by stochastic factors F. Here we will further elaborate the particular assumptions concerning those two cases.

In both cases the conditional correlation coefficient Υ_{ij} is modeled as a function $h_{ij}(Y_t)$ of the stochastic state variables Y, whether or not augmented with the observable factors. In order to keep the correlation coefficient in $[-1,1]$, we apply the following logistic transform

$\mathcal{A}$ on the function $h(Y_t)$:

$$\Upsilon_{ij}(Y) = \mathcal{A}(h_{ij}(Y)) = \frac{1-\exp(-h_{ij}(Y))}{1+\exp(-h_{ij}(Y))}$$

Case A. Dynamic conditional correlation with state variables: $\Upsilon(X_t)$

As our aim is to replicate the stylized fact that correlation between asset returns increases in volatile periods and in extreme market downturns, we model the dynamic conditional correlation coefficient as a function involving the volatility specification considered earlier (3.3.15), as well as the level of the state variables in terms of their probability integral transforms $F(X_i)$. More specifically, we model the function $h_{ij}(\cdot)$ as:

$$h_{ij}(X_t) = \gamma_{ij,0} + \gamma_{ij,1}\max\left(\sigma_1^X(X_t), ..., \sigma_d^X(X_t)\right) + \gamma_{ij,2}\prod_{i=1}^{d} F(X_{it}) \qquad (3.3.16)$$

where $F(X_{it})$ stands for the corresponding univariate NIG CDF. The second term in this specification involves the conditional volatilities of each univariate series. We expect to obtain a positive coefficient $\gamma_{ij,1}$ to reflect the fact that correlation increases in hectic periods. We define this term as the maximum over all individual volatilities in order to allow high volatility in any of the stocks to trigger increased conditional correlation. This specification was also used in Goorbergh et al. (2003) in order to model the dynamics of a conditional copula through Kendall's tau in an option pricing application. The third term is motivated by the fact that conditional correlation shoots up when stock prices jointly and abruptly decline, thus we expect a negative sign for the coefficient $\gamma_{ij,2}$.

Case B. Dynamic conditional correlation with observed factors and latent variables: $\Upsilon(X_t, F_t)$

Instead of letting the dynamics of the conditional correlation parameter be determined exclusively by the state variables that drive the stock price process, we model it instead with observable factors that are believed to drive conditional correlation: the VIX and the CFNAI macroeconomic index. Thus we aim at replicating the stylized fact that correlation increases in volatile markets when the economy is in a bad state. As the economic cycle does not necessarily coincide with bear/bull financial markets, we leave from the previous specification the term that determines the level of the state variable. More specifically, in this case we model the function $h(\cdot)$ as:

$$h_{ij}\left(X_t,F_t\right)=\gamma_{ij,0}+\gamma_{ij,1}F_t^V+\gamma_{ij,2}\prod_{i=1}^{d}F\left(X_{it}\right)+\gamma_{ij,3}F_t^M \tag{3.3.17}$$

where $F_t^V=\log\left(VIX_t\right)$ and $F_t^M=CFNAI$. The second term in this expression involves the VIX and thus tries to account for the fact that conditional correlation will rise in periods of increased volatility, so that we expect a positive sign for $\gamma_{ij,1}$. The third term involves the probability integral transforms of the state variables X and is thus meant to capture the fact that correlation increases in market downturns (which entails an expected negative coefficient $\gamma_{ij,2}$). The last term involves the macroeconomic factor and thus aims at capturing the effect of the economic cycle on conditional correlation. As the CFNAI index is designed to take positive values when the economy is in an upturn and negative values otherwise, we expect to obtain a negative sign for $\gamma_{ij,3}$.

Case C. Dynamic conditional correlation with observed factors: $\Upsilon\left(F_t\right)$

If we alternatively believe that correlation is driven by factors that do not affect directly the stock price process, then we may restrict the specification in (3.3.17) in order to include only observable factors:

$$h_{ij}\left(F_t\right)=\gamma_{ij,0}+\gamma_{ij,1}F_t^V+\gamma_{ij,3}F_t^M \tag{3.3.18}$$

This specification will prove quite useful in determining the portfolio correlation hedging demands, as we will see in the following sections, as it will allow us to explicitly identify them from the rest of the hedging terms of the portfolio. This is due to the fact that the factors determining conditional correlation do not affect in a direct way the stock price process itself.

We assume the following processes for the two factors: a CIR process for F^V and a Vasicek process for F^M:

$$\begin{aligned} dF_t^V &= \kappa^V\left(\theta^V-F_t^V\right)dt+\sigma^V\sqrt{F_t^V}dW_t^X \\ dF_t^M &= \kappa^M\left(\theta^M-F_t^M\right)dt+\sigma^M dW_t^X \end{aligned} \tag{3.3.19}$$

These processes will greatly facilitate the implementation of the portfolio allocation formula, as the Vasicek specification will allow for a closed-form solution for the Malliavin derivative of the macroeconomic factor F^M, while the CIR diffusion term will make possible a variance-reduction technique for the Monte Carlo simulation of the Malliavin derivative of F^V.

3.3.4 The investor's objective function

We consider an investor who maximizes utility over terminal wealth, that we denote by $U(\omega_T)$ by choosing an optimal investment policy $\{\alpha_t\}_{t\in(0,T)}$ that belongs to an admissible set $\mathcal{A}$ for an investment horizon T:

$$\max_{\alpha\in\mathcal{A}} E\left[U(\omega_T)\right] \tag{3.3.20}$$

where the utility function U is strictly increasing, concave and differentiable, and satisfies the conditions $\lim_{x\to\infty} U'(x) = 0$ and $\lim_{x\to 0} U'(x) < \infty$. This standard utility specification includes the case of the Hyperbolic Relative Risk Aversion (HARA) utility function $U(\omega) = \frac{1}{1-\gamma}(\omega+b)^{1-\gamma}$ that we assume for this application. The coefficient of Relative Risk Aversion, defined as $R(\omega) \equiv -\frac{U''(\omega)}{U'(\omega)}\omega$, is equal to $\gamma\frac{\omega}{\omega+b}$ for the HARA case, which boils down to a constant γ for the special case of CRRA utility.

The portfolio policy α is a $(d+1)$-dimensional progressively measurable process that is defined as the proportion of wealth allocated to the risky assets (d stocks and a long term pure discount bond). Thus, the amount invested in the risk-free asset (the money-market account) is $(\omega - \alpha^\intercal 1)$. The portfolio policy generates a wealth process ω whose dynamics are given by:

$$d\omega_t = \omega_t\left\{r_t dt + \alpha_t^\intercal\left[\left(M(t,Y_t) - r_t\iota\right)dt + S(t,Y_t)\,dW_t\right]\right\} \tag{3.3.21}$$

3.3.5 The complete market solution

The complete market setup that we have adopted allows us to solve for the optimal portfolio using the Martingale solution technique that restates the dynamic budget constraint (3.3.21) as a static one and first solves for the optimal terminal wealth, and then finds the optimal portfolio policy that finances it. Thus, following Cox and Huang (1989), optimal terminal wealth is given by $\omega_T^* = I(y\xi_T)^+ = \max\left(I(y\xi_T), 0\right)$, where $I = [U']^{-1}$ denotes the inverse of the marginal utility function, and y satisfies the static budget constraint $E\left[\xi_T I(y\xi_T)^+\right] =$

ω_0, where ω_0 is the initial wealth.

Following Ocone and Karatzas (1991), and using the portfolio decomposition formula of Detemple et al. (2003), we have the following expression for the optimal portfolio policy, that decomposes the portfolio holdings into a Mean Variance part (α^{MV}), an Interest Rate Hedge (α^{IRH}) and a Market Price of Risk hedge (α^{MPRH}):

$$\begin{aligned}\alpha_t^* &= \alpha_t^{MV} + \alpha_t^{IRH} + \alpha_t^{MPRH} \qquad (3.3.22)\\ & \text{where} \\ \alpha_t^{MV} &= \left(\Lambda^{\intercal}(t, Y_t)\right)^{-1} \frac{1}{R(\omega_T)} \Theta(t, Y_t) E_t \left[\xi_{t,T} \frac{\omega_T}{\omega_t} \frac{R(\omega_t)}{R(\omega_T)} 1_{\omega_T > 0}\right] \\ \left(\alpha_t^{IRH}\right)^{\intercal} &= -\left(\Lambda^{\intercal}(t, Y_t)\right)^{-1} E_t \left[\xi_{t,T} \frac{\omega_T}{\omega_t} \left(1 - R(\omega_T)^{-1}\right) I_{\omega_T > 0} H_{t,T}^r\right] \\ \left(\alpha^{MPRH}\right)^{\intercal} &= -\left(\Lambda^{\intercal}(t, Y_t)\right)^{-1} E_t \left[\xi_{t,T} \frac{\omega_T}{\omega_t} \left(1 - R(\omega_T)^{-1}\right) 1_{\omega_T > 0} H_{t,T}^{\Theta}\right]\end{aligned}$$

The terms $H_{t,T}^r$ and $H_{t,T}^{\Theta}$ involve the sensitivities of the short rate and the market price of risk towards shocks in the Brownian motions that drive uncertainty in the model and are defined as follows:

$$H_{t,T}^r = \int_t^T \mathcal{D}_t r_s ds = \int_t^T \partial_2 r(s, Y_s) \mathcal{D}_t Y_s \qquad (3.3.23)$$

$$\begin{aligned} H_{t,T}^{\Theta} &= \int_t^T \left(dW_s + \Theta(s, Y_s) ds\right)^{\intercal} \mathcal{D}_t \Theta(s, Y_s) ds \qquad (3.3.24)\\ &= \int_t^T \left(dW_s + \Theta(s, Y_s) ds\right)^{\intercal} \partial_2 \Theta(s, Y_s) \mathcal{D}_t Y_s ds\end{aligned}$$

where the operator $\mathcal{D}$ is the Malliavin derivative, $\partial_2 f(t, x)$ refers to the derivative with respect of the second argument of $f(t, x)$, and where the second equality was obtained using the chain rule for Malliavin derivatives. For the state variables needed in our application, the Malliavin derivatives are given by:

$$\mathcal{D}_t Y_s = \begin{pmatrix} \mathcal{D}_{1,t} X_{1,s} & \cdots & \mathcal{D}_{d,t} X_{1,s} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \mathcal{D}_{1,t} X_{d,s} & \cdots & \mathcal{D}_{d,t} X_{d,s} & 0 \\ \mathcal{D}_{1,t} F_s^V & \cdots & \mathcal{D}_{d,t} F_s^V & 0 \\ \mathcal{D}_{1,t} F_s^M & \cdots & \mathcal{D}_{d,t} F_s^M & 0 \\ 0 & \cdots & 0 & \mathcal{D}_{d+1,t} Y_s^r \end{pmatrix} = \begin{pmatrix} \mathcal{D}_t X_{1,s} \\ \vdots \\ \mathcal{D}_t F_s^V \\ \mathcal{D}_t F_s^M \\ \mathcal{D}_t Y_s^r \end{pmatrix}$$

The implementation of the above formula follows Detemple et al. (2003) and relies on the fact that the Malliavin derivatives, as well as the state variables, follow stochastic differential equations that can be simulated using standard discretization techniques. Given the particular specification of some of the state variables, we can further apply the Doss transformation[4], reducing the stochastic differential equation of the given state variable to one with a constant diffusion term, which ensures that the Malliavin derivative does not involve a stochastic term. Specific solutions for the Malliavin derivative are given in the appendix.

The long term bond and the interest rate hedging demands

Let us first consider the term $H_{t,T}^r$ that involves the sensitivity of the short rate towards shocks in the underlying Brownian motions. Recall that $r(s, Y_s) = \delta_0 + \delta_1 Y_t^r$, and that the $(d+3)$-dimensional state variable vector, augmented with the observable factors, is defined as $Y \equiv \left(X_1, ..., X_d, F^V, F^M, Y^r\right)^\intercal$. Thus $\partial_2 r(s, Y_s) = (0, ..., 0, \delta_1)$, and using the fact that $\mathcal{D}_{d+1,t} Y_s = (0, ..., 0, \mathcal{D}_{d+1,t} Y_s^r)$, then:

$$H_{t,T}^r = \left(0, ..., 0, \int_t^T \delta_1 \mathcal{D}_{d+1,t} Y_s^r\right)$$

So the long term bond is the sole security in the portfolio that is used to hedge against changes in the short rate.

3.3.6 Correlation hedging

The above portfolio decomposition formula isolates interntemporal hedging demands due to stochastic changes in the short rate or the market price of risk from the mean-variance

[4]See Detemple et al. (2003) for further details.

demand. As in *Cases B* and *C* we have modeled conditional correlation as a function of certain observable factors, the sensitivities of those factors to shocks in the underlying Brownian motions would give rise to hedging demands that can be related (partially for *Case B*) to correlation hedging. As in *Case A* conditional correlation is modeled as a deterministic function of the state variables, determining as well the drift, volatility, and subsequently the market price of risk dynamics, we cannot isolate correlation hedging from the total intertemporal demands in this case. The only way to judge the importance of dynamic correlation modeling for portfolio allocation in this case is to contrast the hedging demands, obtained under a DCC specification with those obtained from a CCC process. We will consider this possibility in the following sections when we consider a real data application.

Isolating the correlation hedging demands involving observable factors

As the primary objective of this chapter is to explicitly isolate the correlation hedging demands in the portfolio that arise from stochastic changes in the conditional correlation, let us now consider the second term $H^{\Theta}_{t,T}$ in the portfolio decomposition formula that handles the sensitivity of the market price of risk towards shocks in the underlying state variables. Let us define the vector Ψ in terms of the market price of risk and the state variables:

$$\Psi_t = (dW_t + \Theta(t, Y_t)ds)^{\intercal}\, \partial_2\Theta(t, Y_t)$$

Note that in *Case B* for the conditional correlation specification, where we have augmented the state variables Y to include observable factors $F = \left(F^V, F^M\right)^{\intercal}$, the vector Ψ will be of dimension $(d+3)$. Then we could represent the $H^{\Theta}_{t,T}$ in terms of Ψ_t and the Malliavin derivatives of the state variables as:

$$H^{\Theta}_{t,T} = \int_t^T \Psi_t \mathcal{D}_t Y_s$$

where $\Psi_t \mathcal{D}_t Y_s$ could be further decomposed as follows:

$$(\Psi_t \mathcal{D}_t Y_s)^\intercal = \begin{pmatrix} \Psi_{1,t} D_{1,t} X_{1,s} + ... + \Psi_{d,t} D_{1,t} X_{d,s} + \Psi_{d+1,t} D_{1,t} F_s^V + \Psi_{d+2,t} D_{1,t} F_s^M \\ \vdots \\ \Psi_{1,t} D_{d,t} X_{1,s} + ... + \Psi_{d,t} D_{d,t} X_{d,s} + \Psi_{d+1,t} D_{d,t} F_s^V + \Psi_{d+2,t} D_{d,t} F_s^M \\ \Psi_{d+3,t} D_{d+1,t} Y_s^r \end{pmatrix}$$

Apparently, the term $H^{\Theta}_{t,T,d+1}$ corresponding to the bond, does not involve any other Malliavin derivatives except that of the state variable Y^r driving the short rate. As for the interest rate hedge, Y^r will be the only state variable whose sensitivity with respect to uncertainty shocks will determine the market price of risk hedging terms for the long term bond.

For each one of the d stocks the term $H^{\Theta}_{t,T,i}$ can be expressed as:

$$\begin{aligned} H^{\Theta}_{t,T,i} &= \int_t^T \Psi_{1,t} \mathcal{D}_{i,t} X_{1,s} + ... + \int_t^T \Psi_{d,t} \mathcal{D}_{i,t} X_{d,s} \\ &\quad + \int_t^T \Psi_{d+1,t} \mathcal{D}_{i,t} F_s^V + \int_t^T \Psi_{d+2,t} \mathcal{D}_{i,t} F_s^M \end{aligned}$$

The last two terms in this expression involve the Malliavin derivatives of the observable factors with respect to the Brownian shocks. As those factors are solely responsible for describing the dynamics of the conditional correlation in the process for asset returns, then the term

$$\begin{aligned} C^{\Theta}_{t,T,i} &= \int_t^T \Psi_{d+1,t} \mathcal{D}_{i,t} F_s^V + \int_t^T \Psi_{d+2,t} \mathcal{D}_{i,t} F_s^M \\ &= V^{\Theta}_{t,T,i} + M^{\Theta}_{t,T,i} \end{aligned} \tag{3.3.25}$$

can be considered as defining the correlation hedging demands for the stocks arising from the necessity to hedge against changes in the observable factors F. Thus we can isolate the effect of the market-wide volatility factor on correlation through $V^{\Theta}_{t,T,i} = \int_t^T \Psi_{d+1,t} \mathcal{D}_{i,t} F_s^V$, and the effect of the macroeconomic state variables through $M^{\Theta}_{t,T,i} = \int_t^T \Psi_{d+2,t} \mathcal{D}_{i,t} F_s^M$. However, as we have defined the conditional correlation dynamics in (3.3.17) as been driven as well by the state variables X through the level of the returns, there will be additional hedging demands, associated with the Malliavin derivatives of X, that cannot be disentangled from the rest of the market price of risk hedging demands. We would have this problem in all cases when conditional correlation is modeled as a function of state variables that are not exclusively 'reserved' for driving its dynamics. If to the contrary we believe that correlation is driven solely by observable factors (eg. by setting $\gamma_{ij,2} = 0$ in (3.3.17)), or by other latent factors that do not enter the specification for the stock prices (3.3.3) except through correlation itself, then $C^{\Theta}_{t,T,i}$ alone will be responsible for the correlation hedging in the portfolio.

Note as well that in *Case A*, where conditional correlation was defined in terms of only the state variables X that drive the stock price dynamics, the term $C^{\Theta}_{t,T,i}$ is set to zero, but that does not entail zero correlation hedging. It rather means that the correlation hedging demands cannot be explicitly isolated in this case. Nevertheless, their importance can be judged by comparing the hedging terms that arise from a constant conditional correlation stock price process to those that arise from the dynamic conditional correlation specification.

Let us now get back to the portfolio decomposition formula (3.3.22). Using (3.3.25) we can now isolate the Market Price of Risk (MPR) hedging terms that arise from hedging changes in the observable factors that drive correlation, that is, the correlation hedging demands:

$$\left(\alpha^{CORR}\right)^{\intercal} = -\left(\Lambda^{\intercal}(t, Y_t)\right)^{-1} E_t\left[\xi_{t,T}\frac{\omega_T}{\omega_t}\left(1 - R\left(\omega_T\right)^{-1}\right) 1_{\omega_T > 0} C^{\Theta}_{t,T}\right] \tag{3.3.26}$$

where $C^{\Theta}_{t,T} = \left(C^{\Theta}_{t,T,1}, ..., C^{\Theta}_{t,T,d}\right)$. This defines the explicitly identifiable correlation hedging demand in our setting. It will amount to the full correlation hedging demand for *Case C* when the factors driving correlation do not affect in a direct way the stock price process.

We can restate the above result in terms of the sensitivity of the cost of optimal wealth to changes in the factors driving the conditional correlation dynamics, as the optimal portfolio policy is indeed obtained as one that finances optimal terminal wealth. Recall that optimal wealth at time t is given by $\omega^*_t = E_t\left[\xi_{t,T}\omega^*_T\right]$, where $\xi_{t,T}\omega^*_T = \xi_{t,T} I\left(y\xi_t\xi_{t,T}\right)^+$ represents its cost. Then for a nonnegative $I\left(y\xi_T\right)$ its sensitivity with respect to fluctuations in the observable factors F is given by:

$$\left[I\left(y\xi_t\xi_{t,T}\right)+y\xi_t\xi_{t,T}I'\left(y\xi_t\xi_{t,T}\right)\right]\left(-\xi_{t,T}\right)\times$$
$$\int_t^T \left(dW_s+\Theta(s,Y_s)ds\right)^{\intercal}\partial_2\Theta(s,Y_s)D_tF_s$$

where we have used (3.3.12) and the fact that $I'(y)=\left(u''\left(I\left(y\right)\right)\right)^{-1}$ which follows from the definition of $I(y)$ as the inverse of the marginal utility. Thus, the portfolio terms that are responsible for the sensitivity of the cost of optimal terminal wealth to fluctuations in the factors are indeed the correlation hedging demands defined in (3.3.26).

3.4 A bivariate application: S&P500 vs. NASDAQ

In order to appreciate the impact of the correlation hedging demands on the optimal portfolio composition in a realistic setting and compare them to the intertemporal hedges that arise due to incorporating tail dependence in the stationary distribution of the process for the state variables, driving asset prices, we offer an application based on real data. We consider a portfolio, formed by a 10-year pure discount bond, as well as two risky funds, represented by old and new economy stocks: S&P 500 and NASDAQ. An application with this choice of a dataset can be found in Detemple et al. (2003). Data is observed at the daily frequency (except for the CFNAI factor, which is observed monthly) and refers to the period 1986-2006.

Without loss of generality, we assume that the coefficients in the short rate specification (3.3.4) are given by $\delta_0=0$ and $\delta_1=1$, so that for the short rate we have that $r\left(t,Y_t^r\right)=Y_t^r$. Given the fact that both the interest rate and the market price of risk of the long term bond are assumed to be stochastic, the optimal portfolio composition for it will involve both the interest rate and the market price of risk hedging terms. For the CIR specification we have chosen there are no closed-form solutions for the hedging terms, as it would have been the case, have we chosen a Vasicek process instead, but nevertheless we can apply a variance stabilization technique following the Doss transformation that renders a constant the diffusion term of the process for Y^r, as explained in the Appendix.

The long term bond is the only risky asset that is responsible for hedging away the source of risk related to the short rate (W^r), as it is the only one exposed to it. The optimal demand for the bond involves a mean-variance component and an intertemporal component used to hedge against fluctuations in the investment opportunity set, induced by W^r:

$$\begin{aligned}
\alpha^*_{b,t} &= \frac{1}{\sigma^B\left(t, Y^r_t\right)} \left\{ \begin{array}{c} \frac{1}{R(\omega_T)} \Theta^B\left(t, Y^r_t\right) E_t\left[\xi_{t,T} \frac{\omega_T}{\omega_t} \frac{R(\omega_t)}{R(\omega_T)} 1_{\omega_T>0}\right] \\ -E_t\left[\xi_{t,T} \frac{\omega_T}{\omega_t}\left(1 - R\left(\omega_T\right)^{-1}\right) I_{\omega_T>0} H^r_{t,T}\right] \\ -E_t\left[\xi_{t,T} \frac{\omega_T}{\omega_t}\left(1 - R\left(\omega_T\right)^{-1}\right) 1_{\omega_T>0} H^{\Theta}_{b,t,T}\right] \end{array} \right\} \\
\text{where } \sigma^B\left(t, Y^r_t\right) &= b\left(\tau\right) \sigma_r \sqrt{Y^r_t} \\
\text{and } \Theta^B\left(t, Y^r_t\right) &= \lambda \sqrt{Y^r_t}
\end{aligned}$$

In this bivariate application the optimal portfolio parts for the two risky funds have a very intuitive representation. As we have assumed that they are not driven by the Brownian that is responsible for interest rate risk, then the diffusion term of the stock price process is a bivariate diagonal matrix:

$$\Lambda^{(I)} = \left[\begin{array}{cc} \sigma^X_1\left(X_t\right) & 0 \\ \Upsilon\left(Y_t\right) \sigma^X_2\left(X_t\right) & \sqrt{1 - \Upsilon\left(Y_t\right)^2} \sigma^X_2\left(X_t\right) \end{array} \right]$$

where $\sigma^X_i\left(X_t\right), i = 1, 2$ is given by (3.3.15) and the conditional correlation $\Upsilon\left(Y_t\right)$ is either a function of the state variables X_t in Case A, a function of both the state variables X_t and the observable factors F_t in Case B, or a function of only the observable factors F_t in Case C. Given this diagonal structure for $\Lambda^{(I)}$, for the two stock prices we obtain:

$$\begin{aligned}
dS_{1t} &= S_{1t}\left\{\mu^S_1\left(X_t\right) dt + \sigma^X_1\left(X_t\right) dW^X_{1t}\right\} \\
dS_{2t} &= S_{2t}\left\{\mu^S_2\left(X_t\right) dt + \Upsilon\left(Y_t\right) \sigma^X_2\left(X_t\right) dW^X_{1t} + \sqrt{1 - \Upsilon\left(Y_t\right)^2} \sigma^X_2\left(X_t\right) dW^X_{2t}\right\}
\end{aligned}$$

Without loss of generality we have assumed a linear function for $\varphi\left(t\right)$ in the general specification in (3.3.11) given by $k_i t, i = 1, 2$, where k_i is a deterministic trend. Note that the second fund (NASDAQ in our example) is the only one affected by W^X_2-risk, i.e. it can be thought of as the incremental risk factor that influences ‘new-economy’ stocks. On the contrary, the W^X_1 risk factor affects both funds in our portfolio. This has some implications on the optimal portfolio choice. As we will see below, the demand for the second fund is entirely driven by fluctuations induced by exposure to W^X_2-risk. Following the optimal allocation rule outlined in (3.3.22), the demand for NASDAQ is given by:

$$\alpha_{2,t}^{*}=\frac{1}{\sigma_{2}^{X}\left(X_{t}\right)\sqrt{1-\Upsilon\left(Y_{t}\right)^{2}}}\left\{\begin{array}{c}\frac{1}{R(\omega_{T})}\Theta_{2}(t,Y_{t})E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\frac{R(\omega_{t})}{R(\omega_{T})}1_{\omega_{T}>0}\right]\\-E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\left(1-R\left(\omega_{T}\right)^{-1}\right)1_{\omega_{T}>0}H_{2,t,T}^{\Theta}\right]\end{array}\right\}$$

where $\Theta_i(t, Y_t)$ is the market price of risk for the i^{th} fund, and $H_{i,t,T}^{\Theta}$ is the term involving the response to fluctuations in the opportunity set driven by the i^{th} Brownian motion. The absence of the interest rate hedge is due to the fact that the state variable underlying the short rate is not dependent on any of the Brownians driving the risky stocks. The demand for S&P 500 is given by:

$$\begin{aligned}\alpha_{1,t}^{*} &= \frac{1}{\sigma_{1}^{X}\left(X_{t}\right)}\left\{\begin{array}{c}\frac{1}{R(\omega_{T})}\Theta_{1}(t,Y_{t})E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\frac{R(\omega_{t})}{R(\omega_{T})}1_{\omega_{T}>0}\right]\\-E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\left(1-R\left(\omega_{T}\right)^{-1}\right)1_{\omega_{T}>0}H_{1,t,T}^{\Theta}\right]\end{array}\right\}\\
&\quad-\frac{\Upsilon\left(Y_{t}\right)}{\sigma_{1}^{X}\left(X_{t}\right)\sqrt{1-\Upsilon\left(Y_{t}\right)^{2}}}\left\{\begin{array}{c}\frac{1}{R(\omega_{T})}\Theta_{2}(t,Y_{t})E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\frac{R(\omega_{t})}{R(\omega_{T})}1_{\omega_{T}>0}\right]\\-E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\left(1-R\left(\omega_{T}\right)^{-1}\right)1_{\omega_{T}>0}H_{2,t,T}^{\Theta}\right]\end{array}\right\}\\
&= \frac{1}{\sigma_{1}^{X}\left(X_{t}\right)}\left\{\begin{array}{c}\frac{1}{R(\omega_{T})}\Theta_{1}(t,Y_{t})E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\frac{R(\omega_{t})}{R(\omega_{T})}1_{\omega_{T}>0}\right]\\-E_{t}\left[\xi_{t,T}\frac{\omega_{T}}{\omega_{t}}\left(1-R\left(\omega_{T}\right)^{-1}\right)1_{\omega_{T}>0}H_{1,t,T}^{\Theta}\right]\end{array}\right\}-\frac{\Upsilon\left(Y_{t}\right)\sigma_{2}^{X}\left(X_{t}\right)}{\sigma_{1}^{X}\left(X_{t}\right)}\alpha_{2,t}^{*}\end{aligned}$$

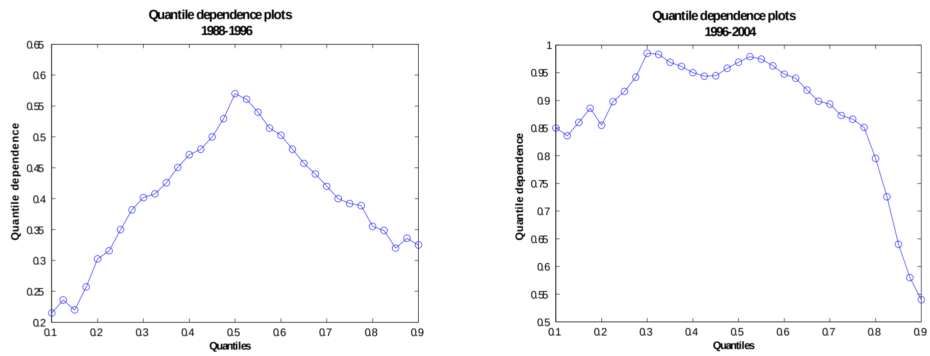
Thus, we can see that for the first fund the optimal portfolio demand has an additional term that involves $\alpha_{2,t}^*$, the optimal holdings of the second fund. It happens because the second fund depends also on W_1^X-risk, so its holding induces also an exposure to it. Consequently, the first fund is used to hedge away this induced exposure, hence the additional term in the optimal portfolio holdings $\alpha_{1,t}^*$. A similar setup with a triangular diffusion term was used in Detemple et al. (2003) in their multiasset application.

Note that the market price of risk hedging demands α^{MPRH} can be decomposed in a similar fashion for the first fund, which will have induced intertemporal hedging demands equal to $-\frac{\Upsilon(Y_t)\sigma_2^X(X_t)}{\sigma_1^X(X_t)}\alpha_{2,t}^{MPRH}$.

3.5 Numerical Results

Before discussing the estimation results for the various diffusion specifications that we have chosen for the state variables X, let us first look at data itself in order to verify whether the stylized facts that we aim at reproducing are indeed present in the data. In the previous

Figure 3.5.1: Quantile dependence plots
Plots of quantile dependence for the de-trended log-prices of S&P 500 vs. NASDAQ for the 1988-1996 and 1996-2004 subperiods.



sections we have seen that dynamic conditional correlation, modeled using a DCC model with exogenous factors, is indeed time-varying and we can distinguish periods of relatively high or low correlation, that we were able to attribute to the influence of the macroeconomic or the volatility factor. In a similar fashion, we split the estimation period in two subsamples, one characterized by decreasing and low volatility and improving macroeconomic conditions (1988-1996), and the other characterized by high volatility and declining and relatively low CFNAI index, pointing towards a declining economy (1996-2004). We then construct quantile dependence plots for the de-trended log-prices of both indices for the corresponding subsamples.

As we can see on Figure 3.5.1, during the first relatively calm period dependence in the extreme quantiles of the joint distribution decreases substantially, even though it does not disappear completely, as one would expect under a Gaussian distributional assumption. As well, a test of tail dependence symmetry, following Hong et al. (2003), does not fail to reject symmetric tails for this particular period, as it can be seen from Table 3.5.1.

On the other hand, the period of (1996-2004) brings about extremely high dependence in the tail quantiles, especially in the left tail, and the dependence symmetry test indeed rejects symmetric tails for the period. Thus, the unconditional distribution of the two risky funds that we have chosen does possess the features that we try to asses, and namely increased dependence when markets experience extreme downturns. Also splitting the sample in two periods with quite distinct characteristics will help us later on to explain the portfolio implications of both conditional correlation and unconditional dependence.

Table 3.5.1: Test of symmetry in the exceedence correlations
The Hong et al. (2003) test of exceedence correlations symmetry in the lower and upper quartiles for the de-trended log-prices of S&P 500 vs. NASDAQ for the 1988-1996 and 1996-2004 subperiods. The test statistic is given by:

$$J = n\left(\rho^{+}-\rho^{-}\right)\Omega^{-1}\left(\rho^{+}-\rho^{-}\right) \xrightarrow{d} \chi^2_m$$

where ρ^{+} and ρ^{-} are the exceedence correlations calculated at the corresponding quantile levels, n is the sample size and m is the number of quantile levels considered.

	1988-1996	1996-2004
Test statistic (J)	6.9048	21.5517
p-values	(0.4389)	(0.0030)

Table 3.5.2: Parameter estimates for the observable factors
Estimated parameters for the observable factors VIX and CFNAI that have the following specifications:

$$\begin{aligned} dF_t^V &= \kappa^V\left(\theta^V - F_t^V\right)dt + \sigma^V\sqrt{F_t^V}dW_t^X \\ dF_t^M &= \kappa^M\left(\theta^M - F_t^M\right)dt + \sigma^M dW_t^X \end{aligned}$$

where $i = \{V, M\}$.

parameter	$CFNAI$	MC s.e.	SIF	VIX	MC s.e.	SIF
κ^i	2.2521	0.0027	0.8153	1.2094	0.0021	0.8002
θ^i	-0.0457	0.0018	1.7702	2.7800	0.0007	0.8863
$\left(\sigma^i\right)^2$	2.9383	0.0005	0.8631	0.1230	0.0000	1.9260

The processes for the observable factors and for the state variables for the risky funds are estimated using Markov Chain Monte Carlo and the Simulation Filter of Golightly and Wilkinson (2006a). This estimation methodology is particularly convenient for highly non-linear multivariate diffusions, as in our case. As well, it allows us to filter out unobservable data points, as is the case of the CFNAI factor, which is observed monthly, whereas the two indices, as well as the VIX factor are observed at the daily frequency. Parameter estimates for the observable factors are given in Table 3.5.2.

Let us not turn to the estimation results for the whole sample period, as well as the two subsamples for the four conditional correlation specifications (DCC, Cases A through C, and CCC) and the three alternative stationary distribution assumptions (no tail dependent Gaussian, symmetric tail dependent Student's t, and asymmetric tail dependent Gaussian-SJC diffusions). As in this application we aim at determining the impact of the stationary

Table 3.5.3: Univariate parameter estimates

Parameter estimates from the univariate Normal Inverse Gaussian (NIG) diffusions with density $f_{NIG}(x;\theta)$, where $\theta = (\alpha, \beta, \delta, \mu)$ is the vector of NIG parameters that satisfy the restrictions, given in the Appendix. The diffusion for each of the state variables X_{it} has the following specification:

$$\begin{aligned} dX_{it} &= b\left(X_{it};\theta_i\right)dt + v\left(X_{it};\theta_i\right)dW_{it} \\ \text{where } b\left(x;\theta\right) &= \frac{1}{2}v\left(x;\theta\right)\frac{d}{dx}\ln\left[v\left(x;\theta\right)f_{NIG}\left(x;\theta\right)\right] \\ v\left(x;\theta\right) &= \sigma^2 f_{NIG}\left(x;\theta\right)^{-\kappa}, \quad \sigma^2 > 0, \kappa \in [0,1] \end{aligned}$$

Monte Carlo standard errors, obtained using the batch-mean approach (multiplied by a factor of 1000) and the simulation inefficiency factor (SIF) are reported for each parameter estimate.

parameter	X_1 $(S\&P500)$	MC s.e.	SIF	X_2 $(NASDAQ)$	MC s.e.	SIF
α	5.6431	0.0601	1.0262	4.2938	0.2138	0.8070
β	-0.6272	0.3091	1.1979	-0.7072	0.4151	0.6343
δ^2	0.0471	0.0016	0.7755	0.0549	0.0026	0.8782
μ	4.6342	0.0083	1.0129	5.1191	0.0146	0.6724
σ^2	0.0268	0.0006	0.8375	0.0222	0.0003	0.2821
κ	0.5776	0.0128	1.0339	0.5349	0.0356	1.2291

distribution and hence tail dependence on the optimal portfolio holdings, regardless of the univariate marginals, we do not proceed to a full-scale optimization of all model parameters, as would be otherwise preferred, but rather undertake a two-step estimation procedure. In a first step, we assume that the two price processes are independent from each other, imposing the independence (or product) copula on their stationary distribution, as well as zero conditional correlation. Thus we are able to estimate them separately, and further use the same marginal distribution parameters for all alternative processes that we consider. In this manner, differences in portfolio demands between the alternative specifications will not depend on the particular parameter choice of the univariate marginals. Parameter estimates are reported in Table 3.5.3. The trend parameters k_i for each of the state variables X_i are estimated separately as a linear trend. Their values are 0.1014 for S&P 500 and 0.1100 for NASDAQ.

In a second step, we assume the marginal parameters as known and we proceed to the estimation of the multivariate processes by assuming all the alternative specifications for the stationary distribution of the conditional correlation. Results are reported in Table 3.5.4.

Table 3.5.4: Parameter estimates from the multivariate diffusion specifications (1986-2006)

Estimates for the parameters of the stationary density, defined in terms of copula functions, and the parameters governing the correlation dynamics for a bivariate diffusion, defined as:

$$dX_t = \mu(X_t)\,dt + \Lambda(X_t)\,dW_t^X$$

$$\text{where } \Lambda = \begin{bmatrix} \sigma_1 \left[\widetilde{f}^1(x1)\right]^{-\frac{1}{2}\kappa_1} & 0 \\ \Upsilon_{12}(X_t)\,\sigma_2 \left[\widetilde{f}^2(x_2)\right]^{-\frac{1}{2}\kappa_2} & \sqrt{1-\Upsilon_{12}^2(X_t)}\sigma_2 \left[\widetilde{f}^2(x_2)\right]^{-\frac{1}{2}\kappa_2} \end{bmatrix}$$

$$\mu_j = \frac{1}{2}q^{-1}\sum_{i=1}^{2}\frac{\partial(\nu_{ij}q)}{\partial x_i}, \quad j = 1,2$$

$$\text{and } q(x_1,...,x_d) \equiv \widetilde{c}(x_1,...,x_d)\prod_{i=1}^{d}\widetilde{f}^i(x_i)$$

where ν_{ij} are entries of the matrix $\Sigma = \Lambda\Lambda^{\intercal}$, and $q(x_1,...,x_d)$ is the stationary density of the diffusion, defined in terms of a copula function $\widetilde{c}$ and the NIG marginal densities $\widetilde{f}^i$. Parameter estimates are given for three cases of copulas: Ga refers to the Gaussian copula, $Ga - SJC$ - to the mixture Gaussian-Symmetrized Joe-Clayton copula, and T - to the Student's t copula. The copula parameters are as follows: ρ is the correlation parameter for the Gaussian or the Student's t copula, ν stands for the degrees of freedom of the Student's t copula, τ_U and τ_L are the upper and lower tail dependence parameters of the Symmetrized Joe-Clayton copula, and ω is the weighting parameter in the Symmetrized Joe-Clayton copula. The parameters that describe the correlation dynamics are $\gamma_i, i = 0,...,3$, consistent with the specification in (3.3.16) for Case A, with (3.3.17) for Case B and with (3.3.18) for Case C. The Constant Conditional Correlation model in Panel 4 assumes that all correlation parameters are zero but γ_0.

Panel 1. Dynamic conditional correlation (Case A)

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.4612	0.3126	0.9440	0.4686	0.2022	0.1966	0.4433	0.6026	1.8164
ν	-	-	-	-	-	-	6.4394	2.0178	0.7087
τ_U	-	-	-	0.5179	0.6057	1.2630	-	-	-
τ_L	-	-	-	0.5003	0.5589	1.2407	-	-	-
ω	-	-	-	0.5599	0.7806	1.6945	-	-	-
γ_0	2.0695	0.0126	0.1636	2.0475	0.0292	0.8041	1.9795	0.0454	1.0962
γ_1	0.4430	1.6643	2.4494	0.6850	0.7402	0.4886	1.3272	0.9481	1.3758
γ_2	-1.4731	0.0422	0.5547	-1.2649	0.0721	0.9250	-0.8214	0.0987	1.3498

Panel 2. Dynamic conditional correlation (Case B)									
param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.4036	0.3654	0.9608	0.4596	0.7086	1.6656	0.3652	0.2750	1.6598
ν	-	-	-	-	-	-	6.6976	9.2680	1.2306
τ_U	-	-	-	0.4669	0.3453	0.6012	-	-	-
τ_L	-	-	-	0.5178	0.3165	1.1565	-	-	-
ω	-	-	-	0.5513	0.7156	0.7900	-	-	-
γ_0	1.7273	0.0166	0.6051	1.7401	0.0252	0.6578	1.7589	0.0381	1.2715
γ_1	0.0060	0.0126	0.9784	0.0034	0.0062	0.3958	-0.0020	0.0145	0.7090
γ_2	-0.2873	0.0642	0.9762	-0.2745	0.0484	0.6097	-0.4227	0.0806	1.1133
γ_3	-0.3086	0.0263	1.0807	-0.3487	0.0240	0.9340	-0.2944	0.0252	1.3209

Panel 3. Dynamic conditional correlation (Case C)									
param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.3734	0.3210	0.3841	0.4984	0.5621	0.9206	0.4146	0.9349	1.8356
ν	-	-	-	-	-	-	6.0105	2.2653	0.4660
τ_U	-	-	-	0.5619	0.3398	0.6210	-	-	-
τ_L	-	-	-	0.4805	0.2818	0.8046	-	-	-
ω	-	-	-	0.4690	0.2544	0.2023	-	-	-
γ_0	1.6288	0.0237	1.0122	1.5920	0.0303	1.7129	1.6122	0.0190	1.1783
γ_1	0.0085	0.0102	0.9935	0.0089	0.0108	0.7495	0.0090	0.0112	2.7264
γ_2	-	-	-	-	-	-	-	-	-
γ_3	-0.2628	0.0394	1.5915	-0.3540	0.0269	0.7109	-0.2510	0.0183	0.6519

Panel 4. Constant conditional correlation									
param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.4565	0.2337	1.2678	0.4918	0.3299	1.1136	0.4052	0.2187	0.5737
ν	-	-	-	-	-	-	4.3149	2.5652	1.6841
τ_U	-	-	-	0.5012	0.6331	2.3965	-	-	-
τ_L	-	-	-	0.5801	0.4020	1.6656	-	-	-
ω	-	-	-	0.3816	0.6329	1.4994	-	-	-
γ_0	1.9955	0.0139	1.8733	2.0374	0.0128	0.9472	2.0470	0.0090	0.7893

Table 3.5.4 (A). Parameter estimates from the multivariate diffusion specifications (1988-1996)

Panel 1. Dynamic conditional correlation (Case A)

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.4130	0.5533	0.8635	0.3971	0.6171	0.8538	0.3951	0.5808	0.6071
ν	-	-	-	-	-	-	5.8728	8.5498	1.0732
τ_U	-	-	-	0.4479	0.4510	0.5219	-	-	-
τ_L	-	-	-	0.4685	0.6630	1.2260	-	-	-
ω	-	-	-	0.5147	0.8890	1.6429	-	-	-
γ_0	1.8897	0.0680	0.9762	1.8835	0.0812	0.9175	1.9037	0.1103	1.1216
γ_1	1.7028	3.9466	2.3019	2.4512	5.0493	2.0684	3.4598	5.1660	1.5876
γ_2	-1.7556	0.6689	1.6051	-1.7040	0.3697	0.5851	-1.3860	0.4324	0.9937

Panel 2. Dynamic conditional correlation (Case B)

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.4011	0.3787	0.4744	0.3705	0.8203	1.3778	0.4590	0.9433	1.4121
ν	-	-	-	-	-	-	6.0486	6.6124	0.3657
τ_U	-	-	-	0.5159	0.9509	0.9491	-	-	-
τ_L	-	-	-	0.5466	0.7998	0.7297	-	-	-
ω	-	-	-	0.5258	1.4451	1.5272	-	-	-
γ_0	2.1724	0.0426	0.4523	2.1661	0.0716	1.0785	2.1827	0.0425	0.5952
γ_1	0.0102	0.0207	1.1832	0.0079	0.0185	0.5042	0.0112	0.0209	0.8377
γ_2	-0.7282	0.3580	1.2605	-0.9716	0.2619	0.5328	-0.7620	0.2754	1.0575
γ_3	-0.2691	0.1471	1.2750	-0.2887	0.1229	0.7116	-0.2734	0.1109	0.9831

Panel 3. Dynamic conditional correlation (Case C)

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.4111	0.5567	0.5478	0.3633	1.4273	1.2948	0.3155	0.5125	0.5295
ν	-	-	-	-	-	-	5.3833	6.6008	1.1860
τ_U	-	-	-	0.6179	0.4655	0.3730	-	-	-
τ_L	-	-	-	0.4446	1.2408	1.4057	-	-	-
ω	-	-	-	0.5042	0.8104	0.6855	-	-	-
γ_0	2.1615	0.0277	0.4684	2.1441	0.0629	1.7323	2.1550	0.0431	1.2044
γ_1	0.0046	0.0294	2.3139	0.0127	0.0192	1.2285	0.0154	0.0159	1.1747
γ_2	-	-	-	-	-	-	-	-	-
γ_3	-0.3223	0.1225	1.9323	-0.2987	0.0682	0.5600	-0.3104	0.1126	2.1642

Panel 4. Constant conditional correlation

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.3348	0.5310	0.7963	0.4497	0.5185	0.7457	0.3677	0.8407	1.5087
ν	-	-	-	-	-	-	5.5060	8.8090	1.9514
τ_U	-	-	-	0.5447	1.0077	1.2661	-	-	-
τ_L	-	-	-	0.5016	1.1308	1.7278	-	-	-
ω	-	-	-	0.5765	0.8678	0.9065	-	-	-
γ_0	1.7174	0.0585	1.8229	1.6532	0.0460	0.9813	1.6437	0.0397	1.0072

Table 3.5.4 (B). Parameter estimates from the multivariate diffusion specifications (1996-2004)

Panel 1. Dynamic conditional correlation (Case A)

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.5637	0.7203	1.2560	0.5274	0.7029	0.5754	0.3722	0.7644	0.5408
ν	-	-	-	-	-	-	4.5172	4.7443	0.6594
τ_U	-	-	-	0.5158	1.1290	1.1144	-	-	-
τ_L	-	-	-	0.4926	0.6007	0.3596	-	-	-
ω	-	-	-	0.4565	0.7126	1.2339	-	-	-
γ_0	1.4097	0.0483	0.6112	1.3723	0.0702	1.0511	1.3127	0.0621	0.6209
γ_1	2.3400	1.0788	1.0603	2.6907	1.2589	0.8808	2.6206	0.6113	0.4612
γ_2	-0.2872	0.1821	0.8152	-0.3649	0.1190	0.3422	-0.1736	0.1280	1.7100

Panel 2. Dynamic conditional correlation (Case B)

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.5380	0.6157	0.5776	0.5383	1.0569	0.6704	0.3368	0.7195	0.8666
ν	-	-	-	-	-	-	4.4252	8.1150	1.5499
τ_U	-	-	-	0.5093	0.4792	0.2987	-	-	-
τ_L	-	-	-	0.5322	0.9009	1.5219	-	-	-
ω	-	-	-	0.5023	0.7294	1.1890	-	-	-
γ_0	1.9191	0.1318	2.0517	1.9198	0.0576	0.4783	1.7604	0.0689	0.6817
γ_1	-0.0134	0.0221	0.5537	-0.0034	0.0157	0.4258	-0.0083	0.0232	0.5074
γ_2	-0.7266	0.1758	0.8608	-0.7292	0.2284	1.6010	-0.6427	0.1392	0.5934
γ_3	-0.0825	0.0983	0.5140	-0.1403	0.0834	0.6450	-0.0741	0.0710	0.4916

Panel 3. Dynamic conditional correlation (Case C)

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.5892	0.8027	1.4944	0.4499	0.7358	0.7108	0.3368	0.7195	0.8666
ν	-	-	-	-	-	-	4.4252	8.1150	1.5499
τ_U	-	-	-	0.5475	0.5474	0.7226	-	-	-
τ_L	-	-	-	0.4939	0.6894	0.7287	-	-	-
ω	-	-	-	0.6078	0.7941	0.6022	-	-	-
γ_0	1.7373	0.1156	0.9672	1.7783	0.0269	0.4362	1.7604	0.0689	0.6817
γ_1	-0.0341	0.0211	0.5716	-0.0215	0.0074	0.2996	-0.0083	0.0232	0.5074
γ_2	-	-	-	-	-	-	-0.6427	0.1392	0.5934
γ_3	-0.2906	0.1588	1.5428	-0.3344	0.0732	0.9997	-0.0741	0.0710	0.4916

Panel 4. Constant conditional correlation

param	Ga	MC s.e.	SIF	Ga-SJC	MC s.e.	SIF	T	MC s.e.	SIF
ρ	0.3533	0.5154	0.7736	0.3853	1.4276	0.6995	0.3981	0.5216	0.3485
ν	-	-	-	-	-	-	6.0479	3.9435	0.2350
τ_U	-	-	-	0.5242	0.7559	0.9003	-	-	-
τ_L	-	-	-	0.5091	0.7778	0.7893	-	-	-
ω	-	-	-	0.5142	0.9299	0.7847	-	-	-
γ_0	1.1262	0.0668	1.1726	1.1751	0.0473	0.6244	1.1200	0.0567	0.9024

Note that the conditional correlation parameters that pertain to volatility (γ_1) (either observed through the VIX factor or modeled through the state variables X) are generally positive through all the specifications, pointing towards an increase in conditional correlation when there is rise in market-wide volatility. An exception to this is the 1996-2004 period, during which the VIX coefficient is negatively estimated for all stationary distributional assumptions. However, γ_1 has the expected positive sign for the conditional correlation specification with no observable factors. On the other hand, the parameter pertaining to the macroeconomic factor (γ_3) is always negatively estimated, pointing towards a decrease in conditional correlation when there is an improvement in macroeconomic conditions, and vice versa.

3.5.1 Correlation hedging demands along realized paths of the state variables

In order to examine the evolution of the portfolio hedging demands for the estimation period, we proceed to a market timing exercise that consists in simulating ahead the Malliavin derivatives of the state variables, the state price density, as well as the portfolio terms involving hedging against changes in the interest rate (3.3.23) and the market price of risk (3.3.24), while keeping the state variables (the latent variables and the observable factors) at their observed values throughout the period[5]. First, we obtain the optimal portfolio terms for the whole period between 1986-2006 for an investor with a constant, moving-window horizon of 4 years. With this we aim at studying differences between the optimal portfolio parts for the alternative specifications considered above for modeling unconditional or conditional dependence, without any influence of the time horizon. Next, we consider an investor who keeps her investment horizon fixed at the end of the period, thus investigating the horizon effect on the optimal portfolio shares.

As during this relatively long 20 year horizon one can distinguish hectic periods, associated with high volatility, negative CFNAI, pointing towards a slow-down in the economy, and subsequently rising conditional correlation, as well as relatively calm periods with low volatility, mostly positive levels of the CFNAI index and thus low conditional correlation, we proceed to a second market timing experiment, considering instead two subperiods of 8 years. The first one spans between 1988 and 1996 and is characterized by increased volatility

[5] As the CFNAI index is observed at a monthly frequency, we filter the unobservable data points at the daily frequency using the MCMC sequential filter.

and a recession in the US economy in the beginning of the period (between July 1990 and March 1991, as determined by NBER), followed by improving macroeconomic conditions (positive and rising CFNAI), as well as relatively low and declining volatility. As it can be seen on Figure 3.2.2, this period is characterized by falling dynamic conditional correlation. On the other hand, the second period, spanning between 1996 and 2004 is characterized by increased volatility for the whole period, a recession towards the end of the period (March 2001 marks the end of a 10-year expansion period, according to NBER, and there is a trough in business activity in November 2001). Figure 3.2.2 shows a rising trend in the dynamic conditional correlations for the period. For both subperiods we consider an investor who has a fixed investment horizon at the end of each period.

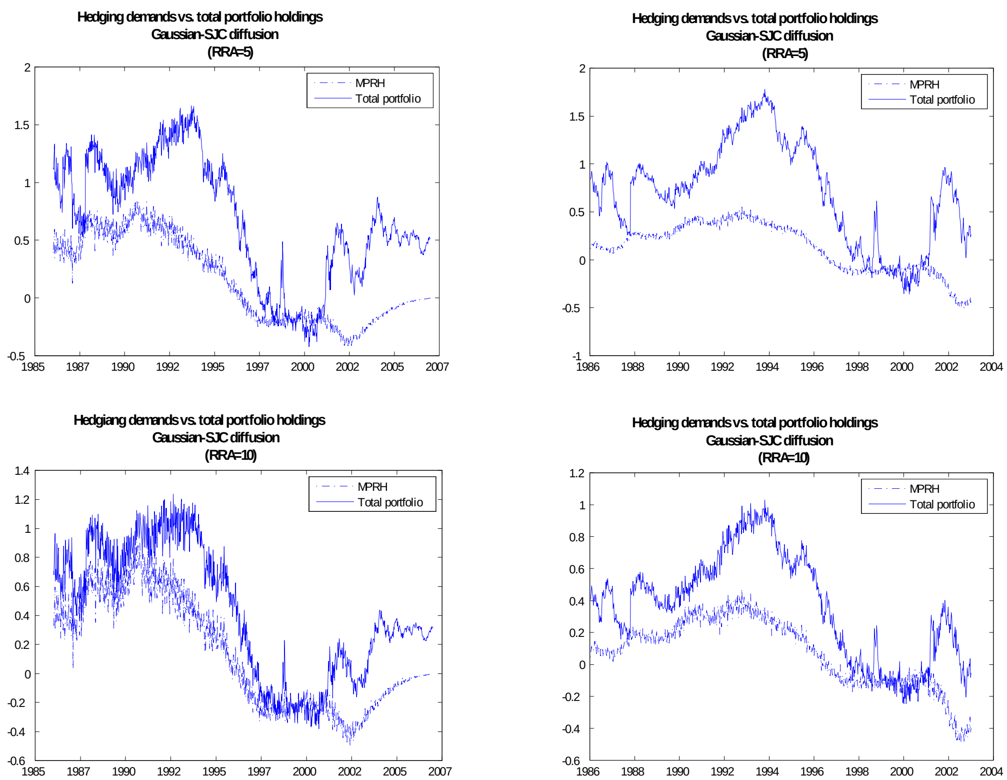
Correlation hedging for the whole estimation horizon

For the first market timing experiment that involves a 20-year investment horizon fixed at the end of the sample, we consider the three cases of modeling the unconditional distribution of the state variables underlying the price processes (non-tail dependent Gaussian, symmetric tail dependent Student's t and asymmetric tail dependent Gaussian-SJC mixture distribution), as well as the three ways to account for dynamically changing conditional correlation with or without observable factors driving it. The same experiment is repeated, but with a moving-window horizon of 4 years. Thus we are able to distinguish the horizon effect in the evolution of the optimal portfolio hedging demands from the effect of the dynamically changing investment opportunity set.

In order to get an impression of the magnitude and the variability of the hedging demands for the risky assets in the portfolio, let us first consider the results displayed on Figure 3.5.2 for a HARA investor with varying degrees of relative risk aversion. The intertemporal hedging demands are a sizeable component of the total portfolio, and they are responsible for a larger portion of the portfolio demands if we increase the level of relative risk aversion of the investor. As well, the hedging demands are larger for longer horizons: an investor with a horizon fixed at the end of the 20-year sample period would have higher hedging demands at each period of time than an investor who has a short rolling-window horizon (4 years in our case). Also the fixed horizon would cause the hedging demands to shrink as we approach it (it is visible during the last 4 years on the left column of Figure 3.5.2), so that

Figure 3.5.2: Total portfolio holdings and intertemporal hedging demands for the two risky stocks over the entire sample

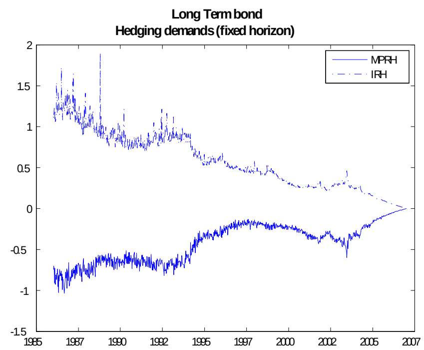
The figure displays the holdings of the two risky stocks in the portfolio for the entire sample period 1886-2006. The total holdings are contrasted with the intertemporal hedging demands, which for the stocks are entirely given by the market price of risk hedges. The figure on the left represents the portfolio holdings for a fixed investment horizon at the end of the 20-year sample. The figure on the right represents the holdings for a moving-window 4-year horizon. The two top figures concern a HARA investor with relative risk aversion of 5, while the bottom two - a HARA investor with relative risk aversion of 10. The data generating process is a Gaussian-SJC diffusion with dynamic correlation (Case B).

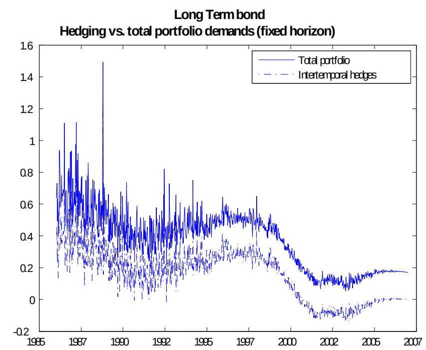


the Mean-Variance component would be increasingly more important in the total portfolio holdings. The results there are based on a Gaussian-SJC diffusion with dynamic correlation driven by observed factors (Case B), but the relative importance of the hedging demands for the other cases is qualitatively the same.

Before we continue with the hedging demands that arise from the different stationary distribution or conditional correlation specifications, let us examine the evolution of the optimal portfolio parts for the long term pure discount bond. As we have already observed in the previous sections, it is the only security in the investor's portfolio in our case that is responsible for hedging interest rate risk.

Figure 3.5.3: Hedging demands for the long term pure discount bond
The top figure displays the hedging demands obtained for the long term pure discount bond for an investment horizon fixed at the end of the 20-year sample: the market price of risk hedge (MPRH) and the interest rate hedge (IRH). The bottom figure plots the total portfolio holdings of the bond against the intertemporal hedging demands which are the sum of IRH and MPRH. HARA investor ($B = -0.1$).





As it can be seen from Figure 3.5.3, the variability of the total portfolio demands is almost entirely driven by the hedging terms. Due to the chosen specification of the market price of risk, we have a negative market price of risk hedging term and a positive interest rate hedge. Due to the fact that the Brownian motion driving the short rate is independent of the Brownian motions driving the rest of the state variables, and that the short rate does not enter the stock price dynamics, the portfolio parts for the bond will remain unchanged for the various specifications for the state variables underlying the stocks that we consider.

Let us now turn to the results for the differences in the hedging demands of the two risky stocks in the portfolio due to the unconditional dependence structure (through the stationary distribution of the process for the state variables underlying stock prices) and due to the dynamics of conditional correlation. On Figure 3.5.4, Panel A we have plotted the correlation hedging demands due to observable factors (CFNAI and VIX) that we have isolated following (3.3.26) for an investor with a fixed horizon at the end of the sample period (left column) and an investor with a rolling-window horizon (right column). On Figure 3.5.4, Panel B we can see the relative importance of the correlation hedging terms due to each one of the factors for the same 20-year investment horizon. The hedge due to the macroeconomic factor is generally negative, reducing the total portfolio demand, while the hedging term due to volatility is positive but very small in absolute value, compared to the CFNAI hedge.

Figure 3.5.4: Correlation hedging demands due to observed factors

Panel A. The figure displays the sum of the hedging demands due to observed factors (CFNAI and VIX) driving conditional correlation for the two risky stocks in the portfolio for the entire sample period 1886-2006. The figure on the left represents the correlation hedging demands for a fixed investment horizon at the end of the 20-year sample. The figure on the right represents the correlation hedging demands for a moving-window 4-year horizon. HARA investor with relative risk aversion of 5. The data generating process is a Gaussian-SJC diffusion with dynamic correlation (Case C).

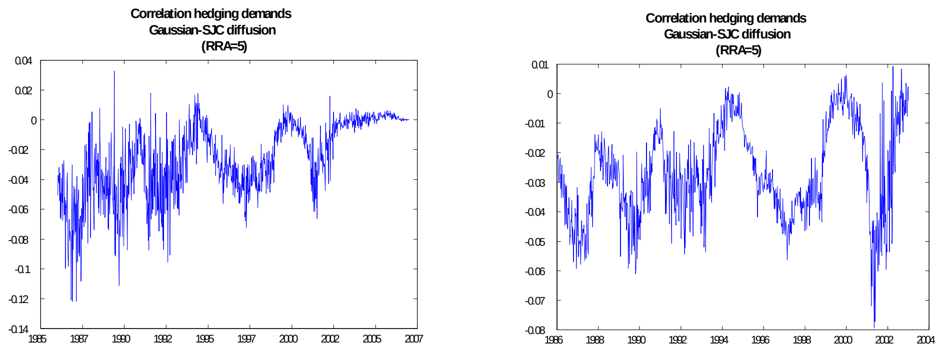
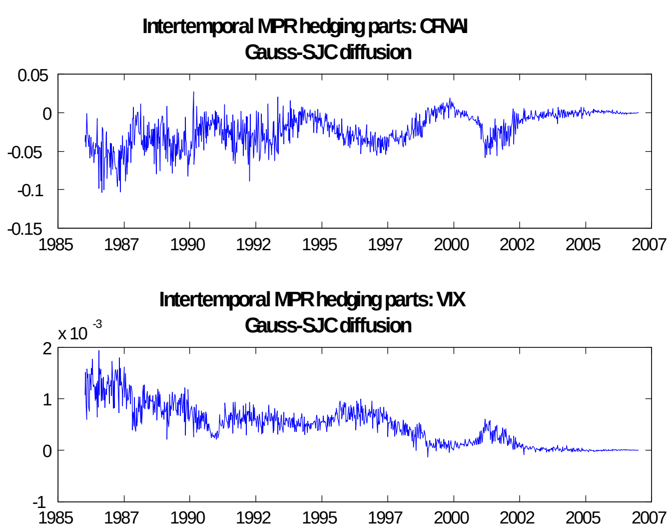


Figure 3.5.4. Panel B. The figure displays the hedging demands due to observed factors driving conditional correlation for the two risky stocks in the portfolio for the entire sample period 1886-2006. The top figure represents the demands due to hedging changes factor that proxies the macroeconomic conditions (CFNAI), while the bottom figure represents the correlation hedging demands due to the factor that proxies market volatility (VIX). HARA investor with relative risk aversion of 5. The data generating process is a Gaussian-SJC diffusion with dynamic correlation (Case B).



The magnitude of these correlation hedging components is quite small compared to the total hedging demands on Figure 3.5.2. They are negative in sign, pointing towards a reduction in the total portfolio holdings. One can as well distinguish periods with peaks in the absolute value of the correlation hedging demands, that can be attributable to some market events (e.g. the market crashes in 1987, 1990-1992, 2001). Those demands are also higher for longer investment horizon, which can be seen by comparing the holdings of the investor with a fixed vs. rolling-window shorter horizon, and they decline to zero as we approach the investment horizon. The results are obtained for the dynamic correlation specification following Case C, that is the case when only the VIX and the CFNAI indices drive conditional correlation. Results for the Case B, as well as Gaussian or the Student's t diffusion are qualitatively the same and are not reported for brevity.

Those hedging demands arise in order to hedge against stochastic changes in the observable factors that proxy volatility or the macroeconomic conditions, and they constitute the total correlation demands in Case C, where the dynamics of conditional correlation are not driven by other state variables. However, as we consider the case of conditional correlation being dependent as well on the level of the state variables X (Case B), then there would be another component in the correlation hedging demands apart from the influence of the factors that is not directly identifiable. In order to gauge its importance, we compare the intertemporal market price of risk hedging parts for a process with dynamic vs. constant conditional correlation. Figure 3.5.5 reports the results for an underlying Gaussian and a Gaussian-SJC diffusion for a fixed investment horizon at the end of the sample period.

The presence of dynamically varying conditional correlation asks for an increase in the intertemporal hedging demands for the Gaussian diffusion, which is mainly driven by NASDAQ, while the hedging demands for S&P 500 are virtually unchanged. At first sight these results are surprising given the evidence that correlation hedging demands due to observable factors for both fixed and rolling window horizon are negative throughout the period, so that we would expect a reduction in the total intertemporal hedging terms for the dynamic conditional correlation case compared to the terms under constant conditional correlation. However, the influence of dynamic correlation does not show up in the correlation hedging term (3.3.25) only through the Malliavin derivatives of the factors. It influences as well the market price of risk $\Theta(t, Y_t)$, which determines the total market price of risk hedging

Figure 3.5.5: Hedging demands along realized paths for the risky stocks for the 20-year fixed investment horizon (Case B)

Plotted are the intertemporal demands (separately for each risky fund and their sum) along realized paths of the state variables for the whole sample period for the risky stocks for a fixed investment horizon at the end of the period The left column plots the intertemporal hedging demands obtained under a DCC specification vs. those under CCC; the right column contrasts hedging terms under constant and time-varying volatility.

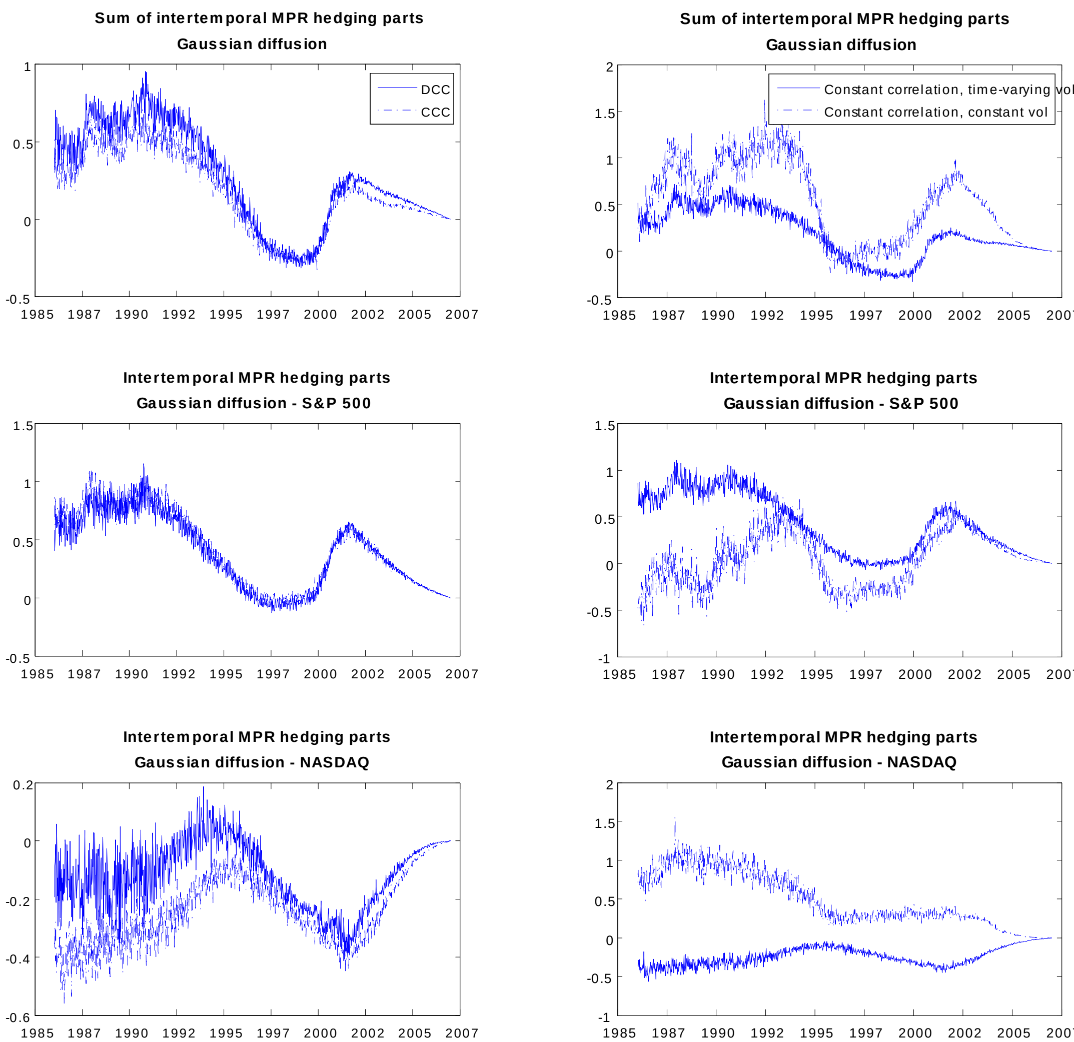


Figure 3.5.5. Panel B. Induced hedging demands (Case B)

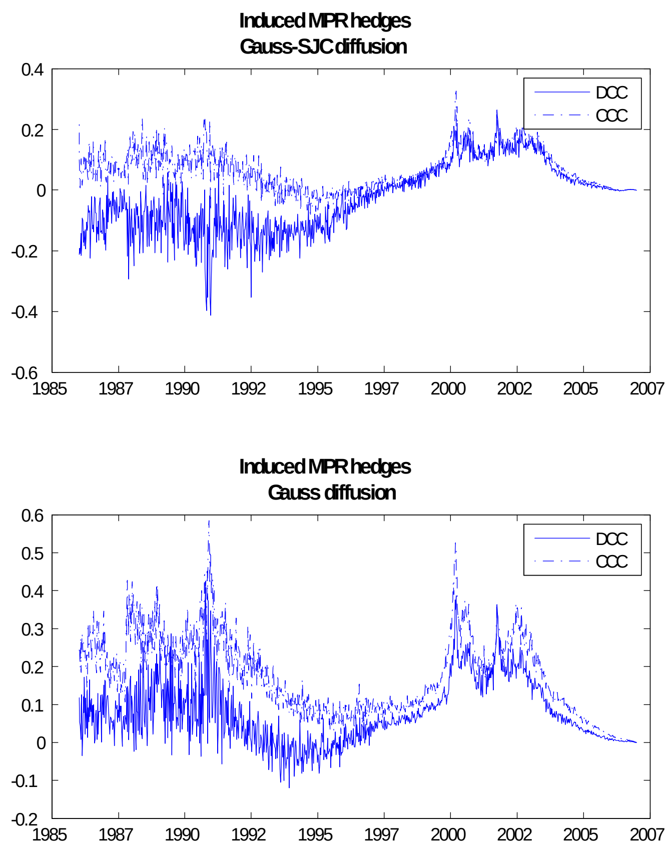
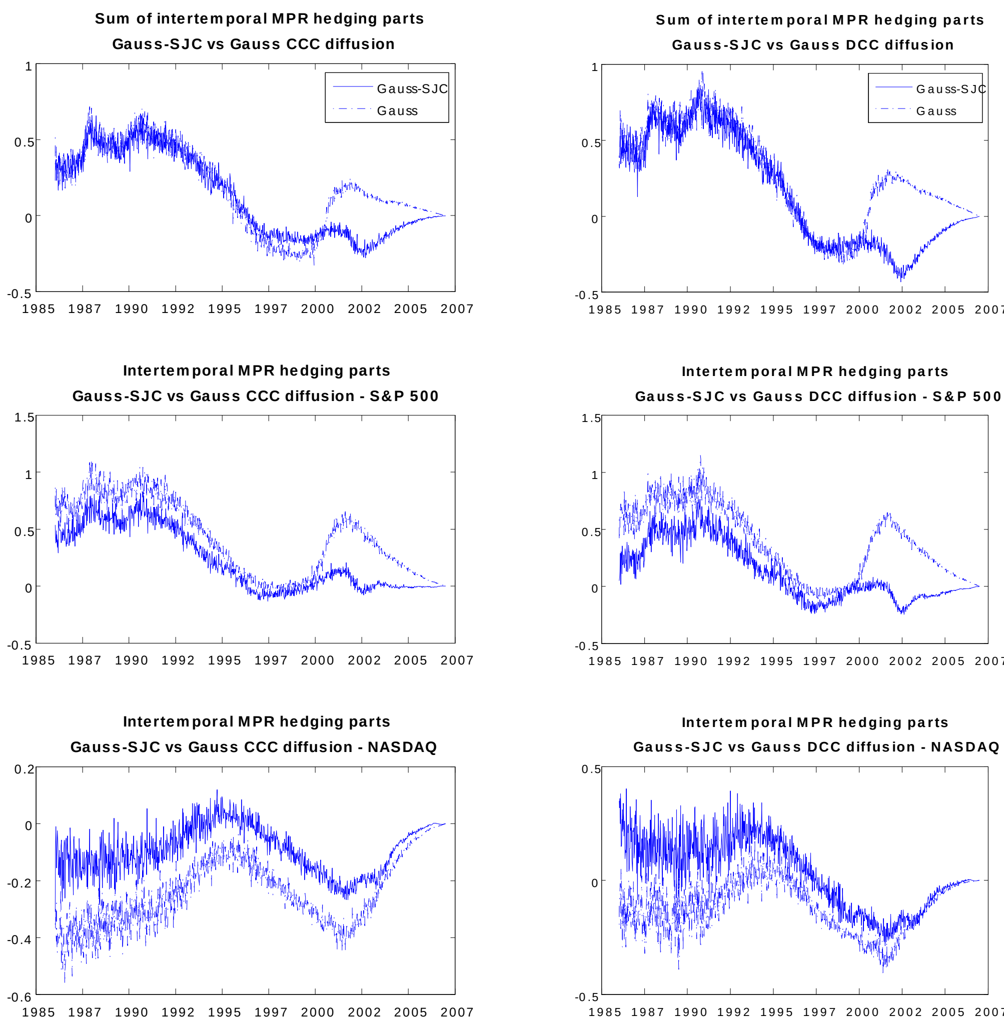


Figure 3.5.5. Panel C. Hedging demands due to differences in the unconditional distribution (tail dependence vs. no tail dependence) for the risky stocks for the 20-year fixed investment horizon for a CCC diffusion (left column) and a DCC diffusion (right column) (Case B)



demands. So while the portfolio term that is due to the need to hedge against stochastic changes in the observable factors driving correlation is indeed correlation hedging demand, the difference in the level of the market price of risk hedge terms between dynamic and constant conditional correlation is not entirely explained by this demand. Hence the possible disparity, even in sign, between the correlation hedging demands and the difference in the level of market price of risk hedges between constant and dynamic conditional correlation diffusions.

It is also of interest to contrast the differences in hedging demands due to dynamic correlation to those due to dynamic volatility, so we have reported on the right column of Figure 3.5.5, Panel A the results for a process for which we have assumed constant volatility and correlation (note that constant volatility is nested in the specification given in (3.3.15) and is achieved by setting the parameter κ to zero). Throughout the sample period the hedging demands for the constant volatility model are significantly higher than those with time-varying volatility, rendering the volatility effect much more pronounced than the effect of conditional correlation. The effect is qualitatively the same for a fixed and a rolling-window investment horizon. Unlike the correlation hedging demands, the hedging parts for the S&P 500 are increased when we allow for variations in volatility, while those of NASDAQ are significantly reduced for the whole period.

An alternative way to illustrate the importance of dynamically changing correlation on intertemporal hedging demands is to look at the induced portfolio holdings of S&P 500 from the position in NASDAQ, as explained in the previous section. On Panel B of Figure 3.5.5 we have plotted the induced MPR hedging demands for S&P 500 for a HARA investor with a 20-year investment horizon. We contrast the induced hedges for a DCC vs. a CCC model under two alternative unconditional distribution assumptions (Gaussian and Gaussian-SJC)[6]. Regardless of the form of the stationary density that we suppose, the induced hedging demands are lower for the DCC case then for the CCC one, pointing towards a reduction in the total portfolio holdings when dynamics of conditional correlation are explicitly accounted for.

Until now we have discussed the magnitude and sign of the hedging demands that arise

[6] Here we have reported results for DCC following Case B. All alternative cases of DCC were considered against the CCC model, and they all yield qualitatively similar results.

due to stochastic changes of the state variables driving conditional correlation which increases in down markets, volatile periods or bad states of the economy. An important question is whether there would be a similar shift in portfolio composition when the unconditional dependence structure is changed, that is the same stylized fact is reproduced through the stationary distribution of the process for the state variables X through a Gaussian copula (no tail dependence) or Gaussian-SJC copula (asymmetric tail dependence). On Figure 3.5.5, panel C we have reported the hedging demands of a Gaussian vs. a Gaussian-SJC diffusion under a CCC assumption, and the hedging demands of a Gaussian vs. a Gaussian-SJC diffusion under a DCC assumption for an investment horizon fixed at the end of the 20-year period. The presence of tail dependence changes the composition of the portfolio by reducing the absolute value of the intertemporal hedging terms. The latter are generally positive for S&P 500 and generally negative for NASDAQ, so tail dependence reduces in absolute value the holdings of both assets, driving them closer to zero. This result is maintained throughout the investment horizon, regardless of the way conditional correlation is modeled. Thus, for portfolio allocation, the impact of tail dependence through the unconditional distribution cannot be swept away by allowing conditional correlation to vary through time, rising in down markets.

The effect of tail dependence is somewhat subdued for the sum of the intertemporal hedges for both assets for the first half of the sample period, while towards the end of the period, mainly after 2000, the effect is more pronounced in the sense that the total intertemporal hedging demands are reduced for the case where we allow for tail dependence. It appears that for different subperiods of this relatively long sample hedging demands may have qualitatively different behavior. In order to gather more insight into the reasons behind differences in those demands, we concentrate our attention on two 8-year subperiods: one relatively calm in the sense of diminishing volatility, economy on the rise, low conditional correlation (1988-1996), and another period characterized by more hectic behavior in terms of high volatility, declining economic indicators and increased conditional correlation (1996-2004).

Correlation hedging for the two subperiods

Comparing the intertemporal hedging demands on Figure 3.5.6 and 3.5.7 for each one of the two subperiods, regardless of the assumptions we have made on the conditional correlation or the unconditional distribution, we see that those demands are generally positive throughout the first relatively calm period of economy on the rise and generally negative for the second hectic period of slowing down economy. There is just one exception to this rule that deserves attention - the hedging demands turn positive towards the second half of the 1996-2004 period for the Gaussian diffusion for both constant and dynamic specifications for the conditional correlation. Thus, failing to account for tail dependence increases the demand for the two risky funds and the fact that we allow for dynamically varying conditional correlation does not change this. It appears, following this preliminary observation, that unconditional dependence has a portfolio impact beyond the one induced by correlation hedging.

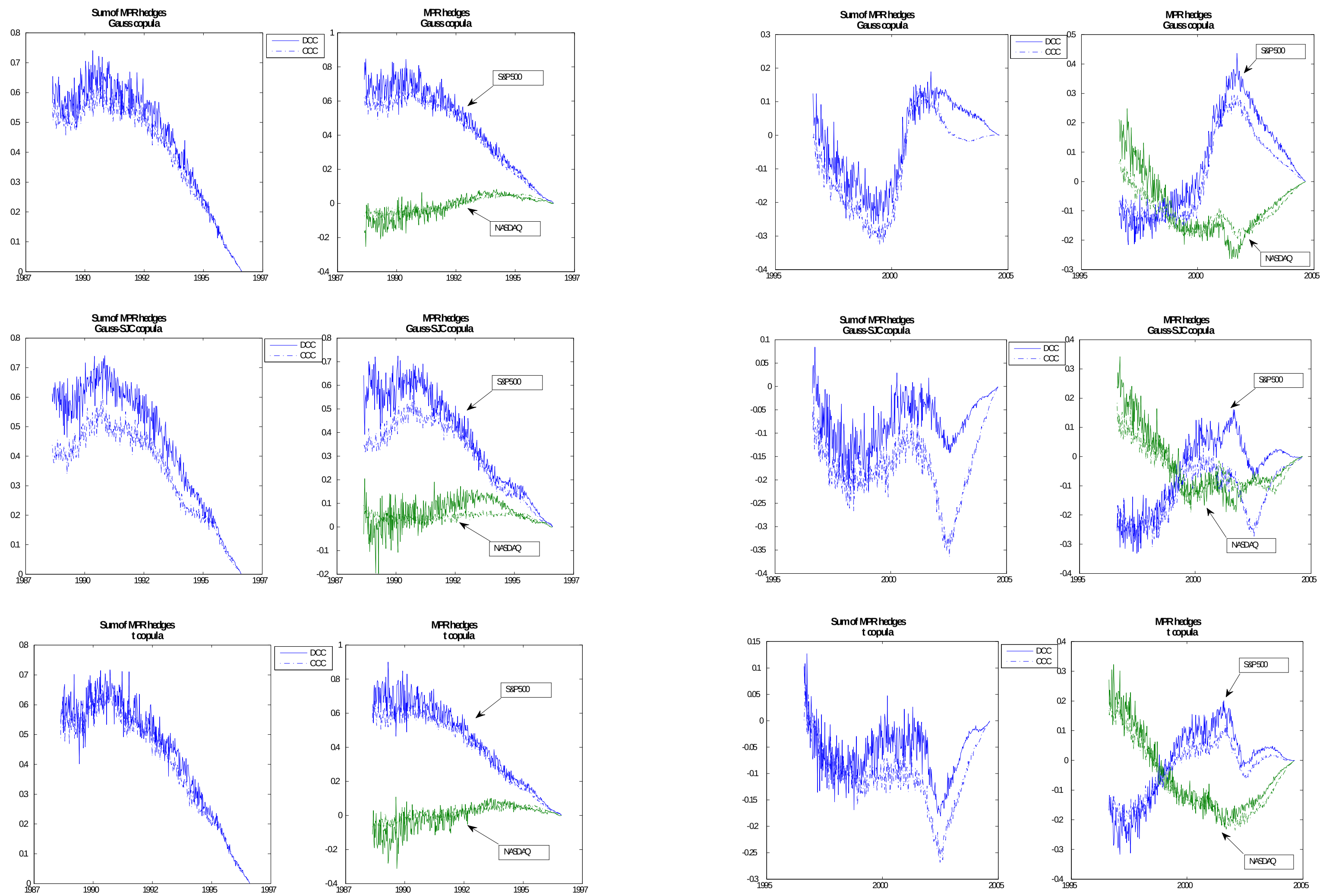
We now turn to a more detailed analysis of the portfolio implications of modeling conditional or unconditional dependence. The first comparison that we consider for the two chosen subperiods is one that is aimed at bringing forward the importance of correlation hedging through contrasting the intertemporal demands for the risky funds under a constant vs. a dynamic conditional correlation specification (for any of the three cases considered). To this end, we have plotted on Figure 3.5.6 the evolution of the hedges for a Gaussian, Gaussian-SJC and a t-diffusion for a HARA investor with a coefficient of relative risk aversion of 5.

For any of the unconditional distribution assumptions during the 1988-1996 period the presence of dynamically varying conditional correlation brings about increased hedging demands. When looking at the individual demands for any of the risky funds, we find that under the DCC assumption those demands are larger in absolute value, generally positive for S&P 500 and generally negative for NASDAQ. During the 1996-2004 period dynamic conditional correlation also leads to higher demands in absolute value for both funds, but the effect on the total hedging demands is more pronounced in the case when conditional correlation depends both on observable factors F and the level of the state variables X (Case B). In this case dynamic correlation leads to an increase in the total hedging demands. Results for conditional correlation specifications under Case A and C are qualitatively the same and

Figure 3.5.6: Market price of risk hedging terms for the two subperiods - effect of the conditional correlation

Panel A. (Case B).

The figure displays the intertemporal hedging demands of a HARA investor with RRA coefficient of 5 and an investment horizon fixed at the end of the period. The results of a DCC specification (Case B) are plotted against those coming from a CCC specification. All alternative unconditional distributions are considered (Gaussian, Gaussian-SJC, Student's t). The two left columns plot the results for the 1988-1996 period, the two right ones - the results for the 1996-2004 period.



are not reported for brevity.

Second, we consider the effect of the unconditional distribution on the hedging demands by comparing the results under the assumption of Gaussianity with those under the two alternatives of allowing for tail dependence - a Gaussian-SJC or a Student's t distribution. With this we aim to determine whether there is any portfolio effect induced by different assumptions on modeling tail dependence beyond the one incurred by dynamic conditional correlation.

On Panel A of Figure 3.5.7 we have plotted the hedging demands of a HARA investor with a relative risk aversion coefficient of 5 who models the stock price process using a Gaussian vs. a Gaussian-SJC diffusion (the effect of disregarding tail dependence) or alternatively a Student's t vs. a Gaussian-SJC diffusion (the effect of disregarding asymmetric tail dependence). In all cases we have constant conditional correlation. Contrary to the results on Figure 3.5.6 which tried to gauge the importance of modeling conditional correlation, here we have the opposite impact of the presence of tail dependence: it leads to smaller hedging demands in absolute value for both risky funds which reduces the total intertemporal demands for the risky assets. Those differences are more pronounced during the 1996-2004 period, and they are quite significant when the investor disregards tail dependence by assuming a Gaussian diffusion (in this case hedging demands grow to be positive in the second half of the period, whereas accounting for tail dependence both through the Gaussian-SJC and the t-diffusions leads to negative hedges).

However, when we allow for dynamically varying correlation some interesting results follow. Looking at Panel B on Figure 3.5.7, the large difference between the alternative unconditional distribution assumptions seems to vanish for the first subperiod. Allowing or not for tail dependence leads to virtually the same hedging demands. So, for this relatively calm period of improving economic conditions towards its end the presence of tail dependence does not lead to any significant change in the portfolio composition beyond the impact of correlation hedging. Still, the picture for the second highly volatile period is quite different. Accounting for tail dependence still leads to a decrease in absolute terms of the hedging components for both risky funds which generally leads to a decrease in the total hedging demand, especially for the Gaussian case. Thus, for a volatile period of deteriorating economic conditions tail dependence has a significant impact on the portfolio

Figure 3.5.7: Market price of risk hedging terms for the two subperiods - effect of the unconditional distribution
Panel A. (CCC).

The figure displays the intertemporal hedging demands of a HARA investor with RRA coefficient of 5 and an investment horizon fixed at the end of the period. The results of a Gaussian-SJC diffusion are plotted against those coming from a Gaussian diffusion (first row), or from a Student's t diffusion (second row). All alternative conditional correlation specifications come from the CCC model according to Case A. The two left columns plot the results for the 1988-1996 period, the two right ones - the results for the 1996-2004 period.

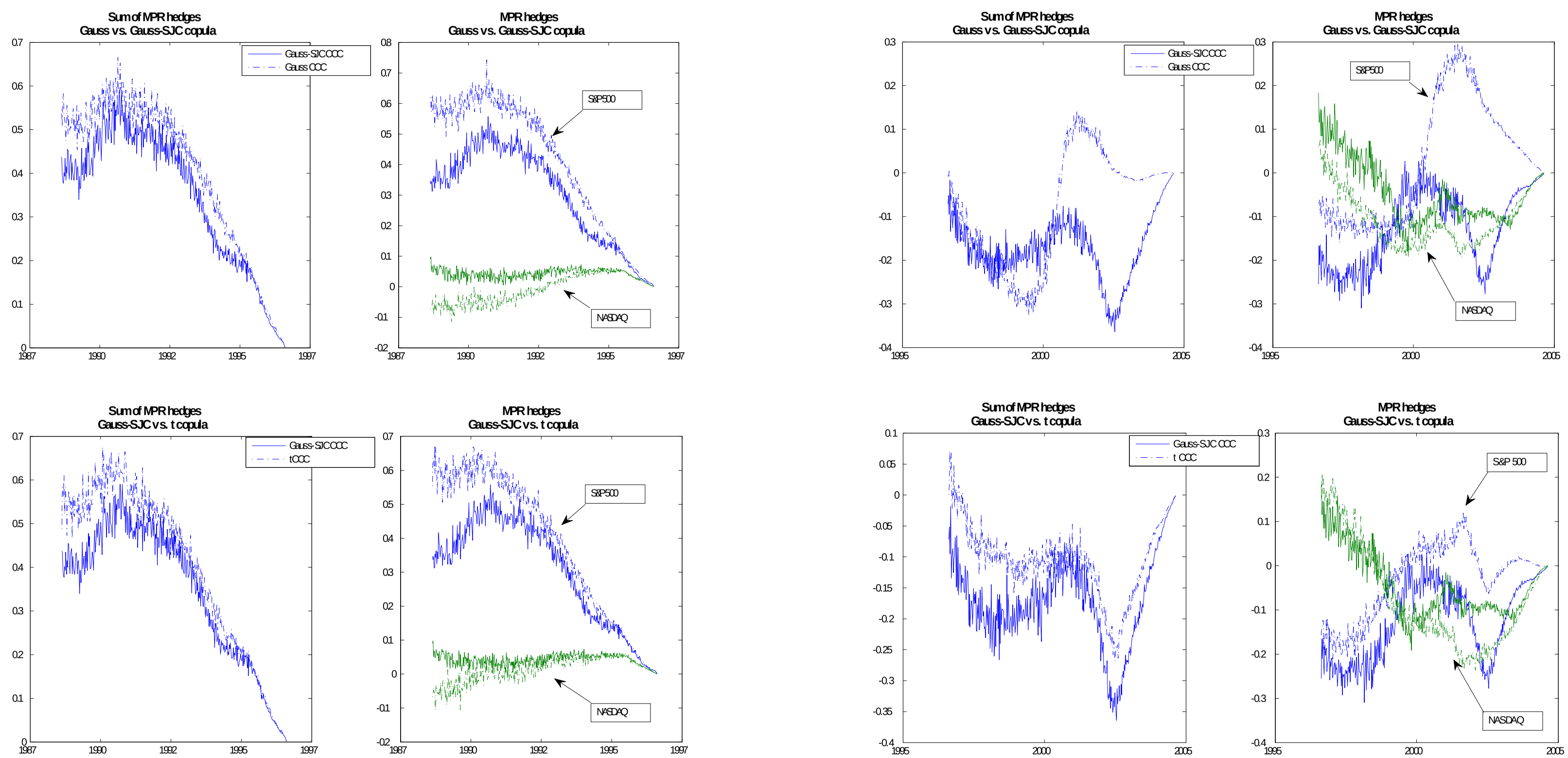
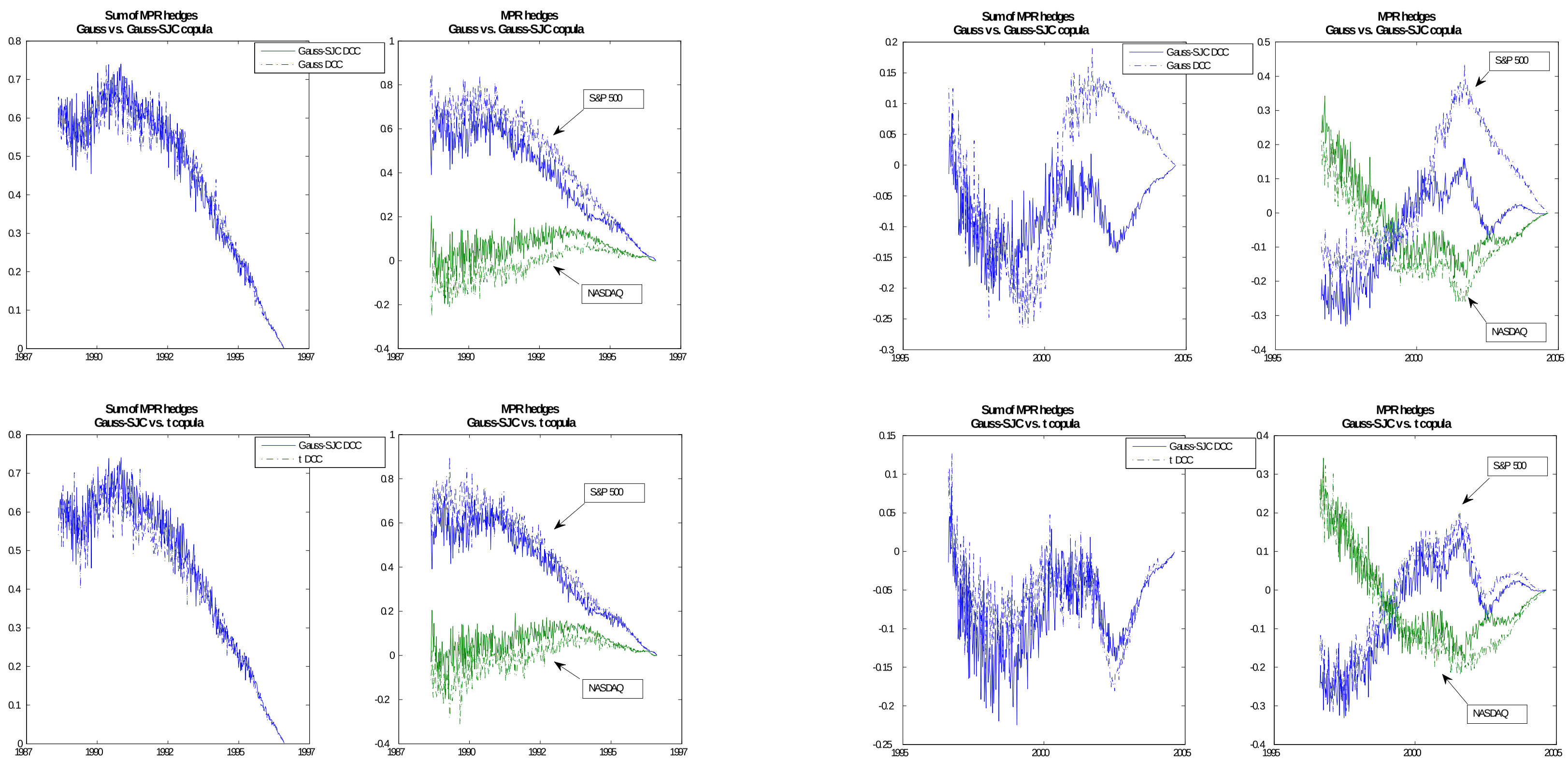


Figure 3.5.7. Panel B. (DCC, Case B)
The figure displays the intertemporal hedging demands of a HARA investor with RRA coefficient of 5 and an investment horizon fixed at the end of the period. The results of a Gaussian-SJC diffusion are plotted against those coming from a Gaussian diffusion (first row), or from a Student's t diffusion (second row). All alternative conditional correlation specifications come from the DCC model according to Case B. The two left columns plot the results for the 1988-1996 period, the two right ones - the results for the 1996-2004 period.



composition, even when dynamic conditional correlation has been accounted for.

3.5.2 Simulations

Having examined the distinct ways that dynamic conditional correlation or tail dependence influence the optimal portfolio decisions for a particular period and for realized paths of the state variables, we now turn to a simulations experiment that determines optimal portfolio shares for varying investment horizons while simulating ahead all the state variables involved. With this we aim to determine whether for the estimated parameters of the corresponding processes the relative importance of conditional and unconditional dependence on portfolio hedging demands will remain qualitatively the same as with the historical data considered.

Thus, we set up a first simulations exercise that aims at determining the importance of correlation hedging demands for a HARA investor who already believes that the process underlying stock prices has asymmetric tail dependence, incorporated through the Gaussian-SJC diffusion. Then we alternate the way to model conditional correlation by letting it be either constant or dynamic. In this way we can analyze the correlation hedging demands that arise beyond those that could be attributed to tail dependence through the unconditional distribution. We use the parameters estimated from a Gaussian-SJC process with DCC following Case B for the whole estimation period as a benchmark. Then, in order to obtain a CCC model, we set all parameters, driving conditional correlation, to zero, except for γ_0. We calibrate this parameter in order to reflect the same average correlation throughout the estimation period as the one implied by the benchmark process. In order to gauge the relative importance of adding each one of the observable factors to the dynamic correlation specification, we alternatively set either γ_1 (the VIX coefficient) or γ_3 (the CFNAI coefficient) to its corresponding value from the benchmark process, while setting all the other parameters to zero except γ_0 that is again calibrated in order to reflect the same average correlation. We then simulate ahead all the state variables involved in each of the four alternative processes, as well as their Malliavin derivatives, in order to obtain the Monte Carlo estimates of their conditional expectations in (3.3.22) and thus the intertemporal hedging demands. Results for investment horizons of 1 and 5 years are reported in Table 3.5.5, Panels A through C and Panel F.

The major conclusion that we may draw from those results is that for all investment

Table 3.5.5: Portfolio hedging terms through simulation

Intertemporal hedging demands for the 2 funds with different specifications for the conditional dependence through the conditional correlation (Gaussian-SJC diffusion with DCC modeled following Case B (observed factors and latent variables). For each investment horizon of 1 and 5 years the table displays the sum of market price of risk hedges (MPRH), the separate demands for S&P 500 and NASDAQ, as well as the two correlation hedges (CorrH), corresponding to the macroeconomic factor F^M and the volatility factor F^V (the latter multiplied by 100).

Panel A. Gaussian-SJC CCC diffusion $(\gamma_1 = \gamma_2 = \gamma_3 = 0)$										
Horizon	1 year					5 years				
	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100
CRRA, γ=5	0.1056	0.0614	0.0442	-	-	0.8623	0.6588	0.2035	-	-
CRRA, γ=10	0.0643	0.0370	0.0273	-	-	0.7341	0.5807	0.1534	-	-
HARA, γ=5,b=-0.2	0.0873	0.0505	0.0368	-	-	0.8201	0.6315	0.1885	-	-
HARA, γ=10,b=-0.2	0.0539	0.0307	0.0232	-	-	0.7113	0.5661	0.1451	-	-
Panel B. Gaussian-SJC DCC diffusion with only VIX driving conditional correlation $(\gamma_2 = \gamma_3 = 0)$										
Horizon	1 year					5 years				
	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100
CRRA, γ=5	0.1060	0.0613	0.0446	-	0.4368	0.8617	0.6575	0.2041	-	0.4103
CRRA, γ=10	0.0646	0.0370	0.0276	-	0.2815	0.7348	0.5809	0.1539	-	0.1492
HARA, γ=5,b=-0.2	0.0878	0.0505	0.0372	-	0.3693	0.8203	0.6314	0.1890	-	0.3392
HARA, γ=10,b=-0.2	0.0541	0.0307	0.0234	-	0.2429	0.7112	0.5657	0.1455	-	0.1078
Panel C. Gaussian-SJC DCC diffusion with only CFNAI driving conditional correlation $(\gamma_1 = \gamma_2 = 0)$										
Horizon	1 year					5 years				
	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100
CRRA, γ=5	0.0353	0.0620	-0.0267	-0.0673	-	0.8030	0.7259	0.0771	-0.0521	-
CRRA, γ=10	0.0204	0.0379	-0.0174	-0.0403	-	0.7356	0.6523	0.0833	0.0024	-
HARA, γ=5,b=-0.2	0.0285	0.0512	-0.0227	-0.0558	-	0.7798	0.7021	0.0777	-0.0373	-
HARA, γ=10,b=-0.2	0.0160	0.0313	-0.0153	-0.0338	-	0.7215	0.6371	0.0844	0.0101	-

Table 3.5.5 (cont.). Intertemporal hedging demands for the 2 stocks with different specifications for the unconditional dependence (Gaussian diffusion with no tail dependence, Student's t diffusion with symmetric tail dependence, and Gaussian-SJC diffusion with asymmetric tail dependence). Dynamic conditional correlation with observed factors (Case B). For each investment horizon of 1 and 5 years the table displays the sum of market price of risk hedges (MPRH), the separate demands for S&P 500 and NASDAQ, as well as the two correlation hedges (CorrH), corresponding to the macroeconomic factor F^M and the volatility factor F^V (the latter multiplied by 100).

Panel D. Gaussian DCC diffusion

Horizon	1 year					5 years				
	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100
CRRA, γ=5	0.0557	0.0652	-0.0094	-0.0469	0.5440	0.9605	0.7419	0.2187	-0.0127	0.1512
CRRA, γ=10	0.0374	0.0431	-0.0057	-0.0277	0.3209	0.8905	0.6694	0.2211	0.0237	-0.2819
HARA, γ=5,b=-0.2	0.0476	0.0555	-0.0078	-0.0387	0.4491	0.9353	0.7167	0.2186	-0.0026	0.0358
HARA, γ=10,b=-0.2	0.0321	0.0372	-0.0052	-0.0232	0.2673	0.8752	0.6559	0.2193	0.0292	-0.3400

Panel E. Student's t DCC diffusion

Horizon	1 year					5 years				
	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100
CRRA, γ=5	0.0442	0.0676	-0.0235	-0.0487	-0.1939	1.0418	0.7998	0.2419	-0.0670	-0.2690
CRRA, γ=10	0.0280	0.0424	-0.0145	-0.0283	-0.1128	0.9319	0.7114	0.2205	-0.0245	-0.1000
HARA, γ=5,b=-0.2	0.0364	0.0562	-0.0198	-0.0400	-0.1596	1.0045	0.7699	0.2346	-0.0538	-0.2182
HARA, γ=10,b=-0.2	0.0237	0.0362	-0.0125	-0.0236	-0.0940	0.9082	0.6915	0.2167	-0.0187	-0.0771

Panel F. Gaussian-SJC DCC diffusion

Horizon	1 year					5 years				
	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100	MPRH Sum	MPRH S&P500	MPRH NASDAQ	CorrH F^M	CorrH F^V*100
CRRA, γ=5	0.0268	0.0594	-0.0326	-0.0652	0.3782	0.8419	0.7265	0.1154	-0.0245	0.1429
CRRA, γ=10	0.0140	0.0356	-0.0216	-0.0386	0.2243	0.7708	0.6500	0.1208	0.0226	-0.1317
HARA, γ=5,b=-0.2	0.0206	0.0484	-0.0277	-0.0538	0.3129	0.8166	0.7007	0.1159	-0.0110	0.0664
HARA, γ=10,b=-0.2	0.0107	0.0297	-0.0190	-0.0323	0.1885	0.7552	0.6350	0.1202	0.0303	-0.1761

horizons considered, as well as for all degrees of relative risk aversion, the market price of risk hedge for the DCC model is the lowest. If we add only the macroeconomic factor to render conditional correlation dynamic, we get results that are quite close to the benchmark model. So for this application the macroeconomic factor seems to be the major driving force to determine the optimal portfolio composition. However, adding only the VIX factor does not change in any substantial way the portfolio holdings and they remain virtually unchanged with respect to the CCC model. As in the portfolio allocation example along realized paths of the state variables, here we also observe a larger spread between the holdings of S&P 500 and NASDAQ for the DCC case with respect to CCC. These results are confirmed for a CRRA as well as HARA investor and are valid for all investment horizons considered, as well as levels of risk aversion. Increasing the level of risk aversion invariably leads to a decrease in the intertemporal hedging demands in absolute terms. It also happens for a HARA investor with a certain subsistence level b below which she is unwilling to fall as compared to a CRRA investor.

A second simulations experiment that we consider aims at determining the importance of the stationary distribution and hence tail dependence for an investor who has already accounted for dynamically varying conditional correlation. We pick again the Gaussian-SJC diffusion with DCC according to Case B as the benchmark case and compare its implied hedging demands with those from a Gaussian or a Student's t alternative. Results are presented on Panels D through F of Table 3.5.5. As in the portfolio example over realized paths of the state variables, the stationary distribution still plays a role in determining the hedging demands, rendering them smaller in the presence of tail dependence. For smaller horizons its effect is smaller than the effect of disregarding conditional correlation, but at the 5-year horizon the Gaussian diffusion renders the highest hedging demands, even higher than the CCC case, which confirms our findings of the market timing exercise.

The above results may be sensitive to the level of conditional correlation that we impose. Thus, we repeat the simulations experiment with a Gaussian-SJC diffusion and DCC following Case B for varying values of the γ_0 parameter for the conditional correlation. For levels of γ_0 of 1, 2 and 3 obtain conditional correlation levels (averaged over the estimation period) of 0.45, 0.75 and 0.90. For each one of those DCC cases we find the appropriate CCC calibration for the conditional correlation parameters, keeping the same average correlation levels. Results are plotted on Figure 3.5.8.

Regardless of the investment horizon, for relatively low correlation levels (0.45) the DCC model implies significantly lower intertemporal hedging demands, compared to a CCC specification, even after tail dependence has been accounted for through the Gaussian-SJC stationary distribution. For extremely high correlation levels (the case of $\gamma_0 = 3$) the roles of DCC and CCC change and now it is the latter that implies lower hedging demands. Depending on the investment horizon, we may have higher or lower hedge levels for a mean conditional correlation of 0.75. This behavior can thus explain the higher hedging demands implied by the DCC specification over a realized path of the state variables that we encountered earlier.

3.5.3 Certainty equivalent cost of ignoring correlation hedging

We follow the common approach in literature on portfolio choice and study the effect of ignoring correlation hedging on the wealth of the investor using the utility loss, or the certainty equivalent cost (see Liu et al., 2003). The approach consists in computing the additional amount of wealth that would be needed for an investor to consider a suboptimal allocation strategy (that results from ignoring correlation hedging) instead of the optimal one (that takes into account the dynamics of conditional correlation), in order to achieve the same expected utility of terminal wealth. In other words, we are looking to determine the amount *ceq* such that:

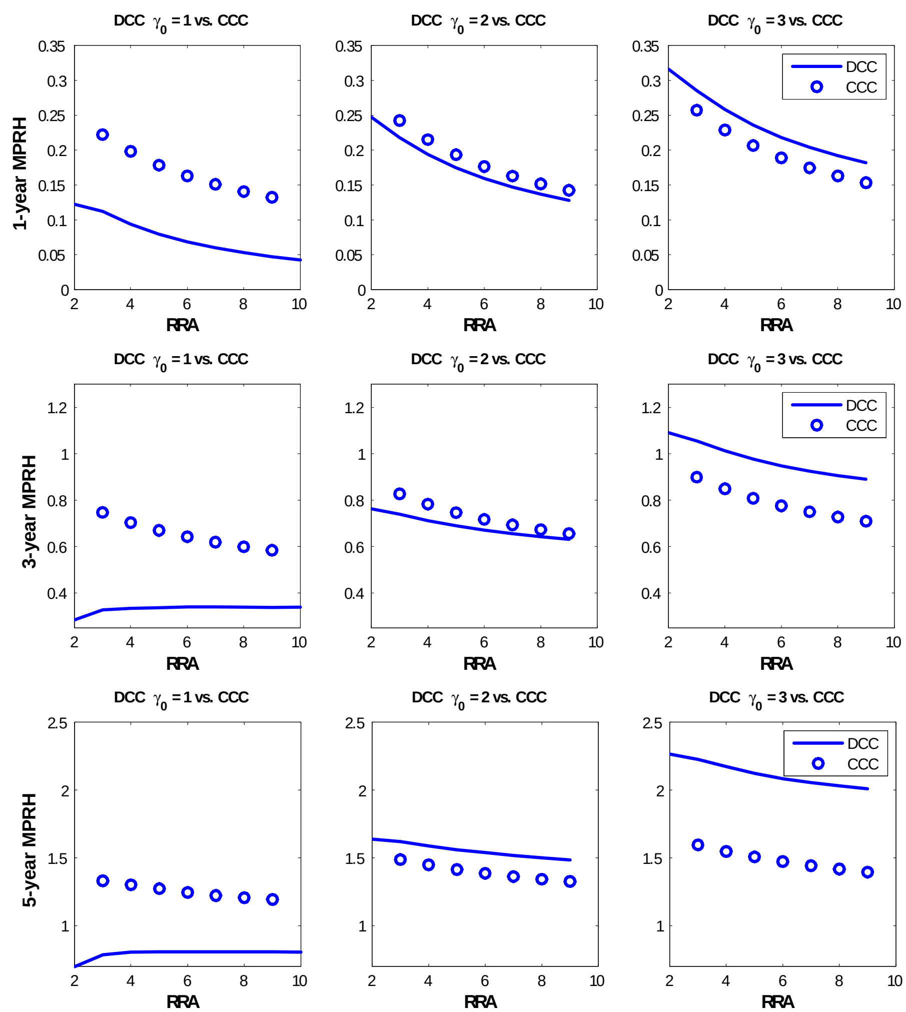
$$E\left[U\left(\omega_T^* \mid \omega_0 = 1\right)\right] = E\left[U\left(\omega_T \mid \omega_0 = 1 + ceq\right)\right]$$

where ω_T^* is the terminal wealth achieved under the optimal investment strategy and ω_T is the terminal wealth under the suboptimal one.

The first question that we address, in accordance with the simulation exercise above, is whether the investor would lose anything if she disregards the dynamics of conditional correlation, modeled using observable factors, given the fact that tail dependence in the unconditional distribution has already been accounted for. Thus, we choose as a benchmark process the Gaussian-SJC diffusion with DCC according to Case B. Then we alternate between setting all conditional correlation parameters to zero except for γ_0 (CCC alternative),

Figure 3.5.8: Dynamic correlation-induced portfolio hedging terms through simulation: the influence of correlation level

Intertemporal hedging demands for a benchmark Gaussian-SJC diffusion with DCC (Case B) vs. a CCC specification with parameter calibrated to match the mean conditional correlation of the corresponding DCC model. Varying average values of conditional correlation through the parameter γ_o. HARA investor with $b = -0.2$ and varying degrees of relative risk aversion, and investment horizon of 1, 3 and 5 years.



letting only γ_3 be zero (conditional correlation being driven by the VIX factor), or letting γ_1 be nonzero (conditional correlation being driven by the macroeconomic factor). In those alternative models the γ_0 parameter is calibrated in order to reflect the same average correlation as the DCC benchmark over the estimation horizon. We consider again a HARA investor with varying degrees of relative risk aversion and a parameter b in the utility function equal to -0.2, 0 or 0.2. The case of $b = 0$ corresponds to a CRRA investor, while if $b < 0$ relative risk aversion is decreasing and convex in wealth, in which case the investor is intolerant towards wealth falling below a certain subsistence level $-b$, and alternatively, if $b > 0$, then relative risk aversion is increasing and concave. Table 3.5.6 summarizes the results on the certainty equivalent cost in each case, calculated in cents per dollar.

The cost of disregarding the dynamics of conditional correlation is comparable to the cost of disregarding the presence of the macroeconomic factor driving its dynamics, so we may conclude that the CFNAI factor is the major player in the present setting in terms of utility loss. The cost decreases with rising levels of the risk aversion coefficient, and is highest for a HARA investor with relative risk aversion that is increasing and concave in wealth. However, the impact of disregarding the VIX factor is almost insignificant.

We next address the alternative problem of finding the utility cost for an investor who disregards the fact that extreme realizations of the assets in her portfolio may be dependent, as modeled through the stationary distribution of X. Results are summarized in Table 3.5.7, where we take as a benchmark process either the DCC Gaussian-SJC diffusion (left column), or the CCC one (right column) against the two Elliptic counterparts. In order to isolate only the impact of the tail dependence through the stationary distribution, conditional correlation parameters for all processes are taken from the Gaussian-SJC type with DCC (Case B).

The main conclusion that we can draw from comparing the wealth loss across the alternative specifications is that the investor loses more from disregarding tail dependence if she has not taken into account the dynamics in conditional correlation. It is an anticipated result, as both ways of modeling dependence through the dynamics of the conditional correlation or through the stationary distribution aim at reproducing the same stylized fact of increased dependence in down markets. Thus if at least one of them is taken into account when making portfolio decisions, the impact of disregarding the other in terms of wealth

Table 3.5.6: Certainty equivalent cost of ignoring dynamic conditional correlation, modeled with observable factors

The benchmark process is a Gaussian-SJC diffusion with DCC according to Case B. All of the alternative processes have a Gaussian-SJC stationary distribution, but their conditional correlation specifications vary from CCC to DCC with no VIX ($\gamma_1 = 0$), and DCC with no CFNAI factor ($\gamma_2 = 0$).All parameters of the stationary distribution are from the Gaussian-SJC type with DCC (Case B), the conditional correlation parameters of the alternative processes were calibrated in order to reflect the same mean conditional correlation as the benchmark process. The Certainty Equivalent Cost is given in cents per dollar. Investment horizon is 5 years.

Panel A. The cost of disregarding DCC (CCC alternative)			
	HARA, $b = -0.2$	CRRA	HARA, $b = 0.2$
$\gamma = 2$	2.3054	2.4039	2.5024
$\gamma = 4$	1.8987	1.9369	1.9751
$\gamma = 6$	1.7983	1.8216	1.8449
$\gamma = 8$	1.7538	1.7706	1.7873
$\gamma = 10$	1.7289	1.7419	1.7549
Panel B. The cost of disregarding the CFNAI factor (DCC with $\gamma_2 = 0$ alternative)			
	HARA, $b = -0.2$	CRRA	HARA, $b = 0.2$
$\gamma = 2$	2.4273	2.5533	2.6792
$\gamma = 4$	1.9309	1.9832	2.0355
$\gamma = 6$	1.7988	1.8315	1.8643
$\gamma = 8$	1.7384	1.7622	1.7860
$\gamma = 10$	1.7039	1.7226	1.7413
Panel C. The cost of disregarding the VIX factor (DCC with $\gamma_1 = 0$ alternative)			
	HARA, $b = -0.2$	CRRA	HARA, $b = 0.2$
$\gamma = 2$	0.0000	0.0000	0.0000
$\gamma = 4$	0.0000	0.0000	0.0000
$\gamma = 6$	0.0000	0.0000	0.0000
$\gamma = 8$	0.0000	0.0000	0.0000
$\gamma = 10$	0.0000	0.0000	0.0000

Table 3.5.7: Certainty equivalent cost of ignoring tail dependence

The benchmark process is a Gaussian-SJC diffusion with DCC according to Case B. The alternative processes have either a DCC specification (left figures) or a CCC specification (right column), and their unconditional distribution varies from Gaussian to Student's t. All parameters of the conditional correlation specification are from the Gaussian-SJC type with DCC (Case B) (left column) and from Gaussian-SJC type with CCC (right column). The Certainty Equivalent Cost is given in cents per dollar. Investment horizon is 5 years.

Panel A. The cost of disregarding tail dependence

	(Gaussian alternative, DCC)			(Gaussian alternative, CCC)		
	HARA	CRRA	HARA	HARA	CRRA	HARA
	b=-0.2	b=0	b=0.2	b=-0.2	b=0	b=0.2
$\gamma = 2$	1.3153	1.5158	1.7162	3.2467	3.8692	4.4916
$\gamma = 4$	0.6384	0.7438	0.8492	1.1366	1.4361	1.7357
$\gamma = 6$	0.3912	0.4619	0.5326	0.4602	0.6562	0.8523
$\gamma = 8$	0.2658	0.3189	0.3719	0.1301	0.2757	0.4212
$\gamma = 10$	0.1902	0.2327	0.2751	0.0000	0.0507	0.1664

Panel B. The cost of disregarding asymmetric tail dependence

	(Student's t alternative, DCC)			(Student's t alternative, CCC)		
	HARA	CRRA	HARA	HARA	CRRA	HARA
	b=-0.2	b=0	b=0.2	b=-0.2	b=0	b=0.2
$\gamma = 2$	0.1886	0.1696	0.1506	0.5891	0.6486	0.7081
$\gamma = 4$	0.4271	0.4416	0.4561	0.4755	0.5176	0.5597
$\gamma = 6$	0.4259	0.4403	0.4546	0.3960	0.4260	0.4559
$\gamma = 8$	0.4121	0.4245	0.4369	0.3509	0.3740	0.3970
$\gamma = 10$	0.3999	0.4106	0.4213	0.3224	0.3411	0.3598

loss will be subdued.

As we saw in the above simulations exercise, the portfolio composition changes considerably for varying levels of the mean conditional correlation, modeled through the parameter γ_0. In order to determine the economic significance of this finding, we determine the certainty equivalent cost for disregarding correlation dynamics for any of the three cases that we considered at the end of the previous section. Results are summarized on Panel A of Figure 3.5.9.

The certainty equivalent cost is lower for the lowest levels of correlation considered ($\gamma_0 = 1$ or average correlation of 0.45 over the estimation horizon) and increases significantly for higher correlation levels. It also increases with the investment horizon. Results are consistent over the utility specifications considered (CRRA and 2 types of HARA utility).

For the above cases we have considered the Case B DCC specification as a benchmark, that is the case when dynamic conditional correlation is driven by both the observable factors F and the state variables X. In order to gauge the economic importance of any of the other DCC specifications, we calculate the wealth loss of an investor who believes that conditional correlation is either driven exclusively by observed factors (Case C) or they do not enter correlation dynamics (Case A), instead of the benchmark Case B. Results for an investment horizon of 5 years are summarized on Panel B on Figure 3.5.9. We find that the difference in terms of wealth loss between cases B and C is negligible, that is the investor does not lose much by just considering the observed factors for the dynamics of conditional correlation. The loss for an investor who totally disregards observed factors is higher, especially for low levels of risk aversion. But for extremely risk averse investors there is virtually no cost for considering any of the alternative DCC models instead of the benchmark one.

Being consistent with the simulations experiment, we consider also the economic loss for disregarding tail dependence, given that the dynamics of conditional correlation have been accounted for. We compute it by comparing the benchmark Gaussian-SJC diffusion with DCC according to Case B with a corresponding Gaussian diffusion with the same correlation dynamics. We do so for varying weights ω of the mixture copula $C^{Ga-SJC} = \omega C^{SJC} + (1-\omega)\, C^{Ga}$. Parameters are taken from the benchmark model over the whole estimation horizon, and the Gaussian correlation parameter is set so that the Kendall's tau implied by the Gaussian copula is equal to the one implied by the SJC copula, so varying

Figure 3.5.9: Certainty Equivalent Cost

Panel. A. Certainty Equivalent Cost of ignoring dynamic conditional correlation, modeled with observable factors for varying mean levels of conditional correlation

The certainty equivalent cost of disregarding dynamic conditional correlation for a benchmark Gaussian-SJC diffusion with DCC (Case B) vs. a Gaussian diffusion with CCC with parameter calibrated to match the mean conditional correlation of the corresponding DCC model. Varying average values of conditional correlation through the parameter γ_o. HARA investor with $b = -0.2$ and varying degrees of relative risk aversion, and investment horizon of 1, 3 and 5 years.

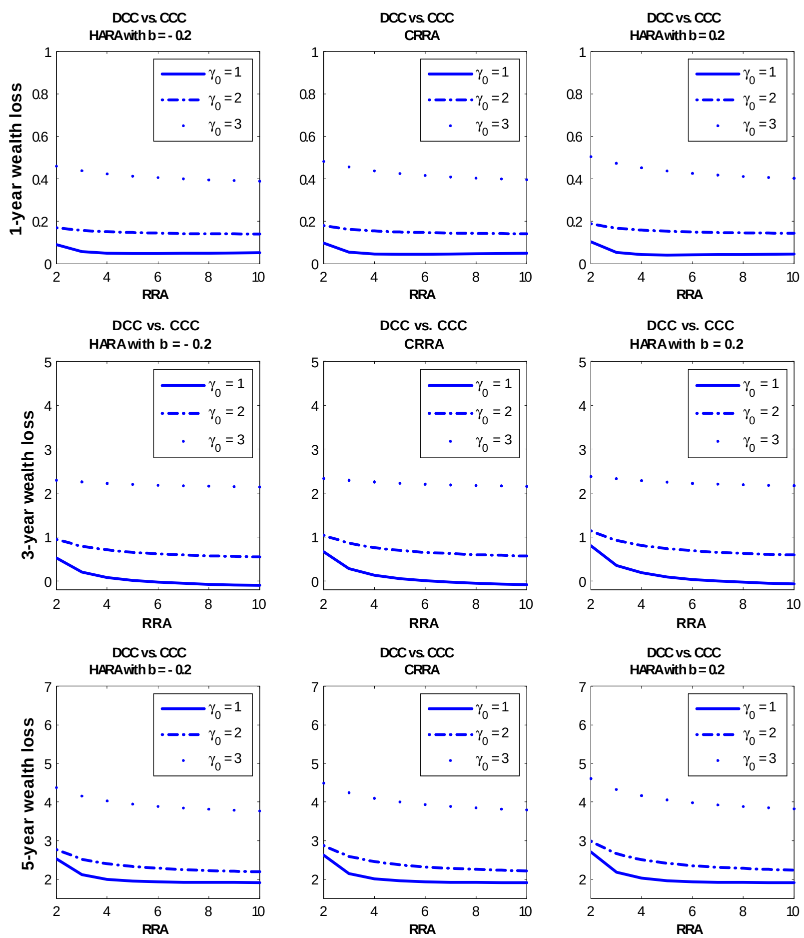
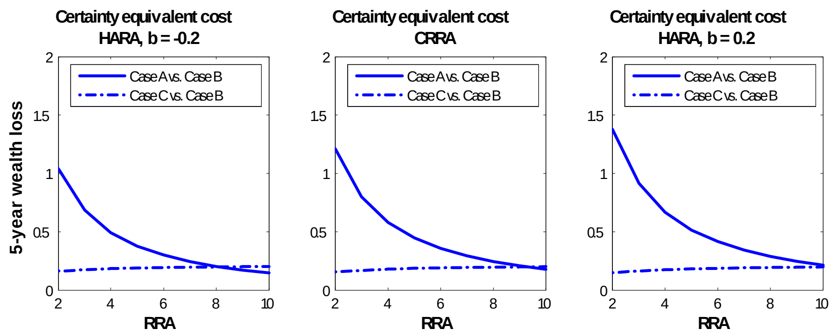


Table 3.5.9. Panel. B. Certainty Equivalent Cost of using alternative DCC specifications The certainty equivalent cost of modeling DCC following Case A or C vs. the benchmark case B for a Gaussian-SJC diffusion. 5-year investment horizon. Parameters for cases A and C are calibrated so as to reflect the same average conditional correlation over the estimation period as that implied by the benchmark case.



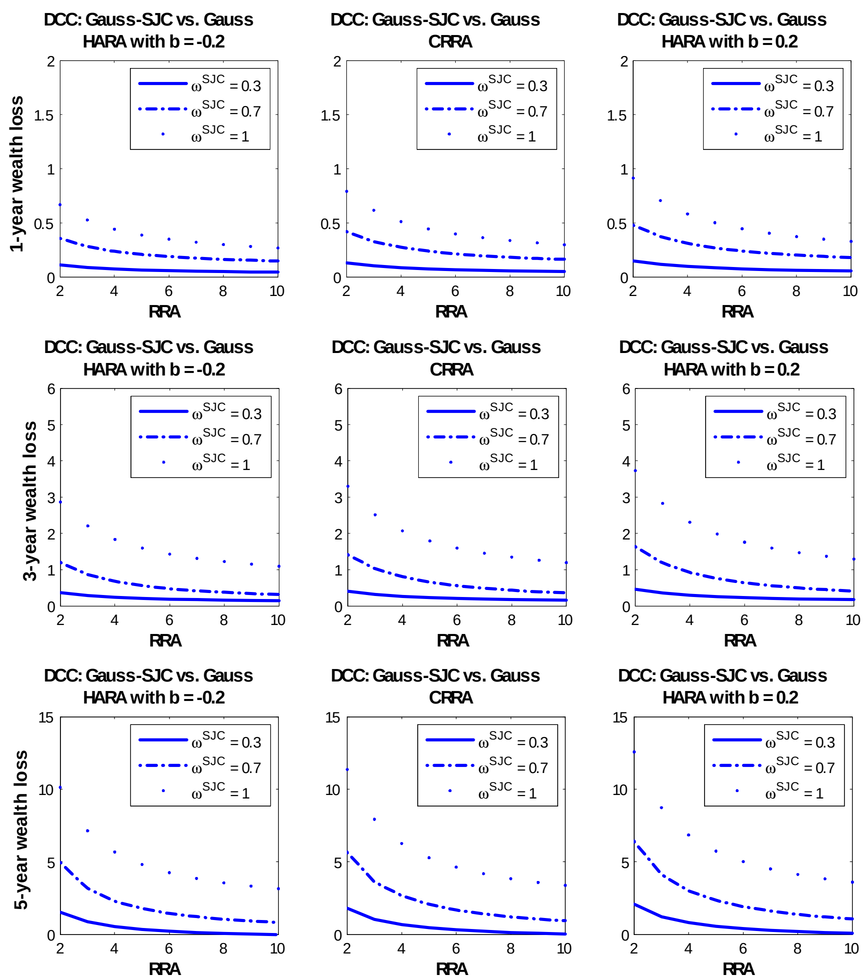
the composition of the Gaussian-SJC copula will not change the Kendall's tau, but only the relative importance of tail dependence. Results are presented on Panel C on Figure 3.5.9. Even if dynamic conditional correlation has already been accounted for, there are substantial economic costs for disregarding tail dependence, reaching over ten cents per dollar for a 5-year investment horizon. They increase with increasing the weight of the SJC copula in the benchmark model (and hence the importance of tail dependence in the data generating process), and are higher for investors with lower levels of risk aversion.

3.6 Conclusion

In this chapter we address the issue of determining the impact of dynamic correlation modeled through observable factors on the portfolio hedging demands. The solution methodology that we apply allows us to disentangle the intertemporal demands due to the need to hedge against stochastic changes in those factors from the rest of the market price of risk hedging terms. We also account for tail dependence that manifests itself through increased co-movements between risky stocks during sharp market downfalls. We find that demands for correlation hedging and intertemporal demands due to high tail dependence have a distinct impact on the optimal portfolio behavior both in terms of optimal portfolio composition and of loss of wealth criterion.

There are a number of ways in which the present study could be extended. First, we could test the sensitivity of the results to an increased number of assets in the portfolio, as we would expect that hedging demands should increase as a result of the higher level of

Table 3.5.9. Panel. C. Certainty Equivalent Cost of disregarding tail dependence
The certainty equivalent cost of disregarding tail dependence by considering a Gaussian DCC diffusion instead of the benchmark data generating process of a Gaussian-SJC DCC diffusion for varying levels of the ω^{SJC} parameter determining the weight of the SJC copula in the mixture distribution. DCC specification follows Case B. Parameters are taken from estimating the benchmark case over the whole estimation horizon, while the correlation parameter of the Gaussian copula is calibrated so that to reflect the same Kendall's tau as the one implied by the SJC copula with the estimated parameters.



uncertainty linked to both the conditional correlation structure and the dependence through the stationary distribution. Second, it would be of interest to extend the dynamic treatment to the dependence structure modeled by the copula, assumed to be fixed in the present setup, in the spirit of dynamic copula models as in Patton (2004) . By letting observable factors affect the evolution of tail dependence we may find similar hedging demands as those implied by dynamic correlation. As well, we have seen that the dependence structure changes dramatically from relatively calm periods of low volatility and rising economic conditions, when it is not far from Gaussian to highly volatile periods marked with recessionary states, when dependence exhibits asymmetries and high tail coefficients. This could motivate us to consider a specification where the copula composition changes from normal to extreme value dependent one through varying weights of the copula.

Finally, for the sake of simplicity, we have assumed so far that the bond and stock dynamics are independent from each other. As there is compelling evidence of co-movement between bond and stock returns that could be linked to common exposure to macroeconomic factors (e.g. Li, 2002), it would be of interest to incorporate this finding in the present portfolio solution setup.

Appendix A

For Chapter 2

A.1 Copula functions

In this chapter we have used he following d-dimensional copula functions.

- **Gaussian copula**

$$\begin{aligned} & C^{Ga}\left(u_1, u_2, ..., u_d \mid R_{Ga}\right) \\ = & \int\limits_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int\limits_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{2\pi \left|R_{Ga}\right|^{1/2}} \exp\left\{-\frac{1}{2} x^{\intercal} R_{Ga}^{-1/2} x\right\} dx_1 ... dx_d \end{aligned}$$

where R_{Ga} denotes the correlation matrix, and $\Phi^{-1}\left(u_i\right)$ is the inverse of the univariate standard normal CDF.

- **Student's t copula**

$$\begin{aligned} & C^{t}\left(u_1, u_2, ..., u_d \mid R_t, \nu\right) \\ = & \int\limits_{-\infty}^{t_\nu^{-1}(u_1)} \cdots \int\limits_{-\infty}^{t_\nu^{-1}(u_d)} \frac{\Gamma\left(\frac{\nu+d}{2}\right) \left|R_t\right|^{1/2}}{\Gamma\left(\frac{\nu}{2}\right) \left(\nu\pi\right)^{d/2}} \left(1 + \frac{1}{\nu} x^{\intercal} R_t^{-1} x\right)^{-\frac{\nu+d}{2}} dx_1 ... dx_d \end{aligned}$$

where R_t denotes the correlation matrix, ν is the degrees of freedom parameter, and $t^{-1}\left(u_i\right)$ is the inverse CDF of the univariate Student's t distribution with ν degrees of freedom.

- **Archimedean copulas**

Copulas in this family are constructed using a continuous and strictly decreasing generator function $\varphi\left(u\right) : [0, 1] \to [0, \infty)$:

$$C\left(u_1, u_2, ..., u_n\right) = \varphi^{-1}\left(\varphi\left(u_1\right) + \varphi\left(u_2\right) + ... + \varphi\left(u_n\right)\right)$$

The generator for the Gumbel copula is given by $\varphi\left(x\right) = \left(-\log\left(x\right)\right)^{\frac{1}{\alpha}}, \alpha \in (0, 1]$, and consequently its form is as follows:

$$C_\alpha^G\left(u_1, u_2, ..., u_n\right) = \exp\left(-\left(\sum_{i=1}^{n} \left(-\log u_i\right)^{\frac{1}{\alpha}}\right)^{\alpha}\right), \quad \alpha \in (0, 1]$$

for a dependence parameter α, common across all random variables. The survival counterpart of the Gumbel copula for the bivariate case is given by:

$$\begin{aligned}
\overline{C}^G_{\overline{\alpha}}(u,v) &= u+v-1+\exp\left(-\left[(-\log(1-u))^{\frac{1}{\overline{\alpha}}}+(-\log(1-v))^{\frac{1}{\overline{\alpha}}}\right]^{\overline{\alpha}}\right), \\
\overline{\alpha} &\in (0,1]
\end{aligned}$$

for a dependence parameter $\overline{\alpha}$. See Theorem 4.7 in Cherubini et al. (2004) for dimensions bigger than 2.

The nested copula construction that we consider consists in consequently nesting bivariate copulas within each other. Thus, for the tri-variate case the copula has the form:

$$C(u_1,u_2,u_3)=\varphi_2^{-1}\left(\varphi_2\left(\varphi_1^{-1}\left(\varphi_1(u_1)+\varphi_1(u_2)\right)\right)+\varphi_2(u_3)\right)$$

where each generating function $\varphi_i(u_i)$ has its own dependence parameter α_i. With this construction we achieve $(n-1)$ different pairs of variables that have distinct dependence, which are still below the general case, but is a considerable improvement compared to the case of homogenous dependence above. The parameters α_i should satisfy certain conditions in order for the above function to be indeed a copula (see Embrechts et al. (2002) for a discussion). For the Gumbel copula this condition amounts to verifying the following: $\alpha_1 \leq \alpha_2$, i.e. dependence should be higher in the more deeply nested copulas (note that for the above parameterization of the Gumbel copula dependence increases for decreasing values of the parameter α).

A.2 Form, properties and subclasses of the univariate Generalized Hyperbolic family of distributions

The family of GH distributions is constructed as normal mean-variance mixtures with the Generalized Inverse Gaussian (GIG) as the mixing distribution. Its probability density function is given by:

$$\begin{aligned}
f_{GH}(x;\alpha,\beta,\delta,\mu) &= c(\lambda,\alpha,\beta,\delta)\left(\delta^2+(x-\mu)^2\right)^{\frac{\lambda-1/2}{2}}\times \\
&\quad K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)e^{\beta(x-\mu)} \\
\text{where } c(\lambda,\alpha,\beta,\delta) &= \frac{\left(\alpha^2-\beta^2\right)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}K_{\lambda}\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \\
x &\in \mathbb{R}
\end{aligned}$$

where $c(\lambda,\alpha,\beta,\delta)$ is the normalizing constant and K_λ is the modified Bessel function of the third kind with index λ, defined as :

$$K_\lambda(x)=\frac{1}{2}\int_0^\infty y^{\lambda-1}e^{-\frac{x}{2}\left(y+y^{-1}\right)}dy,\quad x>0$$

The parameters have the following interpretations in terms of the shape of the distribution: α determines the shape, β - the skewness, μ is a location parameter and δ is a scaling parameter. The parameter domain is:

$$\begin{aligned}
\delta &\geq 0, \alpha > |\beta| \text{ for } \lambda > 0 \\
\delta &> 0, \alpha > |\beta| \text{ for } \lambda = 0 \\
\delta &> 0, \alpha \geq |\beta| \text{ for } \lambda < 0 \\
\mu &\in \mathbb{R}
\end{aligned}$$

The GH family of distributions has the normal distribution as a limiting case for $\delta \to \infty$, $\delta/\alpha \to \sigma^2$, and the Student's t distribution as a limit for $\lambda < 0$, $\alpha = \beta = \mu = 0$ (Barndorff-Nielsen, 1978; Prause, 1999).

Various special cases can be obtained for different parameterization of the GH distribution. For $\lambda = -1/2$ we obtain the Normal Inverse Gaussian (NIG) distribution, whose density is given by:

$$\begin{aligned}
f_{NIG}(x;\alpha,\beta,\delta,\mu) &= c(\alpha,\delta)\left(\delta^2 + (x-\mu)^2\right)^{\frac{1}{2}} \times \\
&\quad K_1\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)} \\
\text{where } c(\alpha,\delta) &= \frac{\alpha\delta}{\pi} \\
x &\in \mathbb{R}
\end{aligned}$$

where $\delta > 0$, $\alpha \geq |\beta| \geq 0$, $\mu \in \mathbb{R}$. Its tail behavior is given by

$$\lim_{x\to\pm\infty} f_{NIG}(x;\alpha,\beta,\delta,\mu) \sim |x|^{-3/2}\, e^{(\mp\alpha+\beta)x}$$

and it has the interesting property of being closed under convolution, so that the sum of two independent random variables that have a NIG distribution $X_i \sim NIG(x;\alpha,\beta,\delta_i,\mu_i), i = 1,2$ is also NIG-distributed: $X_1 + X_2 \sim NIG(x;\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2)$.

In the portfolio application we use several properties of the modified Bessel function that we summarize bellow (following Bibby and Sorensen, 2003):

$$\begin{aligned}
K_{-\lambda}(x) &= K_\lambda(x) \\
K'_\lambda(x) &= -\frac{\lambda}{x}K_\lambda(x) - K_{\lambda-1}(x) \\
K_{n+\frac{1}{2}}(x) &= \sqrt{\frac{\pi}{2x}}\exp(-x)\left[1 + \sum_{i=1}^{n}\frac{(n+i)!}{(n-i)!i!}(2x)^{-i}\right], \quad n = 1,2,3,...
\end{aligned}$$

A.3 The Sequential Markov Chain Monte Carlo estimation algorithm

The algorithm for carrying out the Metropolis-Hastings scheme for sampling from the conditional posterior of parameters and latent data following Golightly and Wilkinson (2006a) can be summarized as follows:

Consider a d-dimensional Itô diffusion given by:

$$dY_t = \mu(Y_t)\,dt + \sigma(Y_t)\,dW_t$$

Let data be observed at times $t_0 < t_1 < ... < t_{n-1} < t_n$ with $\Delta\tau = t_{i+1} - t_i$. We divide each subinterval between observations in equidistant points, so that the augmented data matrix looks like:

$$Y^{aug} = [\ \overline{\overline{Y}}_{t_0,0} \quad Y_{t_0,1} \quad ... \quad Y_{t_0,m} \quad \overline{Y}_{t_1,0} \quad ... \quad \overline{Y}_{t_{n-1},0} \quad ... \quad Y_{t_{n-1},m} \quad \overline{Y}_{t_n,1}\ ],$$

$Y_{t_i,j}$ is a d-dimensional vector of latent data points at time $t_i + j\Delta\tau$ and $\overline{\overline{Y}}_{t_i,0}$ is the vector of observations at time t_i.

Initialization. Set $j = 0$. Initialize the augmented data points for each of the $s = 1, ..., MC$ iterations by linearly interpolating between observations for the first interval. Initialize the parameter set for all s by sampling from a prior density $\pi(\theta)$.

1. For each $s = 1, ..., MC$:

- Propose the parameters θ^* using a kernel density estimate of the marginal parameter posterior $\pi\left(\theta \mid \overline{Y}_{t_j}\right)$ with the kernel shrinkage correction of Liu and West (2001):

$$\begin{aligned}
\theta^* &\sim \phi\left(\alpha\theta_u + (1-\alpha)\overline{\theta}, h^2 V\right) \\
\alpha^2 &= 1 - h^2 \\
h^2 &= 1 - \left((3\delta - 1)/2\right)^2
\end{aligned}$$

for a discount factor δ, where ϕ denotes the Gaussian density, and u is an integer that has been drawn uniformly from $\{1, ..., MC\}$.

- Propose the latent data Y^* for the interval (t_j, t_{j+m}) for each $i = j+1, ..., M-1$ using a Brownian bridge proposal:

$$\begin{aligned}
q\left(Y_{t_{i+1}} \mid Y_{t_i}, \overline{Y}_{t_M}; \theta\right) &= \phi\left(Y_{t_{i+1}}, Y_{t_i} + \widetilde{\mu}_i, \widetilde{\sigma}_i\right) \\
\text{where } \widetilde{\mu}_i &= \frac{1}{M-i}\left(\overline{Y}_{t_M} - Y_{t_i}\right) \\
\widetilde{\sigma}_i &= \Delta t \frac{1}{M-i}(M-i-1)\,\sigma\left(Y_{t_i}\right)
\end{aligned}$$

where ϕ denotes the Gaussian density and $\sigma\left(Y_{t_i}\right)$ is the volatility term of the process for Y.

- Accept the parameter and latent data proposal with probability $\alpha = \min(1, A)$ and set $(Y_s, \theta_s) = (Y^*, \theta^*)$, or else set $(Y_s, \theta_s) = (Y_{s-1}, \theta_{s-1})$. A is given by:

$$A = \frac{\prod_{i=j}^{M-1} \pi\left(Y^*_{t_{i+1}} \mid Y^*_{t_i}; \theta^*\right) \prod_{i=j}^{M-2} q\left(Y_{t_{i+1}} \mid Y_{t_i}, Y_{t_i}, \overline{Y}_{t_M}; \theta\right)}{\prod_{i=j}^{M-1} \pi\left(Y_{t_{i+1}} \mid Y_{t_i}; \theta\right) \prod_{i=j}^{M-2} q\left(Y^*_{t_{i+1}} \mid Y^*_{t_i}, Y^*_{t_i}, \overline{Y}_{t_M}; \theta^*\right)}$$

where $\pi\left(Y_{t_{i+1}} \mid Y_{t_i}; \theta\right)$ is the Euler transition density.

2. Set $j = j + m$ and go to (1).

The resulting draws of latent data and parameters form a Markov chain, whose stationary distribution after an initial burn-in period is the joint posterior of the data and the model parameters:

$$\pi\left(Y,\theta\right) \propto \pi\left(\theta\right) \prod_{t=t_0}^{t_n-1} \left\{ \prod_{j=1}^{m} \pi\left(Y_{t,j+1} \mid Y_{t,j};\theta\right) \right\}$$

The number of imputed data points that are needed could be determined by running the sampler for $m = 1$ and consequently increasing the discretization points until there is no significant change in the posterior parameter samples.

Appendix B

For Chapter 3

B.1 Copula Functions

The following d-dimensional copula functions are used in the chapter.

- **Gaussian copula**

$$
\begin{aligned}
& C^{Ga}\left(u_1, u_2, ..., u_d \mid R_{Ga}\right) \\
= & \int\limits_{-\infty}^{\Phi^{-1}(u_1)} ... \int\limits_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{2\pi \left|R_{Ga}\right|^{1/2}} \exp\left\{-\frac{1}{2} x^{\intercal} R_{Ga}^{-1/2} x\right\} dx_1 ... dx_d
\end{aligned}
$$

where R_{Ga} denotes the correlation matrix, and $\Phi^{-1}(u_i)$ is the inverse of the univariate standard normal CDF.

- **Student's t copula**

$$
\begin{aligned}
& C^{t}\left(u_1, u_2, ..., u_d \mid R_t, \nu\right) \\
= & \int\limits_{-\infty}^{t_\nu^{-1}(u_1)} ... \int\limits_{-\infty}^{t_\nu^{-1}(u_d)} \frac{\Gamma\left(\frac{\nu+d}{2}\right) \left|R_t\right|^{1/2}}{\Gamma\left(\frac{\nu}{2}\right) \left(\nu\pi\right)^{d/2}} \left(1 + \frac{1}{\nu} x^{\intercal} R_t^{-1} x\right)^{-\frac{\nu+d}{2}} dx_1 ... dx_d
\end{aligned}
\tag{B.1.1}
$$

where R_t denotes the correlation matrix, ν is the degrees of freedom parameter, and $t^{-1}(u_i)$ is the inverse CDF of the univariate Student's t distribution with ν degrees of freedom.

- **Symmetrized Joe-Clayton copula**

This copula function was introduced by Patton (2004) and is based on the bivariate Joe-Clayton copula, that is a two-parameter copula function with parameters $\tau_L \in (0, 1)$ and $\tau_U \in (0, 1)$ that are a measure of the lower and upper tail dependence. The Joe-Clayton copula has the following form:

$$\begin{aligned}
& C^{JC}\left(u_1, u_2 \mid \tau_L, \tau_U\right) \\
= \quad & 1-\left\{1-\left[\left(1-\left(1-u_1\right)^\kappa\right)^{-\gamma}+\left(1-\left(1-u_2\right)^\kappa\right)^{-\gamma}-1\right]^{-\frac{1}{\gamma}}\right\}^{\frac{1}{\kappa}} \\
\text{where } \kappa \quad = \quad & \frac{1}{\log_2\left(2-\tau_U\right)} \\
\gamma \quad = \quad & -\frac{1}{\log_2\left(2-\tau_L\right)}
\end{aligned}$$

The symmetrized version of the copula, designed to render it completely symmetric for equal values of the lower and upper tail dependence parameters has the following form:

$$\begin{aligned}
& C^{SJC}\left(u_1, u_2 \mid \tau_L, \tau_U\right) \\
= \quad & \frac{1}{2}\left[C^{JC}\left(u_1, u_2 \mid \tau_L, \tau_U\right)+C^{JC}\left(1-u_1, 1-u_2 \mid \tau_U, \tau_L\right)+u_1+u_2-1\right]
\end{aligned}$$

B.2 Malliavin Derivatives of the State Variables

Recall that the Malliavin derivatives of the state variables $Y \equiv \left(X_1, X_2, F^V, F^M, Y^r\right)$ can be represented as the solutions to a linear stochastic differential equation[1]:

$$D_t Y_s = \sigma^Y\left(t, Y_t\right) \exp\left\{\int_t^s dL_v\right\}$$

where $\sigma^Y\left(t, Y_t\right)$ is the 5×5 matrix of diffusion terms of the state variables, and dL_t is defined by:

$$dL_t \equiv \left(\partial_2 \mu^Y\left(t, Y_t\right)-\frac{1}{2} \sum_{j=1}^5 \partial_2 \sigma_{\cdot j}^Y\left(t, Y_t\right) \partial \sigma_{\cdot j}^Y\left(t, Y_t\right)^{\intercal}\right) dt+\sum_{j=1}^5 \partial_2 \sigma_{\cdot j}^Y\left(t, Y_t\right) dW_{jt}$$

where $\partial_2 \mu^Y\left(t, Y_t\right)$ and $\partial_2 \sigma_j^Y\left(t, Y_t\right)$ denote the derivatives of $\mu^Y\left(t, Y_t\right)$ and $\sigma_{\cdot j}^Y\left(t, Y_t\right)$ with respect to Y_t, and $\sigma_{\cdot j}^Y\left(t, Y_t\right)$ denotes the j^{th} column of the matrix $\sigma^Y\left(t, Y_t\right)$. The particular forms of the drift $\mu^Y\left(t, Y_t\right)$ and the diffusion term $\sigma^Y\left(t, Y_t\right)$ of the state variables are given by:

$$\mu^Y(t, Y)=\left(\begin{array}{c}
\mu_1^X\left(t, X_t, F^V, F^M\right) \\
\mu_2^X\left(t, X_t, F^V, F^M\right) \\
\mu^{F^V}\left(t, F^V\right) \\
\mu^{F^M}\left(t, F^M\right) \\
\mu^{Y^r}\left(t, Y^r\right)
\end{array}\right)$$

where $\mu_i^X\left(t, X_t, F^V, F^M\right), i=1,2$ are given by (3.3.13), $\mu^{F^V}\left(t, F^V\right)=\kappa^V\left(\theta^V-F^V\right)$, $\mu^{F^M}\left(t, F^M\right)=\kappa^M\left(\theta^M-F^M\right)$, $\mu^{Y^r}\left(t, Y^r\right)=\kappa_r\left(\theta^r-Y_t^r\right)$.

[1] See Theorem 1 in Detemple et al. (2003)

$$\sigma^Y(t,Y) = \begin{pmatrix} \sigma^X_{11}(t,X) & \sigma^X_{12}(t,X) & 0 \\ \sigma^X_{21}(t,X) & \sigma^X_{22}(t,X) & 0 \\ \sigma^{F^V}\left(t,F^V\right) & \sigma^{F^V}\left(t,F^V\right) & 0 \\ \sigma^{F^M}\left(t,F^M\right) & \sigma^{F^M}\left(t,F^M\right) & 0 \\ 0 & 0 & \sigma^{Y^r}(t,Y^r) \end{pmatrix}$$

where $\sigma^X(t,X)$ is given by (**??**), $\sigma^{F^V}\left(t,F^V\right) = \sigma^V\sqrt{F^V}$, $\sigma^{F^M}\left(t,F^M\right) = \sigma^M$, and $\sigma^{Y^r}(t,Y^r) = \sigma_r\sqrt{Y^r_t}$.

Given the chosen specifications for the state variables, we can solve separately for the Malliavin derivatives of state variable driving the short rate, as well as for the Malliavin derivatives of the two factors. The processes that we have assumed for the observable factors (F^V for the VIX and F^M for CFNAI), as well as for the state variable Y^r, allow for either closed form solutions for the Malliavin derivatives (in the case of a Vasicek process) or for significant variance reduction in their simulation following the Doss transformation[2] that eliminates the stochastic term in the Malliavin derivative (for a CIR process).

In the Vasicek case, the Malliavin derivative of F^M simplifies significantly to:

$$D_{i,t}F^M_s = \sigma^M \exp\left\{-\kappa^M(s-t)\right\}, i = 1,2$$

For the other two state variables, Y^r and F^V, we have assumed a CIR process, that can be reduced to have constant diffusion term through a suitable change of variable technique, which then eliminates the stochastic terms for the simulation of the corresponding Malliavin derivatives. For a univariate diffusion, this variance stabilizing transformation is described in detail in Proposition 2 of Detemple et al. (2003) and we reproduce it here for completeness.

Consider a state variable Y satisfying a stochastic differential equation

$$dY_t = \mu(t,Y_t)\,dt + \sigma(t,Y_t)\,dW_t$$

We can replace it with a new state variable $Z_t = F(t,Y_t)$ where the function $F : [0,T]\times\mathbb{R} \to \mathbb{R}$ is such that $\partial_2 F = \frac{1}{\sigma^Y}$. Then for a continuously differentiable drift μ, twice continuously differentiable diffusion term σ, that also satisfy the growth conditions that $\mu(t,0)$ and $\sigma(t,0)$ are bounded for all $t \in [0,T]$, then we have for $t \leq s$:

$$\begin{aligned} D_tY_s &= \sigma(t,Y_t)\,D_tZ_s \\ \text{where } D_tZ_s &= \exp\left\{\int_t^s \partial_2 m(v,Z_v)\,dv\right\} \\ m(t,Z) &\equiv \left[\frac{\mu}{\sigma} - \frac{1}{2}\partial_2\sigma + \partial_1 F\right](t,Y) \end{aligned}$$

[2] See Detemple et al. (2003)

Bibliography

Aït-Sahalia, Y.: 1999, Transition densities for interest rate and other nonlinear diffusions, *Journal of Finance* **54**, 1361–1395.

Aït-Sahalia, Y.: 2003, Closed-form likelihood expansions for multivariate diffusions, *Working paper* .

Ang, A. and Bekaert, G.: 2002, International asset allocation with regime shifts, *Review of Financial Studies* **15**, 1137–1187.

Ang, A. and Chen, J.: 2002, Asymmetric correlations of equity portfolios, *Journal of Financial Economics* **63**, 443–494.

Backus, D., Foresi, S. and Telmer, C.: 1998, Affine models of currency pricing: Accounting for the forward premium anomaly. Working Paper.

Barndorff-Nielsen, O. E.: 1977, Exponentially decreasing distributions for the logarithm of the particle size, *Proceedings of the Royal Society London A 353* pp. 401–419.

Barndorff-Nielsen, O. E.: 1978, Hyperbolic distributions and distributions on hiperbolae, *Scandinavian Journal of Statistics* **5**, 151–157.

Barndorff-Nielsen, O. E.: 1995, Normal inverse gaussian processes and the modeling of stock returns, *Rep. No. 300* .

Barndorff-Nielsen, O. E. and Blaesid, P.: 1981, Hyperbolic distributions and ramifications: Contributions to theory and application, *in* C. Taillie, G. Patil and B. Baldessari (eds), *Statistical Distributions in Scientific Work*, Vol. 4, pp. 19–44.

Bauer, C.: 2000, Value at risk using hyperbolic distributions, *Journal of Economics and Business* **52**, 435–453.

Bibby, B. M. and Sorensen, M.: 1995, Martingale estimating functions for discretely observed diffusion processes, *Bernoulli* **1**, 17–39.

Bibby, B. M. and Sorensen, M.: 1997, A hyperbolic diffusion model for stock prices, *Finance Stoch.* **1**, 25–41.

Bibby, B. M. and Sorensen, M.: 2001, Simplified estimating functions for diffusion models with a high-dimensional parameter, *Scandinavian Journal of Statistics* **28**(1), 99–112.

Bibby, B. M. and Sorensen, M.: 2003, Hyperbolic processes in finance, *in* S. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance*, Elsevier Science, pp. 211–248.

Bollerslev, T., Engle, R. and Woolridge, J.: 1988, A capital asset pricing model with time-varying covariances, *Journal of Political Economy* **96**, 116–131.

Boyle, P. and Imai, J.: 2002, Asset allocation using quasi monte carlo methods. Working paper.

Bru, M. F.: 1991, Wishart processes, *Journal of Theoretical and Applied Probability* **4**, 725–743.

Buraschi, A., Porchia, P. and Trojani, F.: 2007, Correlation risk and optimal portfolio choice. University of St. Gallen Working paper series in Finance, Paper No. 60.

Carr, P. and Wu, L.: 2004, Time-changed levy processes and option pricing, *Journal of Financial Economics* **71**(1), 113–141.

Chan, K., Karolyi, A., Longstaff, F. and Sanders, A.: 1992, An empirical comparison of alternative models of the short-term interest rate, *The Journal of Finance* **47**, 1209–1227.

Chen, X., Hansen, L. P. and Scheinkman, J. A.: 2002, Principal components and the long run, *Working paper* .

Cherubini, U., Luciano, E. and Vecchiato, W.: 2004, *Copula Methods in Finance*, John Wiley and Sons, England.

Chunhachinda, P., Dandapani, K., Hamid, S. and Prakash, A.: 1997, Portfolio selection and skewness: Evidence from international stock markets, *Journal of Banking and Finance* **21**, 143–167.

Cox, J. and Huang, C. F.: 1989, Optimal consumption and portfolio policies when asset prices follow a diffusion process, *Journal of Economic Theory* **49**, 33–83.

Cvitanic, J., Polimenis, V. and Zapatero, F.: 2008, Optimal portfolio allocation with higher moments, *Annals of Finance* **4**, 1–28.

Dai, Q. and Singleton, K.: 2000, Specification analysis of affine term structure models, *The Journal of Finance* **55**, 1943–1978.

Das, S. and Uppal, R.: 2004, Sistemic risk and international portfolio choice, *Journal of Finance* **59**, 2809–2834.

Detemple, J., Garcia, R. and Rindisbacher, M.: 2003, A monte carlo method for optimal portfolios, *Journal of Finance* **58**, 410–446.

Drouet-Mari, D. and Kotz, S.: 2001, *Correlation and Dependence*, Imperial College Press, London.

Duffee, G.: 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* **57**(1), 405–443.

Duffie, D., Pan, J. and Singleton, K.: 2000, Transform analysis and asset pricing for affine jump diffusions, *Econometrica* **68**(6), 1343–1376.

Durham, G. and Gallant, A.: 2002, Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes, *Journal of Business and Economic Statistics* **20**(3), 297–316.

Eberlein, E. and Keller, U.: 1995, Hyperbolic distributions in finance, *Bernoulli* **1**, 281–299.

Elerian, O., Chib, S. and Shephard, N.: 2001, Likelihood inference for discretely observed non-linear diffusions, *Econometrica* **69**, 959–993.

Embrechts, P., McNeil, A. and Strautman, D.: 2002, Correlation and dependence in risk management: Properties and pitfalls, *in* M. Dempster (ed.), *Risk Management: Value at Risk and Beyond*, Cambridge University Press, Cambridge, pp. 176–223.

Engle, R.: 2002, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics* **20**, 339–350.

Eraker, B.: 2001, MCMC analysis of diffusion models with applications to finance, *Journal of Business and Economic Statistics* **19**(2), 177–191.

Eraker, B., Johannes, M. and Polson, N.: 2003, The impact of jumps in returns and volatility, *Journal of Finance* **53**, 1269–1300.

Erb, C., Harvey, C. and Viskanta, T.: 1994, Forecasting international equity correlations, *Financial Analyst Journal* **November/December**, 32–45.

Fermanian, J.-D. and Wegkamp, M.: 2004, Time-dependent copulas. Working paper.

Florens-Zmirou, D.: 1993, On estimating the diffusion coefficient from discrete observations, *Journal of Applied Probability* **30**, 790–804.

Frees, E. W. and Valdez, E.: 1998, Understanding relationships using copulas, *North American Actuarial Journal* **2**(1), 1–25.

Genest, C. and MacKay, J.: 1986, The joy of copulas: Bivariate distributions with uniform marginals, *Amer. Statist.* **40**, 280–283.

Goetzmann, W. N., Li, L. and Rouwenhorst, K. G.: 2005, Long-term global market correlations, *Journal of Business* **78**, 1–38.

Golightly, D. and Wilkinson, J.: 2006a, Bayesian sequential inference for nonlinear multivariate diffusions, *Statistics and Computing* **16**, 323–338.

Golightly, D. and Wilkinson, J.: 2006b, Bayesian sequential inference for stochastic kinetic biomechanical network models, *Journal of Computational Biology* **13**(3), 838–851.

Goorbergh, R. V. D., Genest, C. and Werker, B.: 2003, Multivariate option pricing using dynamic copula models. CentER Discussion paper No. 2003-122.

Gumbel, E. J.: 1960, Bivariate exponential distributions, *Journal of the American Statistical Association* **55**, 698–707.

Hansen, L. P. and Scheinkman, J. A.: 1995, Back to the future: Generating moment implications for continuous-time markov processes, *Econometrica* **63**(4), 767–804.

Heston, S.: 1993, A closed-form solution for options with stochastic volatility with application to bond and currency options, *The Review of Financial Studies* **6**(2), 327–343.

Hong, Y., Jun, T. and Zhou, G.: 2003, Asymmetric correlations of equity portfolios, *Journal of Financial Economics* **63**, 443–494.

Jaschke, S. R.: 1997, A note on stochastic volatility, GARCH models, and hyperbolic distributions. Working paper, SFB 373, Humboldt-Universitat Berlin.

Joe, H.: 1997, *Multivariate Models and Dependence Concepts*, Chapman and Hall, London.

Johannes, M., Polson, N. and Stroud, J.: 2004, Sequential parameter estimation in stochastic volatility models with jumps. Working paper.

Jondeau, E. and Rockinger, M.: 2002, Conditional dependency of financial series: The copula-GARCH model. FAME research paper No 69.

Jondeau, E. and Rockinger, M.: 2005, Conditional asset allocation under non-normality: How costly is the mean-variance criterion. FAME Research Paper No. 132.

Jones, C.: 2003, Nonlinear mean reversion in the short-term interest rate, *The Review of Financial Studies* **16**(3), 793–843.

Joy, C., Boyle, P. and Tan, K.: 1996, Quasi-monte carlo methods in numerical finance, *Management Science* **42**(6), 926–938.

Kim, S., Shephard, N. and Chib, S.: 1998, Stochastic volatility: Likelihood inference and comparison with ARCH models, *Review of Economic Studies* **65**, 361–393.

Kunz, A.: 2002, Extremes of multivariate diffusions in finance: A data analysis. Working paper, Center of Mathematical Sciences, Technische Universität München.

Ledoit, O., Santa-Clara, P. and Wolf, M.: 2003, Flexible multivariate GARCH modeling with an application to international stock markets, *Review of Economics and Statistics* **85**, 735–747.

Lewis, M. and Raftery, A.: 1997, Estimating bayes factors via posterior simulation with the laplace-metropolis estimator, *Journal of the American Statistical Association* **92**(438), 648–655.

Li, L.: 2002, Macroeconomic factors and the correlation of stock and bond returns. Yale ICF Working Paper No. 02-46.

Liu, J.: 2007, Portfolio selection in stochastic environments, *Review of Financial Studies* **20**, 1–39.

Liu, J., Longstaff, F. and Pan, J.: 2003, Dynamic asset allocation with event risk, *Journal of Finance* **58**, 231–259.

Liu, J. and West, M.: 2001, Combined parameter and state estimation in simulation-based filtering, *in* A. Doucet, N. de Freitas and N. Gordon (eds), *Sequential Monte Carlo Methods in Practice*, Springer, Berlin.

Longin, F. and Solnik, B.: 1995, Is the correlation in international equity returns constant: 1960-1990?, *Journal of International Money and Finance* **14**, 3–26.

Longin, F. and Solnik, B.: 2001, Extreme correlation of international equity markets, *Journal of Finance* **56**, 649–675.

Moskowitz, T.: 2003, An analysis of covariance risk and pricing anomalies, *Review of Financial Studies* **16**, 417–457.

Nelsen, R. B.: 1999, *An Introduction to Copulas*, Springer, New York.

Nelson, D. B.: 1990, ARCH models as diffusion approximations, *Journal of Econometrics* **45**, 7–38.

Ocone, D. and Karatzas, I.: 1991, A generalized clark representation formula with applications to optimal portfolios, *Stochastics and Stochastic Reports* **34**, 187–220.

Patton, A.: 2004, Modeling asymmetric exchange rate dependence, *Working Paper, LSE* .

Pedersen, A. R.: 1995, Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes, *Bernoulli* **1**, 257–279.

Poon, S. H., Rockinger, M. and Tawn, J.: 2004, Extreme value dependence in financial markets: Diagnostics, models, and financial implications, *Review of Financial Studies* **17**(2), 581–610.

Prakash, A., Chang, C. and Pactwa, T.: 2003, Selecting a portfolio with skewness: Recent evidence from US, european, and latin american equity markets, *Journal of Banking and Finance* **27**, 1375–1390.

Prause, K.: 1999, *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*, PhD thesis, University of Freiburg.

Ribiero, R. and Veronesi, P.: 2002, The excess comovement of international stock markets in bad times: A rational expectations equilibrium model. Working paper, University of Chicago.

Roberts, G. O.: 1996, Markov chain concepts related to sampling algorithms, *in* W. Gilks, S. Richardson and D. J. Spiegelhalter (eds), *Markov Chain Monte Carlo in Practice*, markov chain monte carlo in practice edn, Chapman and Hall, London, pp. 45–57.

Roberts, G. O. and Strammer, O.: 2001, On inference for partially observed non-linear diffusion models using the metropolis-hastings algorithm, *Biometrika* **88**(3), 603–621.

Rydberg, T.: 1999, Generalized hyperbolic diffusion processes with applications in finance, *Mathematical Finance* **9**(2), 183–201.

Stock, J. and Watson, M.: 1999, Forecasting inflation. NBER Working Paper 7023, available at http://www.nber.org/papers/w7023.

Stroud, J., Polson, N. and Müller, P.: 2004, Practical filtering for stochastic volatility models, *in* A. Harvey, S. J. Koopman and N. Shephard (eds), *State Space and Unobserved Components Models*, Cambridge Press, pp. 236–247.

Tierny, L.: 1994, Markov chains for exploring posterior distributions. Technical report No 560 School of Statistics, University of Minnesota.

Tse, Y. K., Zhang, X. and Yu, J.: 2004, Estimation of hyperbolic diffusion using MCMC method, *Quantitative Finance* **4**, 158–169.

Whelan, N.: 2004, Sampling from archimedean copulas, *Quantitative Finance* **4**, 339–352.