Nonlinear Term Structure Dependence: Copula Functions, Empirics, and Risk Implications*

Markus Junker
caesar Financial Engineering
D-53175 Bonn, Germany
junker@caesar.de

Alexander Szimayer
The University of Western Australia
Crawley, WA 6009, Australia
aszimaye@ecel.uwa.edu.au

Niklas Wagner
Munich University of Technology
D-80290 Munich, Germany
niklas.wagner@wi.tum.de

Version: July 2003

*The authors would like to thank participants at the 2002 Copula Workshop, the Frankfurt Math-Finance Colloquium and at the 2003 CEPR European Summer Symposium for helpful comments and discussions. Special thanks are to Robert Durand and Claudia Klüppelberg. A first draft of the paper was written and the dataset was obtained while the third author was visiting Stanford GSB. The authors are grateful to David Tien for kindly providing additional data. All errors soley remain with the authors. E-mail correspondence: niklas.wagner@wi.tum.de

Nonlinear Term Structure Dependence: Copula Functions, Empirics, and Risk Implications

Abstract

This paper documents nonlinear cross-sectional dependence in the term structure of U.S. Treasury yields and points out risk management implications. The analysis is based on a Kalman filter estimation of a two-factor affine model which specifies the yield curve dynamics. We then apply a broad class of copula functions for modeling dependence in factors spanning the yield curve. Our sample of monthly yields in the 1982 to 2001 period provides evidence of upper tail dependence in yield innovations; i.e., large positive interest rate shocks tend to occur under increased dependence. In contrast, the best fitting copula model coincides with zero lower tail dependence. This asymmetry has substantial risk management implications. We give an example in estimating bond portfolio loss quantiles and report the biases which result from an application of the normal dependence model.

Key words: affine term structure models, nonlinear dependence, copula functions, tail dependence, value-at-risk

JEL classification: C13, C16, G10, G21

1 Introduction

The class of affine term structure models (ATSMs) as proposed by Duffie and Kan (1996) and further characterized by Dai and Singleton (2000), has recently become a benchmark in modeling the term structure of default-free interest rates. Within the model class, the term structure is characterized by the current realizations as well as the dynamics of a set of state variables. Logarithmic bond prices are then affine functions of these state variables. The class appeals by its analytical tractability and contains the well-known models by Vasicek (1977), Cox et al. (1985), Chen and Scott (1992), and Longstaff and Schwartz (1992), for example. However, recent empirical evidence indicates that term structure data do not fully confirm the ATSM class. A series of articles document distinct nonlinearities e.g. in the drift and volatility function of the short-rate, particularly implying that mean-reversion in the short-rate depends on its level; see for example Aït-Sahalia (1996) and Stanton (1997). Ang and Bekaert (2000) focus on these findings and develop a Markovian switching-model which captures such nonlinearities. Also, empirical results on one- and two-factor ATSMs by Duan and Simonato (1999) indicate a rejection of the affine model assumption when tested against local alternatives.[1] In general, findings of nonlinearity in the term structure of interest rates are important for at least three reasons. First, only an exact assessment of state variable dynamics and their dependence allows for an accurate modeling of the term structure. Second, derivatives pricing is frequently based on assumptions imposed by the class of ATSMs. And lastly, effective bond portfolio risk management builds upon models which give reliable risk implications.

While previous empirical studies have focused on time-series nonlinearities and discontinuities in the process dynamics, this paper analyzes nonlinear cross-sectional dependence between factors that span the yield curve. We show that the dependence structure of the long and short end of the yield curve exhibits nonlinearity which can be characterized under a particular focus on extremal dependence. The starting point of our model is the benchmark-class of ATSMs. Based on this theory, the term structure dynamics in our study are given by a Gaussian two-factor generalized Vasicek model. This model was applied for example by Babbs and Nowman (1998) who find that the two-factor approach provides a good description of the yield curves for a broad sample of mature bond markets. Formulating a discrete time model in state-space representation allows for pa-

[1]Other specifications of the term structure include Ahn et al. (2002) as well as alternative formulations of the short rate e.g. by Chan et al. (1992) and Aït-Sahalia (1996). For an extense survey of models see also Dai and Singleton (2002). Besides the linear structure, the distributional assumptions imposed by ATSMs is critical: Björk et al. (1997) extend the diffusion driven ATSMs by allowing for jumps. Eberlein and Raible (1999) study term structure models driven by general Lévy processes.

rameter estimation. We then focus on cross-sectional dependence in the term structure by modeling general forms of dependence in discrete factor innovations by a broad choice of copula functions. While a Gaussian factor model allows for correlated factors only, copula functions as outlined in Joe (1997) and Nelsen (1999), generalize the dependence concept by separating the treatment of dependence and marginal behavior. Based on the model, it is possible to characterize dependence in the center of the distribution independently from dependence in the distribution tails. Hence, we can impose various combinations of symmetric as well as asymmetric tail dependence on the factor innovations. Recent studies which apply copula functions in finance such as for example Ané and Kharoubi (2001) and Scaillet (2002) indicate that the concept appeals in modeling complex dependence structures. Malevergne and Sornette (2002) argue that the hypothesis of the normal copula cannot be rejected for a variety of financial returns, including stock and exchange rate returns. However, this finding may relate to the amount of data available and to issues of power of the testing procedures in the presence of tail dependence. Indeed, the authors also find that alternative copula models cannot be rejected either. Conditional copula functions are studied by Patton (2001) who models conditional dependence in U.S.-Dollar exchange rate returns and by Rockinger and Jondeau (2001) who examine conditional dependence in international stock market returns. Within this literature, there is still debate on which copula models are most appropriate. To our knowledge, no evidence for the dependence structure within the term structure of interest rates has yet been provided.

In our empirical investigation we focus on the term structure of U.S.-Treasuries which represent the largest government bond market worldwide. We use a sample of monthly yield curve observations as in the empirical studies for example by Ang and Bekaert (2000) and De Jong (2000). The sample covers the 20-year period from October 1982 to December 2001. We form two 10-year subsamples in order to check for the robustness of the empirical results. The empirical investigation in the paper is then organized in two steps. In the first step, we use the class of affine term structure models to specify the yield curve dynamics. In particular, we choose a two-factor generalized Vasicek model characterized by a jointly normal bivariate factor process. We then extract factors representing yields, namely the interest rates on zero-coupon bonds with one year and five years to maturity. The model parameters are estimated by Kalman filtering as supported by maximum likelihood arguments; this was also done in previous studies such as for example Lund (1997), Duan and Simonato (1999), and Dewachter et. al. (2002). In the second step, we model the dependence structure within the yield curve. We thereby focus on the dependence relation between short-term and long-term interest rates as represented by the two yield factors. To this aim, a broad set of different copula functions is used.

Based on our empirical findings, we show that the class of elliptical copula functions – including symmetric copulas such as the normal and the Student-t– has characteristics which violate the observed complex dependence structure. Hence, the yield factor dependence cannot be characterized by a correlation coefficient as in the normal model nor with a symmetric Student-t model. While the copula function of the normal distribution does not allow for dependence in the tails, the Student-t copula does not allow for asymmetric tail dependence. However, dependence models contained in the class of Archimedean copulas can indeed capture dependence in the yield curve which is characterized by distinct asymmetry and upper tail dependence. The Gumbel as well as a transformed Frank copula turn out to be more suitable choices than the student-t copula. Considering all candidate models used in our study, we find the transformed Frank copula to be the most appropriate model. Moreover, the goodness-of-fit tests for the two subsamples indicate that our main conclusions are robust within the observation period.

Given our empirical findings, we demonstrate the risk management implications in a bond portfolio setting. Based on the affine model of factor dynamics and the alternative copula models of factor dependence, we study the pricing effects of nonlinear dependence in the yield factors. Particularly, we use the ATSM implication that bond prices are exponential affine functions of the state variables. By sampling from one year and five year yield factors under the fitted copula functions we then estimate loss quantiles for bond portfolios with alternative durations. Our analysis highlights that the normal copula function –which is implied by the assumption of linear dependence in affine term structures– yields a substantial bias in the assessment of portfolio risk. When compared to the transformed Frank copula which captures the asymmetric dependence in the data, we report a bias structure in the upper and lower bond portfolio loss quantiles which yields values as high as 6 percent as compared to the normal model.

The remainder of this paper is organized as follows. In the next section, we outline the model used in the analysis. Term structure time series dynamics are given with the class of ATSMs. Cross-sectional dependence in bivariate term structure innovations is modeled by two candidate classes of copula functions. The empirical investigation and the estimation results are given in Section 3. The application to bond portfolio risk management which points out risk implications of nonlinear factor dependence is given in Section 4. Section 5 concludes.

2 The Term Structure Model

The starting point of our model is the class of benchmark ATSMs. We model the term structure time series dynamics within a continuous time two-factor generalized Vasicek model. A state-space representation allows for observational noise and prepares estimation based on discrete time vector autoregressive version of the model. We then focus on cross-sectional dependence in the term structure by modeling dependence in factor innovations by copula functions. The functions stem from two broad classes of copula functions. Based on the copula model, we can impose various combinations of symmetric as well as asymmetric tail dependence on the factor innovations.

2.1 Term Structure Dynamics

2.1.1 Affine Term Structure Models

The affine term structure model is a class of models in which the yields to maturity are affine functions of some state variable vector. The state vector X is assumed to obey the following dynamics

$$dX(t) = \kappa(\theta - X(t))\,dt + \Sigma\sqrt{S(t)}\,dW(t), \tag{1}$$

where W is a d-dimensional standard Brownian motion, θ is a d-vector, κ and Σ are $d \times d$ matrices, and $S(t)$ is a $d \times d$ diagonal matrix with diagonal elements which are affine functions of the state vector X. Provided that a parameterization is admissible, the price of a zero bond $P(t, \tau)$ in time t with time to maturity τ can be expressed as

$$P(t, \tau) = \exp\left(A(\tau) + B(\tau)^\top X(t)\right), \tag{2}$$

where A is a scalar function, and B is a d-dimensional vector function. The instantaneous interest rate is, as usual, defined as

$$r(t) = -\lim_{\tau \searrow 0} \frac{\ln P(t, \tau)}{\tau}. \tag{3}$$

Duffie and Kan (1996) show that $P(\cdot, \cdot)$ is generically exponential affine, i.e. in the form of equation (2), if and only if the mean and variance in equation (1), and the short rate r are affine functions in the state variable X. Moreover, A and B in equation (2) are obtained as solutions to ordinary differential equations. Let $R(t, \tau)$ denote the time-t continuously compounded yield on a zero bond with maturity τ. The yield to maturity of this bond is

$$R(t, \tau) = -\frac{\ln P(t, \tau)}{\tau}. \tag{4}$$

2.1.2 The Gaussian Two-Factor Model

The special case of the two-factor generalized Vasicek model is given by

$$\begin{aligned} r(t) &= R_0 + X_1(t) + X_2(t), \\ dX(t) &= -\kappa\, X(t)\, dt + \Sigma\, dW(t), \end{aligned} \tag{5}$$

where W is a 2-dimensional standard Brownian motion, and

$$\kappa = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}, \text{ and } \Sigma = \begin{pmatrix} \sigma_1 & 0 \\ \rho\,\sigma_2 & \sqrt{1-\rho^2}\,\sigma_2 \end{pmatrix}.$$

The parameter R_0 is the mean level of the instantaneous rate r, the state processes X_1 and X_2 fluctuate around zero with mean reversion rates κ_1, κ_2, and diffusion coefficients σ_1, σ_2, and correlation ρ. Details on the functions A and B describing the term structure implied by the two-factor model are given in Duffie and Kan (1996) and Babbs and Nowman (1998); see also Appendix A.1. These functions are given by the factor parameters defined above and by γ_1 and $\gamma_2 \in \mathbb{R}$, which represent the risk premia of factor one and factor two, respectively.

2.1.3 State-Space Representation

Estimation of the above term structure model can be carried out via transformation to state-space representation; see for example Babbs and Nowman (1998, 1999) and Duan and Simonato (1999) for term structure estimation applications and Harvey (1989) for a general treatment of state-space models.

Assume that the yields for different maturities are observed with error. After the addition of measurement error, the yield to maturity, using the bond pricing formula (2), can be written as

$$R(t,\tau) = \frac{-A(\tau)}{\tau} + \frac{-B(\tau)^\top X(t)}{\tau} + \varepsilon(t,\tau), \tag{6}$$

where $\varepsilon(t,\tau)$ is assumed to be a normally distributed error term with mean zero and standard deviation $\sigma_{\varepsilon_\tau}$. Hence, given that N bond yields for different maturities are observed, the N corresponding yields have the following representation:

$$\begin{bmatrix} R(t,\tau_1) \\ \vdots \\ R(t,\tau_N) \end{bmatrix} = \begin{bmatrix} \frac{-A(\tau)}{\tau_1} \\ \vdots \\ \frac{-A(\tau)}{\tau_N} \end{bmatrix} + \begin{bmatrix} \frac{-B(\tau)^\top}{\tau_1} \\ \vdots \\ \frac{-B(\tau)^\top}{\tau_N} \end{bmatrix} X(t) + \begin{bmatrix} \varepsilon(t,\tau_1) \\ \vdots \\ \varepsilon(t,\tau_N) \end{bmatrix}. \tag{7}$$

In terms of the state-space model, this equation is referred to as the measurement equation. To obtain the transition equation for the state-space model, the expressions for the conditional mean and variance for the unobserved state variable process over a discrete time interval of length h have to be derived. Define $m(X(t);h) = \mathbb{E}\{X(t+h)\,|X(t)\}$ and $\Phi(X(t);h) = \mathrm{Var}(X(t+h)\,|X(t))$, then the transition equation reads

$$X(t+h) = m(X(t);h) + \Phi(X(t);h)^{1/2}\,\eta(t,h), \tag{8}$$

where $\eta(t,h)$ is a d-vector of Gaussian white noise with $\Phi(X(t);h)^{1/2}$ denoting the Cholesky decomposition of $\Phi(X(t);h)$.

The two-factor model (5) defines the state variables as Gauss-Markov processes and thus the conditional mean and the conditional variance are:

$$m(x;h) = (m^h_{i,j})_{2\times 2}\cdot x = \begin{pmatrix} e^{-\kappa_1 h} & 0 \\ 0 & e^{-\kappa_2 h} \end{pmatrix} x, \tag{9}$$

$$\Phi(x;h) = (\Phi^h_{i,j})_{2\times 2} = \begin{pmatrix} \frac{\sigma_1^2}{2\kappa_1}(1-e^{-2\kappa_1 h}) & \frac{\rho\,\sigma_1\,\sigma_2}{\kappa_1+\kappa_2}(1-e^{-(\kappa_1+\kappa_2)\,h}) \\ \frac{\rho\,\sigma_1\,\sigma_2}{\kappa_1+\kappa_2}(1-e^{-(\kappa_1+\kappa_2)\,h}) & \frac{\sigma_2^2}{2\kappa_2}(1-e^{-2\kappa_2 h}) \end{pmatrix}. \tag{10}$$

Given observations of the yield vector in (6) and under a discrete sampling scheme with interval h, the exact likelihood function can be established based on the Kalman filter estimate of the unobservable state variable process X.

2.2 Nonlinear Term Structure Dependence

2.2.1 The Discrete-Time Factor Process

Section 2.1 above outlined the two-factor affine term structure model which we apply in our study for capturing the term structure dynamics. The generalized Vasicek model (5) is based on continuous time factor dynamics dX driven by two-dimensional Brownian motion.

The factor process given by transition equation (8) is linear in the drift and non-stochastic in the diffusion coefficient. Hence, given (9) and (10), a discrete-time sample of X under $h=1$, dropping h superscripts, is given by a vector autoregressive process of order one

$$\begin{aligned} X_{1,t} &= m_{1,1}\,X_{1,t-1} + (\Phi^{1/2})_{1,1}\eta_{1,t} + (\Phi^{1/2})_{1,2}\eta_{2,t}, \\ X_{2,t} &= m_{2,2}\,X_{2,t-1} + (\Phi^{1/2})_{2,1}\eta_{1,t} + (\Phi^{1/2})_{2,2}\eta_{2,t}, \end{aligned} \tag{11}$$

with $t = 0, 1, ..., T$. The variables $\eta_{i,t}$, $i = 1, 2$ are uncorrelated iid standard normal innovations. In this setting, factor dependence is completely characterized by the correlation coefficient ρ. Generalizing the above model, we now rewrite the discrete-time factor dynamics as

$$\begin{aligned} X_{1,t} &= m_{1,1}\, X_{1,t-1} + Z_{1,t}, \\ X_{2,t} &= m_{2,2}\, X_{2,t-1} + Z_{2,t}, \end{aligned} \tag{12}$$

and assume that the innovations $(Z_{1,t}, Z_{2,t})$ are iid vectors with common joint distribution function $H(z_1, z_2)$. This relaxes the assumption of joint normality as imposed by the class of ATSMs.

2.2.2 Copula Functions

Based on (12), copula functions allow us to treat general versions of factor dependence in the two-factor generalized Vasicek model (5). The copula concept dates back to seminal papers by Hoeffding and Sklar; recent methodological overviews are given for example by Joe (1997) and Nelsen (1999). For the present application, we restrict the exposition to the two-dimensional case.

Let F_{Z_1} and F_{Z_2} denote the continuous marginal distribution functions of Z_1 and Z_2, i.e. $H(z_1, \infty)$ and $H(\infty, z_2)$, respectively. By transformation we obtain uniform random variables as $U = F_{Z_1}(Z_1)$ and $V = F_{Z_2}(Z_2)$. The copula function $C : [0,1]^2 \to [0,1]$ for the bivariate random vector (Z_1, Z_2) is defined as the joint distribution function of the uniform random vector $(U, V) = (F_{Z_1}(Z_1), F_{Z_2}(Z_2))$, that is, $C(u, v) = \mathbb{P}[U \leq u, V \leq v]$. Hence, it follows

$$H(z_1, z_2) = C(F_{Z_1}(z_1), F_{Z_2}(z_2)), \tag{13}$$

which is known as Sklar's Theorem. The result generally implies that for multivariate distribution functions the univariate margins and the dependence structure can be separated. Given that the marginal distribution functions are continuous, dependence is represented by a unique copula function C.

Apart from a separate treatment of dependence and marginal behavior, copula functions may characterize dependence in the center of the distribution differently while showing identical limiting properties in characterizing dependence in the distribution tails, as well as vice versa. Given the stylized fact of fat-tails in financial return distributions, tail dependence is therefore an interesting characteristic of copula functions. One can distinguish lower and upper tail dependence as defined below.

Definition 2.1 *The copula function C is* lower tail dependent *if*

$$\lim_{u \to 0} \frac{\mathbb{P}[U \leq u, V \leq u]}{u} = \lim_{u \to 0} \frac{C(u,u)}{u} = \lambda_L, \quad \lambda_L \in (0,\,1],$$

and C is upper tail dependent *if*

$$\lim_{u \to 1} \frac{\mathbb{P}[U > u, V > u]}{1-u} = \lim_{u \to 1} \frac{1-u-u+C(u,u)}{1-u} = \lambda_U, \quad \lambda_U \in (0,\,1].$$

Since the tail dependence measures λ_L and λ_U are limit properties of a copula we can write $\lambda_L = \lambda_L(C)$ ($\lambda_U = \lambda_U(C)$) or $\lambda_L = \lambda_L(\theta)$ ($\lambda_U = \lambda_U(\theta)$) if C is member of a parametric family C_θ with parameter vector θ. For the sake of simplicity, we may write λ whenever $\lambda_L = \lambda_U$.[2] We next introduce and characterize two standard classes of copula functions.

a) Elliptical copulas

The class of the elliptical copulas is widely used as a benchmark model. Elliptical copulas are commonly defined as copulas of elliptical distributions. In particular, this includes the copula of the student-t and the normal distribution function:

(a) The t-copula C_t is given by

$$C_t(u,v) = T_{\nu,\rho}\left(T_\nu^{\leftarrow}(u), T_\nu^{\leftarrow}(v)\right), \tag{14}$$

where $T_{\nu,\rho}$ is the bivariate standardized student-t distribution function with ν degrees of freedom and correlation ρ, while T_ν denotes the univariate standardized student-t distribution function. The upper and lower tail dependence parameter λ for $\nu > 2$ is:[3]

$$\lambda = 2\left(1 - T_{\nu+1}\left(\sqrt{\nu+1}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)\right). \tag{15}$$

[2] Note that one can not be sure from a finite iid sample observation whether the underlying copula function is tail dependent or not. However, recent empirical studies e.g. by Ané and Kharoubi (2001) and Malevergne and Sornette (2002) exhibit that the concept of tail dependence is a useful tool to describe observed dependence structures in financial data.

[3] See Embrechts et al. (2002). Note that the above expression for λ even holds in the case $0 < \nu \leq 2$ then with a different interpretation of ρ.

(b) For $\nu \to \infty$ the t-copula degenerates to the copula of the normal distribution

$$C_N(u,v) = N_\rho\left(N^{\leftarrow}(u), N^{\leftarrow}(v)\right), \tag{16}$$

where $N_\rho(\cdot)$ and $N(\cdot)$ denote the standard bivariate and the standard univariate normal distribution functions, respectively. From equation (15), it is obvious that zero tail dependence, i.e. $\lambda = 0$, results.

b) Archimedean copulas

Elliptical copulas as outlined above are restricted to symmetry. For this reason, we outline the more general class of Archimedean copulas. They are described by a generator function φ as given in the proposition below.

Proposition 2.2 *Let $\varphi : [0,\, 1] \to [0,\, \infty]$ be continuous and strictly decreasing with $\varphi(1) = 0$. The function $C : [0,\, 1]^2 \to [0,\, 1]$ given by*

$$C(u,v) = \varphi^{[-1]}\left(\varphi\left(u\right) + \varphi\left(v\right)\right) \tag{17}$$

is a copula if and only if φ is convex.

Here $\varphi^{[-1]} : [0,\, \infty] \to [0,\, 1]$ denotes the pseudo-inverse of φ. The copula constructed by (17) is called Archimedean. The function φ is called generator of the copula. A generator φ is called strict if $\varphi(0) = \infty$ and in this case $\varphi^{[-1]} = \varphi^{-1}$.The following Archimedean copulas are utilized in this paper:

(a) The independence copula with generator $\varphi_\Pi(q) = -\ln q$ and

$$C_\Pi\left(u, \nu\right) = u\nu. \tag{18}$$

The copula exhibits neither lower nor upper tail dependence, i.e.: $\lambda_L = \lambda_U = 0$.

(b) The Gumbel copula with generator $\varphi_G(q) = \left(-\ln q\right)^\delta$ where $\delta \in [1,\, \infty)$ and

$$C_G\left(u, v\right) = \exp\left(-\left[\left(-\ln u\right)^\delta + \left(-\ln v\right)^\delta\right]^{\frac{1}{\delta}}\right). \tag{19}$$

It exhibits asymmetric tail dependence with zero lower tail dependence $\lambda_L = 0$ and upper tail dependence $\lambda_U = 2 - 2^{\frac{1}{\delta}}$. Note that overall dependence can be modeled only if upper tail dependence is non-zero, i.e. if $\delta > 1$.

(c) The Frank copula with generator $\varphi_F(q) = -\ln \frac{e^{-\vartheta \cdot q}-1}{e^{-\vartheta}-1}$ where $\vartheta \in (-\infty, \infty) \setminus \{0\}$ and

$$C_F(u,v) = -\frac{1}{\vartheta} \ln \left(1 + \frac{\left(e^{-\vartheta u} - 1\right)\left(e^{-\vartheta v} - 1\right)}{e^{-\vartheta} - 1} \right). \tag{20}$$

This copula is neither lower nor upper tail dependent, i.e. as for the independence copula we have: $\lambda_L = \lambda_U = 0$.

(d) In order to broaden the class of copula functions which may proof suitable for our modeling needs, we use a transformation rule as introduced by Nelsen (1999). The rule states that if φ is a generator and $\delta \geq 1$, then $\varphi_\delta(q) = \varphi(q)^\delta$ is also a generator.[4] Once we apply the transformation rule to the Frank copula C_F, the transformed Frank copula C_{TF} has generator $\varphi_{TF} = (\varphi_F)^\delta$ where the parameter vector is $\omega = (\vartheta, \delta) \in (-\infty, \infty) \setminus \{0\} \times [1, \infty)$. It follows that C_{TF} is given by:

$$\begin{aligned} C_{TF}(u,v) &= \varphi_{TF}^{-1}\left(\varphi_{TF}(u) + \varphi_{TF}(v)\right) \\ &= -\tfrac{1}{\vartheta} \ln \left[1 + \left(e^{-\vartheta} - 1\right) \exp\left[-\left(\left(-\ln\left[\tfrac{e^{-\vartheta \cdot u}-1}{e^{-\vartheta}-1}\right]\right)^\delta + \left(-\ln\left[\tfrac{e^{-\vartheta \cdot v}-1}{e^{-\vartheta}-1}\right]\right)^\delta \right)^{\frac{1}{\delta}} \right]\right]. \end{aligned} \tag{21}$$

It can be shown that the transformed Frank copula C_{TF} has zero lower tail dependence, $\lambda_L = 0$, while it is upper tail dependent with $\lambda_U = 2 - 2^{\frac{1}{\delta}}$; in contrast to the Gumbel copula C_G defined in (19), it allows for overall dependence even if upper tail dependence is zero with $\delta = 1$ where it follows $C_{TF} = C_F$. Also, C_{TF} converges to the Gumbel copula C_G for $\vartheta \to 0$; see Junker and May (2002) for details.

3 Empirical Analysis of Nonlinear Term Structure Dependence

Our program for the empirical analysis is as follows. First we briefly introduce the zero-coupon yield dataset. Then we estimate the term structure parameters of the two-factor model based on Kalman filter approach as outlined in Section 2.1.2. The empirical analysis of dependence between unpredictable innovations in the long end and the short end of the yield curve is based on an examination of our different parametric copula functions as given in Section 2.2.2. We argue that the theoretical properties of the copula functions

[4]Hence, the Gumbel copula C_G for example, follows immediately from the independence copula C_Π.

given above, jointly with careful empirical testing, allow us to identify a suitable model which is consistent with the dependence in the yield structure.

3.1 The Sample

As pointed out in the introduction, our empirical analysis of the U.S.-Treasury term structure is based on a sample of monthly zero-coupon yields. The yield observations are obtained from the refined Fama-Bliss zero-coupon yield dataset as introduced in Fama and Bliss (1987). The maturities range from one to five years. The sample covers the period October 1982 to December 2001 with 231 monthly observations. Of course, the amount of sample information comes with a trade-off concerning stationarity. We did therefore not extend the sample back to periods in the early eighties when a much different economic policy regime prevailed. Still, with the given dataset covering nearly twenty years, a check of robustness of the empirical results with respect to sample choice is important. We hence form two subsamples covering the October 1982 to December 1991 and the January 1992 to December 2001 period, which yields 111 monthly observations and 120 monthly observations, respectively.

In the following, we consider the monthly zero-coupon yields $R(t, \tau_i)$, $t = 0,\ \ldots,\ 230$ with τ_i denoting the i-year to maturity bond with $i = 1,\ 2,\ 3,\ 4,\ 5$. All yields are given on an annualized continuously compounded basis. The length of the discrete sampling interval, h, equals 1 month. Table 1 reports summary statistics for the entire sample period as well as for the two subperiods.

(Table 1 about here)

The statistics in Table 1 exhibit an on average increasing yield curve. The sample autocorrelation coefficients indicate typical first order linear dependence in monthly bond yields. Comparing the results for the subsamples with those for the entire period indicates lower levels of interest rates together with lower levels of volatility in yield changes for the second subsample.

3.2 Term Structure Model Estimation

This section presents the term structure estimation results based on Kalman filtering as outlined in Section 2.1.2. In general, the application of the Kalman filter requires the state process X to have normally distributed innovations which is typically violated in estimation applications. This means that the parameter estimation approach is based on a quasi-likelihood function.[5]

(Table 2 about here)

Based on the Kalman filter approach, maximum likelihood (ML) estimation yields an estimate of the parameter vector $\psi = (R_0, \kappa_1, \kappa_2, \gamma_1, \gamma_2, \sigma_1, \sigma_2, \rho)$ of the two-factor generalized Vasicek model (5). We assume a diagonal covariance structure of the measurement errors $\varepsilon(t, \tau)$ in (7) where the diagonal elements are denoted by $\sigma^2_{\varepsilon_i}$. The estimation results are given in Table 2. All parameter estimates contained in ψ, apart from those of γ_1, turn out to be significantly different from zero at usual confidence levels. The estimated standard errors $\widehat{\sigma}_{\varepsilon_i}$ are relatively homogeneous for all maturities with a slight tendency for larger measurement error variability for the 1-year maturity yields. These results are are in line with those of previous empirical research.

3.3 Derivation of Term Structure Innovations

Given the estimate of ψ in Table 2, we can derive unpredictable innovations for our term structure sample. By choosing two observable yields, namely the short end τ_s-year and the long end τ_l-year maturity yield, $R(t, \tau_s)$ and $R(t, \tau_l)$, the dynamics of the two-dimensional yield factor X can be expressed in terms of the estimated term structure parameters; see Appendix A.2 for details on the yield factor representation. For the Gaussian two-factor model it follows

$$\begin{pmatrix} R(t, \tau_s) \\ R(t, \tau_l) \end{pmatrix} = \mu_R + A_R \begin{pmatrix} R(t-1, \tau_s) \\ R(t-1, \tau_l) \end{pmatrix} + \begin{pmatrix} \epsilon_{\tau_s, t} \\ \epsilon_{\tau_l, t} \end{pmatrix}, \tag{22}$$

where μ_R and A_R are functions of the parameter vector ψ and especially, $(\epsilon_{\tau_s,t}\epsilon_{\tau_l,t})^\top$ is a linear transformation of $(Z_{\tau_s,t}Z_{\tau_l,t})^\top$, i.e. $(\epsilon_{\tau_s,t}\epsilon_{\tau_l,t})^\top = \mathcal{B}\ (Z_{\tau_s,t}Z_{\tau_l,t})^\top$. Accordingly, the

[5] Inference based on the Kalman iteration and likelihood maximization faces two specification issues. Firstly, the Kalman Filter estimates of X_t do not exactly correspond to the conditional expectations given the observed yields since the filter relies on a linear projection. Secondly, in a non-Gaussian model, the filtering errors –the differences between X_t and the linear projections– are not normally distributed. Brandt and He (2002) discuss the first-order approximation for non-normalities introduced by affine factor dynamics. Duan and Simonato (1999) discuss the estimation of square-root models by Kalman filtering and show in a simulation study that the biases are small, and, as also the results in Lund (1997) indicate, economically insignificant.

common joint distribution function $G(\epsilon_{\tau_s}, \epsilon_{\tau_l})$ of the innovations is completely determined by $H(z_{\tau_s}, z_{\tau_l})$ and $\mathcal{B}$ which is a function of ψ, see Appendix A.2 for derivation of $\mathcal{B}$.

In the following we choose the one-year and the five-year maturity yield, $R(t,1)$ and $R(t,5)$, to represent the short and the long yield factor, respectively. Based on the Kalman filter estimate $\widehat{\psi}$, we can derive the estimates $\widehat{\mu}_R$ and $\widehat{A}_R$. The time $t-1$ conditional expectation $\mathbb{E}_{t-1,\widehat{\psi}}$ is defined by equation (22). We then obtain the sequence of bivariate empirical yield innovations as

$$\begin{pmatrix} \widehat{\epsilon}_{1,t} \\ \widehat{\epsilon}_{5,t} \end{pmatrix} = \begin{pmatrix} R(t,1) - \mathbb{E}_{t-1,\widehat{\psi}}\{R(t,1)\} \\ R(t,5) - \mathbb{E}_{t-1,\widehat{\psi}}\{R(t,5)\} \end{pmatrix}, \quad t = 1, ..., T, \tag{23}$$

with $T = 230$.

3.4 Analysis of the Term Structure Innovations

Estimation of the copula parameter vector based on sample innovations is widely used in empirical research. Given the parameter estimates for the term structure dynamics above, the empirical marginal distribution functions are determined for each component of the bivariate yield innovation series (23). We thereby check for the joint normality assumption which is imposed by the ATSM. In the second step, parametric estimation of the copula functions is carried out. In order to avoid parametric model misspecification at the margins, we base our inference on the empirical marginal distributions and then derive the parameter estimates for the copula functions. The copula functions introduced in Section 2.2.2 are our respective candidate dependence models for the bivariate yield innovation series.

3.4.1 Distributional Properties of the Term Structure Innovations

The Kalman filtering estimates of Section 3.2 result in the bivariate yield innovation series $(\widehat{\epsilon}_{1,t}, \widehat{\epsilon}_{5,t})_{t=1,...,T}$ as defined by equation (23).

a) Univariate Properties

Closer inspection of the series' univariate distributional properties reveals that the assumption of uncorrelated normally distributed innovations appears to be a suitable approximation.

(Figure 1 about here)

(Figure 2 about here)

The empirical marginal (univariate) distribution functions $F_{\cdot,T}$ are determined for each component separately. They are given as

$$F_{\cdot,T}(x) = \frac{1}{T}\sum\nolimits_{t=1}^{T} I_{\{\cdot_t \leq x\}}. \tag{24}$$

Figure 1 shows QQ-plots of the marginal distributions of the innovations where the quantiles of the empirical distributions are plotted against those of the standard normal distribution. The two plots indicate a reasonably close approximation by the normal distribution, where the fit in the lower tail is better for the long maturity factor innovations. Additional results from the univariate chi-square test indicate that normality cannot be rejected with p-values of 0.40 for $\widehat{\epsilon}_1$ and 0.15 for $\widehat{\epsilon}_5$. We next consider the innovations' time-series properties. Figure 2 shows the sample autocorrelation functions for the univariate and the squared univariate series with lags up to order 23. The estimated autocorrelations for the raw innovations stay within the 95% confidence intervals with one exception in the $\widehat{\epsilon}_1$-plot which is an expected violation under the given confidence level. For the squared innovations we find six coefficients to exceed the 95% confidence interval in both plots. These irregular exceedances are evidence of some heteroskedasticity in both series, though they are weak as compared to what is otherwise typically observed for financial return series.

b) Multivariate Properties

While the above results support the marginal model assumptions of the affine model, a brief graphical analysis of the joint distribution of the factor innovations in Figure 3 casts doubt on the assumption of joint normality of the factor innovations.

(Figure 3 about here)

Figure 3 gives a scatterplot representation of the joint density of the yield innovations for the overall sample and the two subsamples. Apart from standard scatterplot representations of $(\widehat{\epsilon}_{1,t}, \widehat{\epsilon}_{5,t})$, the figure contains plots of the mapped observations

$$(\widehat{\epsilon}_{1,t}, \widehat{\epsilon}_{5,t})_{t=1,\ldots,T} \longmapsto (\widehat{u}_t, \widehat{v}_t)_{t=1,\ldots,T} = \left(F_{\epsilon_1,T}\left(\widehat{\epsilon}_{1,t}\right), F_{\epsilon_5,T}\left(\widehat{\epsilon}_{5,t}\right)\right)_{t=1,\ldots,T}. \tag{25}$$

As an empirical application of Sklar's theorem (13), the mapped observations are defined on the uniform space $[0,1] \times [0,1] = [0,1]^2$. The $(\widehat{\epsilon}_{1,t}, \widehat{\epsilon}_{5,t})$-plots in the figure reveal a somewhat stronger linear factor dependence during the second subsample period. All three $(\widehat{u}_t, \widehat{v}_t)$-plots reveal some clustering of observations in the upper right-hand corner of $[0,1]^2$ which indicates upper tail dependence. In contrast, graphical inspection gives little evidence of lower tail dependence as would be indicated by clustering in the lower left-hand corners of the plots. In the next three paragraphs, we estimate our parametric copula functions and analyze which models may capture the observed factor dependence.

3.4.2 Copula Estimation Methodology

Given the above results supporting the univariate model assumptions of the affine model, we assume in the following that the marginal distributions of the yield innovations are normal with mean zero and variances $\sigma_{1,\cdot}^2$ and $\sigma_{5,\cdot}^2$, respectively. Hence, we can write Sklar's Theorem (13) as

$$G_{\cdot}\left(\epsilon_1, \epsilon_5\right) = C_{\cdot}\left(N\left(\epsilon_1/\sigma_{1,\cdot}\right),\ N\left(\epsilon_5/\sigma_{5,\cdot}\right)\right), \tag{26}$$

where the notation ' $\cdot$ ' indicates the choice of one of the respective copula functions C_t, C_N, C_{TF} and C_G. For the different copula functions, the parameter vectors are given as $\omega_t = (\rho, \nu)$, $\omega_N = (\rho)$, $\omega_{TF} = (\vartheta, \delta)$, and $\omega_G = (\delta)$. We use ML-estimation to obtain simultaneous estimates of the parameters $(\omega_{\cdot}, \sigma_{1,\cdot}, \sigma_{5,\cdot})$ of the joint distribution function $G_{\cdot}$. Note that these estimates are optimal for the overall joint distributional assumption imposed by $G_{\cdot}$, which includes the marginal distributions as well as the dependence structure. With the joint density function $g_{\cdot}(\epsilon_1, \epsilon_5)$ derived from (26) the log likelihood function reads

$$\begin{aligned} &\ln L_{\cdot}\left(\omega_{\cdot}, \sigma_{1,\cdot}, \sigma_{5,\cdot};\ \epsilon_{1,t}, \epsilon_{5,t}\right) = \\ &= \sum_{t=1}^{T} \ln\left[\frac{1}{\sigma_{1,\cdot}\sigma_{5,\cdot}}\ c_{\cdot}\left(N\left(\epsilon_{1,t}/\sigma_{1,\cdot}\right),\ N\left(\epsilon_{5,t}/\sigma_{5,\cdot}\right)\right)\ N'\left(\epsilon_{1,t}/\sigma_{1,\cdot}\right)\ N'\left(\epsilon_{5,t}/\sigma_{5,\cdot}\right)\right], \end{aligned} \tag{27}$$

where N' denotes the density of the standard normal distribution and $c_{\cdot}$ is one of our respective copula densities. As the estimates of the copula parameters $\widehat{\omega}_T$ will have ML-properties, the estimates of the tail dependence parameters, $\widehat{\lambda}_T = \lambda(\widehat{\omega}_T)$ will be consistent and asymptotically normally distributed with

$$\sqrt{T}(\widehat{\lambda}_T - \lambda) \xrightarrow{d} N\left(0, \sigma_\lambda^2\right). \tag{28}$$

Given that $\lambda(\cdot)$ are suitably smooth functions, the variance σ_{λ}^2 above can be approximated by a first order Taylor series expansion of the form

$$\sigma_{\widehat{\lambda}}^2 = \sum_{i=1}^{p} \left(\frac{\partial \lambda(\widehat{\omega}_T)}{\partial \omega_i} \right)^2 \sigma_{\widehat{\omega}_i}^2,$$

where $p = \#\{\omega\}$ denotes the number of parameters of the copula model and $\sigma_{\widehat{\omega}_i}$ denotes the ML standard errors of the respective copula parameter estimates.

3.4.3 Diagnostics for the Estimated Copula Models

Based on our estimates of the parametric copulas, we compare the in-sample model fit based on a set of different goodness-of-fit test procedures. These include seven statistics which are given as follows.

A general goodness-of-fit test is the bivariate version of the well-known χ^2-test which in our application is defined on the space $[0,1]^2$. With a grid of k cells $c_i \subset [0,1]^2$ of identical size, we calculate the chi-square test statistic as

$$\chi_{\mathrm{df}}^2 = \sum_{i=1}^{k} \left(\frac{\mathbb{E}_{G.}(c_i) - \#\{(u_t, v_t),\ t = 1, ..., T : (u_t, v_t) \in c_i\}}{\mathbb{E}_{G.}(c_i)} \right)^2, \tag{29}$$

with $\mathbb{E}_{G.}(c_i)$ denoting the number of expected observations in a cell c_i under the model $G.$. The statistic follows a χ^2-distribution with $k-1-\#\{(\omega_., \sigma_{1,.}, \sigma_{5,.})\}$ degrees of freedom (df). Based on a grid of $6 \times 6 = 36$ cells for the overall sample, this implies 32 degrees of freedom for the normal and the Gumbel copula and 31 degrees of freedom for the t-copula and the transformed Frank copula. For the two subsamples, we use a grid of $4 \times 4 = 16$ cells and adjust the degrees of freedom accordingly.

Three additional tests of the overall model fit are based on the maximized log-likelihood function $\ln L.$. These include the Akaike information criterion, $AIC = -2\ln L. + 2p$, and the Bayesian information criterion, $BIC = -2\ln L. + p\ln T$. While the Bayesian criterion puts a heavier penalty on the number of model parameters, both statistics are based on probability of the observations within a given model. In contrast to that, the entropy criterion measures the probability of a given model. The model entropy is given as the expected value of the negative logarithm of the maximized density function, $EN = \mathbb{E}(-\ln g.(\epsilon_1, \epsilon_5))$, where we approximate the expectation by Monte Carlo simulation.

The χ^2-test can be interpreted as a measure of the differences in the densities. Alternative goodness of fit statistics in the literature, such as the Kolmogorov/Smirnov-test, are based on distances of observed deviations between the empirical and the parametric distribution function. As both, the empirical and the theoretical distribution function have to converge to zero at the lower tail and one at the upper tail, their representation of fit in the tails is weak by construction. A test statistic which is superior in this respect dates back to Anderson and Darling (1952). The Anderson and Darlington test uses relative instead of absolute deviations between the distribution functions and thereby gives a better representation of the fit in the tails. We use the integrated version outlined by the authors and denote it by AD.

Our last goodness-of-fit diagnostic particularly focuses on the fit in the distribution tails. The AD statistic, due to its use of the cumulative distribution function, has the drawback of a smoothing effect particularly present in the upper tail. When considering model fit in the tails, we therefore apply a diagnostic which is based not on the overall probability deviations, but on the probability deviations at a particular quantile of the joint distribution function. Let C_T denote the empirical copula function

$$C_T(u,v) = \frac{1}{T}\#\left\{(u_t, v_t),\ t = 1, ..., T : (u_t, v_t) \leq (u, v)\right\}, \tag{30}$$

and $\overline{C}_T$ denote the empirical survival copula function

$$\overline{C}_T(u,v) = \frac{1}{T}\#\left\{(u_t, v_t),\ t = 1, ..., T : (u_t, v_t) > (u, v)\right\}. \tag{31}$$

With these empirical copulas we measure deviations at the upper and the lower tail independently. The relative lower tail probability deviation PD_p is defined as the deviation of the model probabilities from the empirical probabilities measured at a point (q, q) in the lower corner of the set $[0, 1]^2$. Here, $q = C_T^{\leftarrow}(p)$, $0 \leq p \leq 1$ and $C_T^{\leftarrow}$ is the inverse of the diagonal section of the empirical copula function C_T. The relative lower tail probability deviation PD_p is given as

$$PD_p = \frac{G.\left(F_{\epsilon_1,T}^{\leftarrow}\left(C_T^{\leftarrow}(p)\right),\ F_{\epsilon_5,T}^{\leftarrow}\left(C_T^{\leftarrow}(p)\right)\right) - p}{p}. \tag{32}$$

Based on survival functions, the upper tail relative probability deviation is defined by the survival probability deviation $\overline{PD}_p$ at the point (q, q) in upper corner of $[0, 1]^2$. It is given as

$$\overline{PD}_p = \frac{\overline{G}.\left(\overline{F}_{\epsilon_1,T}^{\leftarrow}\left(\overline{C}_T^{\leftarrow}(p)\right),\ \overline{F}_{\epsilon_5,T}^{\leftarrow}\left(\overline{C}_T^{\leftarrow}(p)\right)\right) - p}{p}. \tag{33}$$

Setting p equal to a small positive value, the probability deviations PD_p and $\overline{PD}_p$ allow us to measure deviations in the tails. Note that pT observations are available for the calculation of the empirical distribution function and –given that pT is sufficiently large– ensure convergence towards the theoretical distribution function.

3.4.4 Copula Estimation Results

Tables 3, 4 and 5 give the estimation results for the overall sample 1982-2001 and the two subsamples 1982-1991 and 1992-2001, respectively. We give the standard deviation estimates for the marginal distributions as well as the estimates of the parameters in the copula parameter vectors ω_t, ω_N, ω_{TF}, and ω_G. In Table 6 we compare the goodness-of-fit for the competing copula models for the overall sample as well as for the two subsamples. For the evaluation we use the seven statistics χ^2, AIC, BIC, EN, AD, PD and $\overline{PD}$ as defined above.

(Table 3 about here)

(Table 4 about here)

(Table 5 about here)

Table 3 summarizes the estimation results for the joint distribution functions in the overall sample period. The normal copula yields an estimate of the correlation coefficient of 0.85 which indicates a quite strong positive linear dependence in the yield factors. The subsample results in Tables 4 and 5 indicate comparable linear dependence, with an estimate of 0.89 in the first and 0.79 in the second subsample. The results also show that the estimates of the standard deviations for the marginal distributions vary somewhat depending on the copula model, where the Gumbel copula assigns the largest standard deviations to the margins. Considering the subsamples, this highlights that the standard deviations in the yield factor realizations were higher in earlier subsample period 1982-1991 than in the later 1992-2001 period. This is also visible in the plots of Figure 3. A graphical illustration of the estimated 1982-2001 joint distribution functions under the different copula functions is given in Figure 4. The plotted joint density contour lines visualize the dependence implications of the different copulas, while the plotted yield innovations allow for a first visual inspection of model fit.

(Figure 4 about here)

(Table 6 about here)

We next turn to the goodness-of-fit tests in Table 6. Given the size of the data set we have to point out that it is impossible to strictly reject any of the copula models. However, it turns out that the transformed Frank copula shows best overall fit.

Starting with the overall sample, our results clearly indicate that the transformed Frank copula is the superior dependence model. All seven statistics including those which penalize for the number of model parameters (χ^2, AIC and BIC) happen to favour the C_{TF}-model where the second best model follows with some diagnostic distance. Considering the symmetric models, the student-t shows advantages in the chi-square and the entropy statistic, but not for AIC and BIC. It is remarkable that the AD test always shows very high deviations values for the student model as compared to the other models. Considering Figure 4, an explanation for this finding is that the contours of the student model narrow most quickly in the overall region of the lower left quadrant $[-\infty, 0] \times [-\infty, 0]$ of the joint distribution function causing large relative deviations in the empirical versus theoretical distributions for moderate negative values. Additionally, due to the sample size, the AD statistic should be interpreted with some caution; the number of 230 observations may not fully guarantee convergence of the empirical distribution functions, which is a requirement for the AD statistic. For the probability deviations in the tails, PD and $\overline{PD}$, we choose $p = 0.05$ which yields 11 observations for the calculation of the marginal distribution function. The PD and $\overline{PD}$ results indicate that the transformed Frank copula has lowest deviations from the empirical observations in the upper as well as the lower tail. The symmetric models tend to overestimate the probability of observations in the lower tail, which is demonstrated by large positive values for PD. The Gumbel copula shows a tendency to overestimate the probability of observations in the upper tail showing a large deviation $\overline{PD}$; note that the Gumbel copula models overall dependence and upper tail dependence jointly via the δ parameter which implies strong upper tail dependence under strong linear dependence and vice versa.

We next turn to the subsamples, i.e. the 1982-1991 and 1992-2001 subsample results. Note that the PD and $\overline{PD}$ statistics are now based on $p = 0.1$ which, under a subsamples sizes of roughly $T/2$, implies a number of tail observations roughly equal to these for the full sample diagnostics. For the subsamples, the assignment of the best goodness-of-fit statistics in Table 6 shows notable variations across the models. This is due to the substantial decrease in sample size, which makes the interpretation of the results less conclusive than for the overall sample. However, the transformed Frank copula still obtains the best results when evaluated by the number of best fit-results among all models. Also, the C_{TF}-copula function is never assigned one of the worst-fit results, which does not hold for the other models. Given a smaller data set, the statistics χ^2, AIC and

BIC considerably penalize the two-parameter copula functions. At the same time, the C_t parameter estimate of ν in Table 4 as well as the C_{TF} parameter estimate of δ in Table 5, exhibit high standard errors with corresponding low respective t-values. A relatively stable pattern in Table 6 is provided by the $\overline{PD}$ statistics; high deviations for the overall sample as well as for the subsamples point out that a drawback of the Gumbel copula is its tendency to overestimate the upper tail.

Turning to the first subsample, 1982-1991, C_{TF} shows the best fit according to the EN, AD and PD statistics. The one-parameter copula models C_N and C_G also perform relatively well. The normal model even gives best fit according to the chi-square test statistic. In the second subsample, 1992-2001, as mentioned above, we report lower estimates of the volatility in the marginal distributions. However, our results do not indicate that the dependence structure is much different in the two subsamples. The C_{TF}-copula again yields results better than for the other models with best fit as measured by the BIC, the EN and the upper tail fit $\overline{PD}$. The student-t model has the worst AD statistic; still it has the best fit according to the χ^2 and AIC measures. The normal copula performs notably well having the best AD statistic and none of the worst fir results. To summarize the subsample comparison results, we can state that – given a high variability in the statistics – the Gumbel copula provides a second best fit in the first subperiod while the normal copula provides a second best fit in the second subperiod. In both subperiods however, the statistics indicate that the transformed Frank copula has best overall fit.

4 Application: Measuring Bond Portfolio Risk

Based on the affine term structure model of Section 2.1.2, the term structure of interest rates is completely described by two risk factors. Clearly, the dependence characteristics of the joint distribution of the 1-year yield and the 5-year yield influences the risk measurement of portfolios. In this section we analyze the different distributional specifications' impact on risk management decisions.

We start with a graphical illustration of the copula function estimation results of Section 3. Figure 5 contains four plots of the fitted conditional densities of the 5-year yield given a fixed realization of the 1-year yield. Each plot represents one of the four different copula models. As can be seen, the conditional densities show large structural differences especially including the probability of joint upper or lower tail events. For example, given a negative shock to the short rate of –0.02 in the transformed Frank model, the conditional density for the five year yield has high variance while, given a positive shock, the conditional density has low variance. As is illustrated, the structure looks much different for the symmetric models.

(Figure 5 about here)

We next apply the estimated dependence structures to quantify the risk of different bond portfolios and compare the results. We consider portfolios which invest in the 1-year zero bond at price $P(t, 1)$ and the 5-year zero bond at price $P(t, 5)$ and then study the return of this investment after 1 month of time. Denote by R_Λ the return of the portfolio which has initial portfolio duration Λ. As a risk measure ϱ we utilize Value-at-Risk (VaR), i.e. the quantile of the profit-and-loss distribution of the portfolio. When adjusted for the expected portfolio return VaR is

$$\varrho = F^{\leftarrow}_{R_\Lambda}(\alpha) - \mathbb{E}\{R_\Lambda\},$$

where α is the confidence level and F_{R_Λ} is the cumulative distribution function of R_Λ. We introduce the superscript '+' to the risk measure ϱ when measuring the risk of a long position, and the superscript '−' for measuring the risk of a short position, respectively. Additionally, the subscript at the risk measure ϱ indicates the copula applied for defining the dependence structure. The confidence levels we discuss are $\alpha = 99\%$ and $\alpha = 99.9\%$ for which the VaR numbers ϱ^+_N, ϱ^+_t, ϱ^+_G, ϱ^+_{TF}, and ϱ^-_N, ϱ^-_t, ϱ^-_G, ϱ^-_{TF} are calculated.

We compare the risk measures ϱ by fixing the risk measure induced by the normal copula which is the standard risk measure, and calculate the relative deviations from this measure. The relative deviations are

$$\Delta_\cdot = \frac{\varrho_\cdot - \varrho_N}{\varrho_N},$$

where' $\cdot$ ' indicates the choice of one of the three copula functions C_t, C_{TF}, and C_G. The relative deviations $\Delta_\cdot$ for long and short bond portfolio holdings as a function of the initial duration Λ are plotted in Figure 6.

(Figure 6 about here)

The results for holding a long position in the interest rate portfolio $\Delta^+_\cdot$ are displayed at the top of Figure 6. The results for t-copula model Δ^+_t are indicated by the solid line. As can be seen at the top left in Figure 6, the t-copula produces VaR numbers which are close to the normal copula model for the 99% confidence level. When increasing the confidence level to 99.9% at the top right in Figure 6, the maximum relative deviation increases from 0.2% to 3% which reflects the property of the t-copula to adopt to the (upper) tail dependence existing in our data set. A similar pattern is observed for the model given by transformed Frank copula (see the dashed line). The relative deviation Δ^+_{TF} takes a maximum value of approximately 4% for the 99.9% confidence level, and the VaR is persistently larger than the numbers based on the t-copula model. The Gumbel copula (dashed-dotted line) generates the highest VaR. The maximum relative deviation Δ^+_G is 2.5% for the 99% confidence level, and around 4.5% for the 99.9% confidence level, respectively.

At the bottom of Figure 6 the results $\Delta^-_{\cdot}$ are shown for a short position in bond portfolios with initial duration Λ. The relative deviation of the t-copula model Δ^-_t has similar characteristics as in the case of the long position. The VaR turns out to be relatively close to the VaR given under the normal copula, where the positive deviations tend to become overall larger when the confidence level is increased from 99% to 99.9%. The t-copula quantiles exceed the normal ones because the t-copula features lower tail dependence which is not present in the data. In contrast to the t-copula, the transformed Frank copula and the Gumbel copula both produce negative relative deviations of the VaR measures when compared to the normal copula. The maximum relative deviation is around 3% for the 99% confidence level and around 6% for the 99.9% confidence level.

The above findings can be interpreted as follows. In Section 3.4.4, the transformed Frank copula proved to be the dependence model which reflects the observed dependence structure in the most appropriate way. Assuming that the data are generated by a joint distribution with normal margins and a transformed Frank copula then implies that the normal copula produces a systematic bias in measured VaR.

For long bond portfolio positions, the normal copula tends to underestimate VaR where the lack in risk capital may approximately amount to up to 4% in our example. Clearly, the negative bias in VaR produced by the normal copula is related to the upper tail dependence which is present in the data but not a characteristic of the normal dependence model. The t-copula results in VaR numbers which are much closer to the transformed Frank numbers than the normal numbers with a maximum deviation of approximately 1%. This finding is due to the upper tail dependence which is incorporated in the t-copula. The Gumbel copula features characteristics which are present in the analyzed data set: upper tail dependence and asymmetry. The VaR numbers are of reasonable quality especially for the high confidence levels of 99.9%. For the 99% confidence level, the Gumbel copula produces the maximum relative deviation to transformed Frank numbers. For this particular case, the Gumbel copula performs poorly compared to the alternative dependence models (Figure 6, top, left).

For short bond portfolio positions, the normal copula overestimates VaR. The maximum relative deviation takes a value around 6%. Though the data were not found to show lower lower tail dependence in Section 3.4.4, which is in accordance with the normal copula, bias is again present having opposite sign. The explanation for this finding is reasonably simple. As the upper and lower tail of the normal copula are estimated simultaneously, the realized estimate is a balanced result of both shortcomings of the normal copula, namely its of lack of tail dependence and its symmetry characteristic. Also, due to its symmetric structure, the absolute biases generated by the t-copula are high when compared to the transformed Frank model. Hence, the t-copula turns out to produce overestimated VaR numbers for short positions. As it turns out, even for moderate confidence levels of 99%, the copula functions' ability to reproduce a complex observed dependence structure becomes important: The Gumbel copula is a parsimoniously parameterized model which captures upper tail dependence. The relative deviations for the best-fitting transformed Frank model indicate that the normal model can produce VaR biases of up to 6% in the given example.

5 Conclusion

As is well-known, the concept of linear dependence breaks down under non-normality. However, as the present investigation documents, statistical theory offers more flexible models of dependence which are relevant to financial modeling.

Based on the benchmark model given by the affine class of term structures which assumes joint normality in yield innovations, this paper analyses cross-sectional nonlinearity in the term structure of U.S.-Treasury yields. The nonlinearities documented in the data represent a profound statistical characteristic which is shown to be of economic significance as well. Deviations from linear dependence have implications on risk management when financial risk is for example measured by the commonly used VaR methodology. Most strikingly, we conclude that the normal copula as a benchmark model of dependence imposes two main problems, namely absence of tail dependence and symmetry, which both prevent accurate risk measurement. Our findings are not limited to bond pricing and bond portfolio VaR applications. The model bias due to the normality assumption should be even more pronounced when the pricing implications for nonlinear contracts, e.g., for interest rate derivatives, are considered.

A Appendix

A.1 Affine Models

We briefly review the concept of ATSMs following Duffie and Kan (1996), and Dai and Singleton (2000). In contrast to the considerations above, we start with an equivalent martingale measure Q, and later on we work out the link to the real world under the so-called physical measure P. Absent arbitrage opportunities, the time t price of a zero-coupon bond that matures at time $t+\tau$, denoted $P(t,\tau)$, is given by

$$P(t,\tau) = \mathbb{E}_t^Q \left\{ \exp\left(- \int_t^{t+\tau} R(s)\, ds \right) \right\} , \tag{34}$$

where $\mathbb{E}_t^Q\{\,\cdot\,\}$ denotes a conditional expectation under the risk neutral measure Q. A d-factor affine term structure model is obtained under the assumption that the instantaneous short rate R is an affine function of a d-dimensional vector process of state variables $X = (X_1, ..., X_d)$:

$$R(t) = R_0 + a_1 X_1(t) + ... + a_d X_d(t) = R_0 + a^\top X(t)\,, \tag{35}$$

and that X follows an affine diffusion:

$$dX(t) = \kappa_Q(\theta_Q - X(t))\, dt + \Sigma \sqrt{S(t)}\, dW_Q(t)\,, \tag{36}$$

where W_Q is a d-dimensional standard Brownian motion under the measure Q, θ_Q is a d-vector, κ_Q and Σ are $d \times d$ matrices, and $S(t)$ is a $d \times d$ diagonal matrix with the ith diagonal element:

$$\{S(t)\}_{ii} = \alpha_{S\,i} + {\beta_{S\,i}}^\top X(t)\,, \tag{37}$$

Provided that a parameterization is admissible, we know from Duffie and Kan (1996):

$$P(t,\tau) = \exp\left(A(\tau) + B(\tau)^\top X(t) \right) , \tag{38}$$

where A and B are satisfy the ordinary differential equations:

$$\frac{dA(\tau)}{d\tau} = \theta_Q^\top \kappa_Q^\top B(\tau) + \frac{1}{2} \sum_{i=1}^{d} \left[\Sigma^\top B(\tau) \right]_i^2 \alpha_{S\,i} - R_0\,, \tag{39}$$

$$\frac{dB(\tau)}{d\tau} = -\kappa_Q^\top B(\tau) + \frac{1}{2} \sum_{i=1}^{d} \left[\Sigma^\top B(\tau) \right]_i^2 \beta_{S\,i} - a\,. \tag{40}$$

The particular specification of the Q-dynamics of X in Equation (36) and the definition of R in Equation (35) allow the exponential affine representation of the bond prices in Equation (38). It is well known that the measure Q is generated by a change of measure with respect to the empirical/physical measure P that describes the behavior of the stochastic factors in the "real world". To obtain an affine structure for X under both measures P and Q, we restrict the measure change in terms of the market price of risk Λ to:

$$\Lambda(t) = \sqrt{S(t)}\gamma \,, \tag{41}$$

where γ is a d-vector of constants. Thus, the P-dynamics of the state process X are:

$$dX(t) = \kappa(\theta - X(t))\, dt + \Sigma\sqrt{S(t)}\, dW(t) \,, \tag{42}$$

where W is a d-dimensional standard Brownian motion under P and:

$$\kappa = \kappa_Q - \Sigma\Psi \,, \tag{43}$$

$$\theta = \kappa^{-1}\left(\kappa_Q\theta_Q + \Sigma\phi\right). \tag{44}$$

The ith row of the $d\times d$ matrix Ψ is filled in by $\gamma_i {\beta_S}_i^\top$ and ϕ is a d-vector whose ith element is given by $\gamma_i {\alpha_S}_i$. The functions A and B describing the two-factor generalized Vasicek model of Babbs and Nowman (1998) are given by the parameters $R_0 \in \mathbb{R}$, $\kappa_1, \kappa_2, \sigma_1, \sigma_2 \in \mathbb{R}^+$, $\rho \in]-1, 1[$, and $\gamma_1, \gamma_2 \in \mathbb{R}$ in the following way:

$$A(\tau) = \sum_{i=1,2}\left[-\pi_i\left(B_i(\tau) + \tau\right) - \frac{\sigma_i^2}{4\,\kappa_i}\, B_i(\tau)^2\right] - R_0\tau + A_\rho(\tau) \,, \tag{45}$$

$$B(\tau) = \left(B_1(\tau), B_2(\tau)\right)^\top = \left(\frac{e^{-\kappa_1\tau} - 1}{\kappa_1}, \frac{e^{-\kappa_2\tau} - 1}{\kappa_2}\right)^\top \,, \tag{46}$$

where $\pi_1 = \frac{\sigma_1\,\gamma_1}{\kappa_1} - \frac{\sigma_1^2}{2\,\kappa_1^2}$, $\pi_2 = \frac{\sigma_2\,(\rho\,\gamma_1 + \sqrt{1-\rho^2}\,\gamma_2)}{\kappa_2} - \frac{\sigma_2^2}{2\,\kappa_2^2}$, and:

$$A_\rho(\tau) = \frac{\rho\,\sigma_1\,\sigma_2}{\kappa_1 + \kappa_2}\left[\frac{1}{\kappa_1}(B_1(\tau) + \tau) + \frac{1}{\kappa_2}(B_2(\tau) + \tau) - B_1(\tau)\,B_2(\tau)\right]. \tag{47}$$

A.2 Yield-Factor Representation

A distinct feature of ATSM framework is that the latent state variables can be transferred to an appropriate set of yields, see Duffie and Kan (1996). Moreover, the affine structure of the latent variables is preserved for the yields, and the yields can be viewed at as a new set of state variables provided some technical conditions hold. Given a d-factor ATSM with state variable $X = (X_1, ..., X_d)^\top$. For a set of maturities $(\tau_1, ..., \tau_d)$ the corresponding yields $Y = (Y_1, ..., Y_d)^\top$ are given by Equation (38):

$$Y(t) = A + B\,X(t) \tag{48}$$

where

$$A = \begin{pmatrix} -\frac{A(\tau_1)}{\tau_1} \\ \vdots \\ -\frac{A(\tau_d)}{\tau_d} \end{pmatrix}, \quad B = \begin{pmatrix} -\frac{B_1(\tau_1)}{\tau_1} & \dots & -\frac{B_d(\tau_1)}{\tau_1} \\ \vdots & \ddots & \vdots \\ -\frac{B_1(\tau_d)}{\tau_d} & \dots & -\frac{B_d(\tau_d)}{\tau_d} \end{pmatrix}. \tag{49}$$

Provided B is non-singular, we can state the analogue of Equation (42), i.e. the state equation for the yield vector Y:

$$dY(t) = \tilde{\kappa}(\tilde{\theta} - Y(t))\,dt + \widetilde{\Sigma}\sqrt{\widetilde{S}(t)}\,dW(t)\,, \tag{50}$$

where

$$\tilde{\kappa} = B\,\kappa\,B^{-1}\,, \quad \tilde{\theta} = B\,\theta + A\,, \quad \widetilde{\Sigma} = B\,\Sigma\,, \text{ and } \widetilde{S}(t) = \{\tilde{\alpha}_i + \tilde{\beta}_i^\top Y(t)\}\,, \tag{51}$$

and $\tilde{\alpha}_i = \alpha_{S\,i} - {\beta_{S\,i}}^\top B^{-1} A$, and $\tilde{\beta}_i = {B^{-1}}^\top \beta_{S\,i}$. We briefly discuss the yield dynamics implied by the two-factor generalized Gaussian model. Here, we find $\theta = 0$, and $S(t) = \mathbb{I}_d$ what results into $\tilde{\kappa} = B\,\kappa\,B^{-1}$, $\tilde{\theta} = A$, $\widetilde{\Sigma} = B\,\Sigma$, and $\widetilde{S}(t) = \mathbb{I}_d$:

$$dY(t) = B\,\kappa\,B^{-1}(A - Y(t))\,dt + B\,\Sigma\,dW(t)\,, \tag{52}$$

where $B\,\kappa\,B^{-1}$ describes the mean reversion including cross-dependencies between Y_1 and Y_2, and the covariance is given by $B\,\Sigma\,\Sigma^\top\,B^\top$.

Setting $(R(\cdot, \tau_s), (R(\cdot, \tau_l))^\top = (Y_1, Y_2)$, it follows from the affine structure

$$\begin{pmatrix} R(t, \tau_s) \\ R(t, \tau_l) \end{pmatrix} = \mathcal{A} + \mathcal{B}X_t = \begin{pmatrix} -A(\tau_s)/\tau_s \\ -A(\tau_l)/\tau_l \end{pmatrix} + \begin{pmatrix} -B(\tau_s)^\top/\tau_s \\ -B(\tau_l)^\top/\tau_l \end{pmatrix} X_t.$$

Due to the autoregressive structure of X in (12) this results in

$$\begin{pmatrix} R(t, \tau_s) \\ R(t, \tau_l) \end{pmatrix} = (\mathbf{I}_2 - \mathcal{B}\,A\,\mathcal{B}^{-1})\,\mathcal{A} + \mathcal{B}\,A\,\mathcal{B}^{-1} \begin{pmatrix} R(t-1, \tau_s) \\ R(t-1, \tau_l) \end{pmatrix} + \mathcal{B} \begin{pmatrix} Z_{\tau_s,t} \\ Z_{\tau_l,t} \end{pmatrix},$$

which is Equation (22) where $\mu_R = (\mathbf{I}_2 - \mathcal{B}\,A\,\mathcal{B}^{-1})\,\mathcal{A}$ and $A_R = \mathcal{B}\,A\,\mathcal{B}^{-1}$.

References

[1] **Ahn, D. H., Dittmar, R. F., Gallant A. R. (2002)**: Quadratic Term Structure Models: Theory and Evidence, *Review of Financial Studies* 15: 243-288

[2] **Aït-Sahalia, Y. (1996)**: Testing Continuous-Time Models of the Spot Interest Rate, *Review of Financial Studies* 9: 385-426

[3] **Anderson, T. W., Darling, D. A. (1952)**: Asymptotic Theory of Certain Goodness of Fit Criteria Based on Stochastic Processes, *Annals of Mathematical Statistics* 23: 193-212

[4] **Ané, T., Kharoubi, C. (2001)**: Dependence Structure and Risk Measure, *Journal of Business*, forthcoming

[5] **Ang, A., Bekaert, G. (2000)**: Short Rate Nonlinearities and Regime Switches, *Journal of Economic Dynamics and Control*, forthcoming

[6] **Babbs, S. H., Nowman, K. B. (1998)**: Econometric Analysis of a Continuous Time Multi-Factor Generalized Vasicek Term Structure Model: International Evidence, *Asia-Pacific Financial Markets* 5: 159-183

[7] **Babbs, S. H., Nowman, K. B. (1999)**: Kalman Filtering of Generalized Vasicek Term Structure Models, *Journal of Financial and Quantitative Analysis* 34: 115-130

[8] **Björk, T., Kabanov, Y., Runggaldier, W. (1997)**: Bond Market Structure in the Presence of Marked Point Processes, *Mathematical Finance* 7: 211-239

[9] **Brandt, M. W., He, P. (2002)**: Simulated Likelihood Estimation of Affine Term Structure Models from Panel Data, Working Paper, University of Pennsylvania

[10] **Chan, K. C., Karolyi, G. A., Longstaff, F. A., Sanders, A. B. (1992)**: An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance* 47: 1209-1227

[11] **Chen, R., Scott, L. (1992)**: Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure, *Review of Financial Studies* 5: 613-636

[12] **Cox, J. C., Ingersoll, J. E., Ross, S. A. (1985)**: A Theory of the Term Structure of Interest Rates, *Econometrica* 53: 385-407

[13] **Dai, Q., Singleton, K. J. (2000)**: Specification Analysis of Affine Term Structure Models, *Journal of Finance* 55: 1943-1978

[14] **Dai, Q., Singleton, K. J. (2002)**: Fixed Income Pricing, Mimeo, NYU and Stanford University

[15] **De Jong, F. (2000)**: Time-Series and Cross-Section Information in Affine Term-Structure Models, *Journal of Business and Economic Statistics* 18: 300-314

[16] **Dewachter, H., Lyrio, M., Maes, K. (2002)**: The Effect of Monetary Union on German Bond Markets, Working Paper, University of Leuven

[17] **Duan, J. C., Simonato, J. G. (1999):** Estimating Exponential-Affine Term Structure Models by Kalman Filter, *Review of Quantitative Finance and Accounting* 13: 111-135

[18] **Duffie, D., Kan, R. (1996)**: A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6: 379-406

[19] **Eberlein, E., Raible, S. (1999)**: Term Structure Models driven by General Lévy Processes, *Mathematical Finance* 9: 31-53

[20] **Embrechts, P., McNeil, A., Straumann, D. (2002)**: Correlation and Dependence in Risk Management: Properties and Pitfalls, Preprint, ETH Zürich

[21] **Fama, E. F., Bliss, R. R. (1987)**: The Information in Long-Maturity Forward Rates, *American Economic Review* 77: 680-692

[22] **Harvey, A. C. (1989)**: Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge

[23] **Joe, H. (1997)**: Multivariate Models and Dependence Concepts, Chapmann & Hall, London

[24] **Junker, M., May, A. (2002)**: Measurement of Aggregate Risk with Copulas, Preprint, Research Center Caesar, Bonn

[25] **Longstaff, F. A., Schwartz, E. S. (1992)**: Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *Journal of Finance* 47: 1259-1282

[26] **Lund, J. (1997)**: Non-Linear Kalman Filtering Techniques for Term-Structure Models, Working Paper, Aarhus School of Business

[27] **Malevergne, Y., Sornette, D. (2002)**: Testing the Gaussian Copula Hypothesis for Financial Assets Dependences, Working Paper, UCLA

[28] **Nelsen, R. B. (1999)**: An Introduction to Copulas, Springer, New York

[29] **Patton, A. J. (2001)**: Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula, Working Paper, U.C. San Diego

[30] **Rockinger, M., Jondeau, E. (2001)**: Conditional Dependency of Financial Series: An Application of Copulas, CEPR and HEC Working Paper

[31] **Scaillet, O. (2002)**: A Nonparametric Analysis of Stock Index Return Dependence through Bivariate Copulas, *European Investment Review* 1: 7-16

[32] **Stanton, R. (1997)**: A Nonparametric Model of the Term Structure Dynamics and the Market Price of Interest Rate Risk, *Journal of Finance* 52: 1973-2002

[33] **Vasicek, O. A. (1977)**: An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5: 177-188

[34] **White, H. (1982)**: Maximum Likelihood Estimation of Misspecified Models, *Econometrica* 50: 1-25

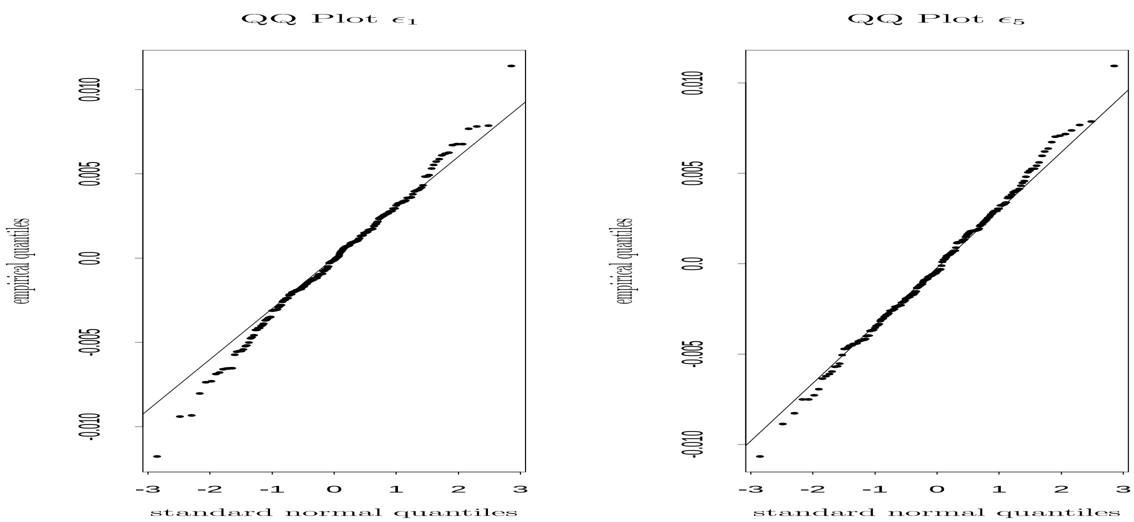

Figure 1: QQ-plots of the marginal distributions of $\widehat{\epsilon}_1$ and $\widehat{\epsilon}_5$ each against the standard normal distribution. Sample period 1982 to 2001.

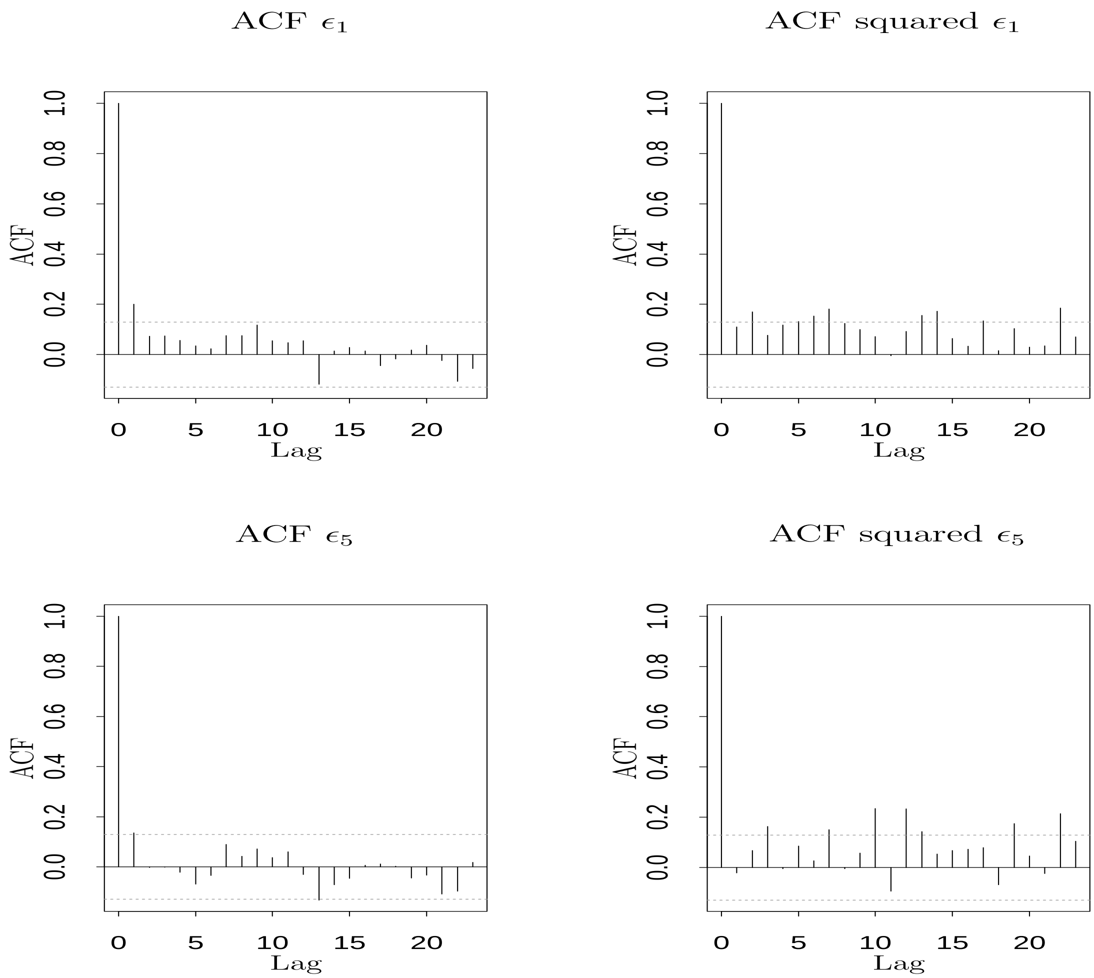

Figure 2: Plots of the autocorrelations of the innovations $\widehat{\epsilon}_1$ and the squared innovations $\widehat{\epsilon}_5^2$. The dotted straight line indicates the 95% confidence interval under the Null of uncorrelatedness. Sample period 1982 to 2001.

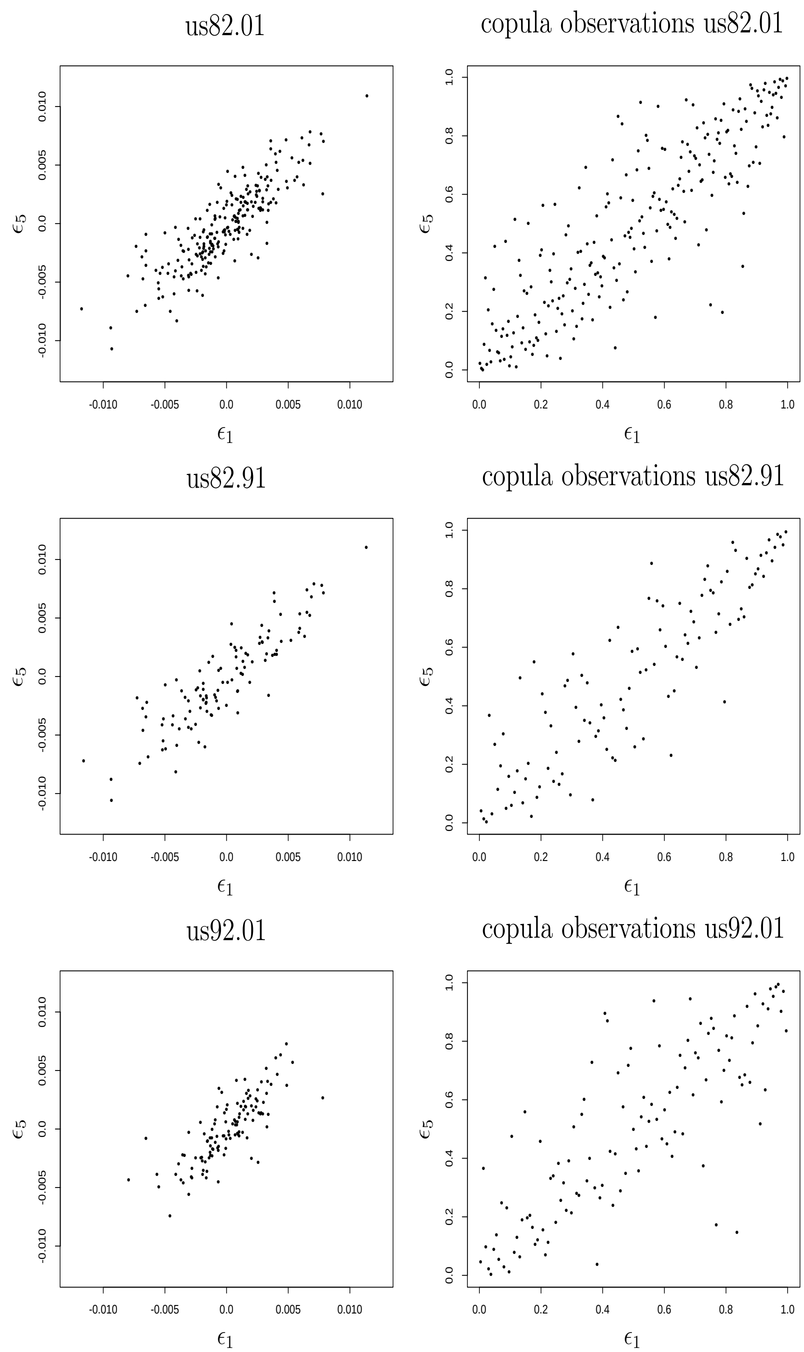

Figure 3: Scatterplot representation of the yield innovations and the yield innovation copula densities. Sample period 1982 to 2001 and subsamples 1982 to 1991 and 1992 to 2001.

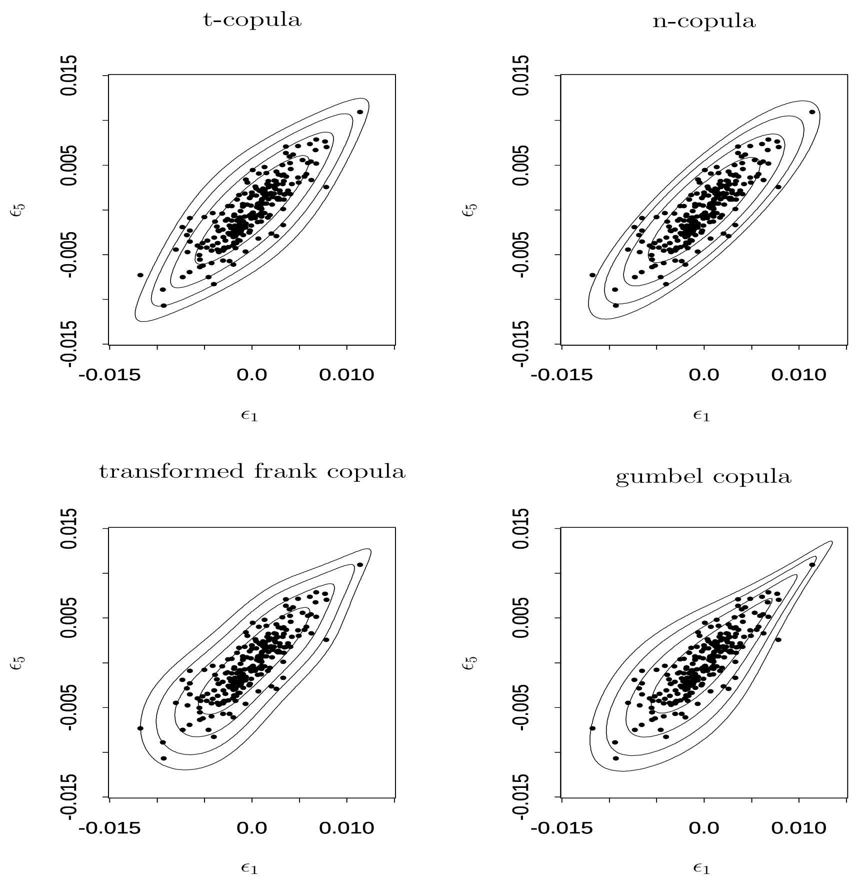

Figure 4: Contourlines of the estimated densities under the different copula assumptions with normal margins. Sample period 1982 to 2001.

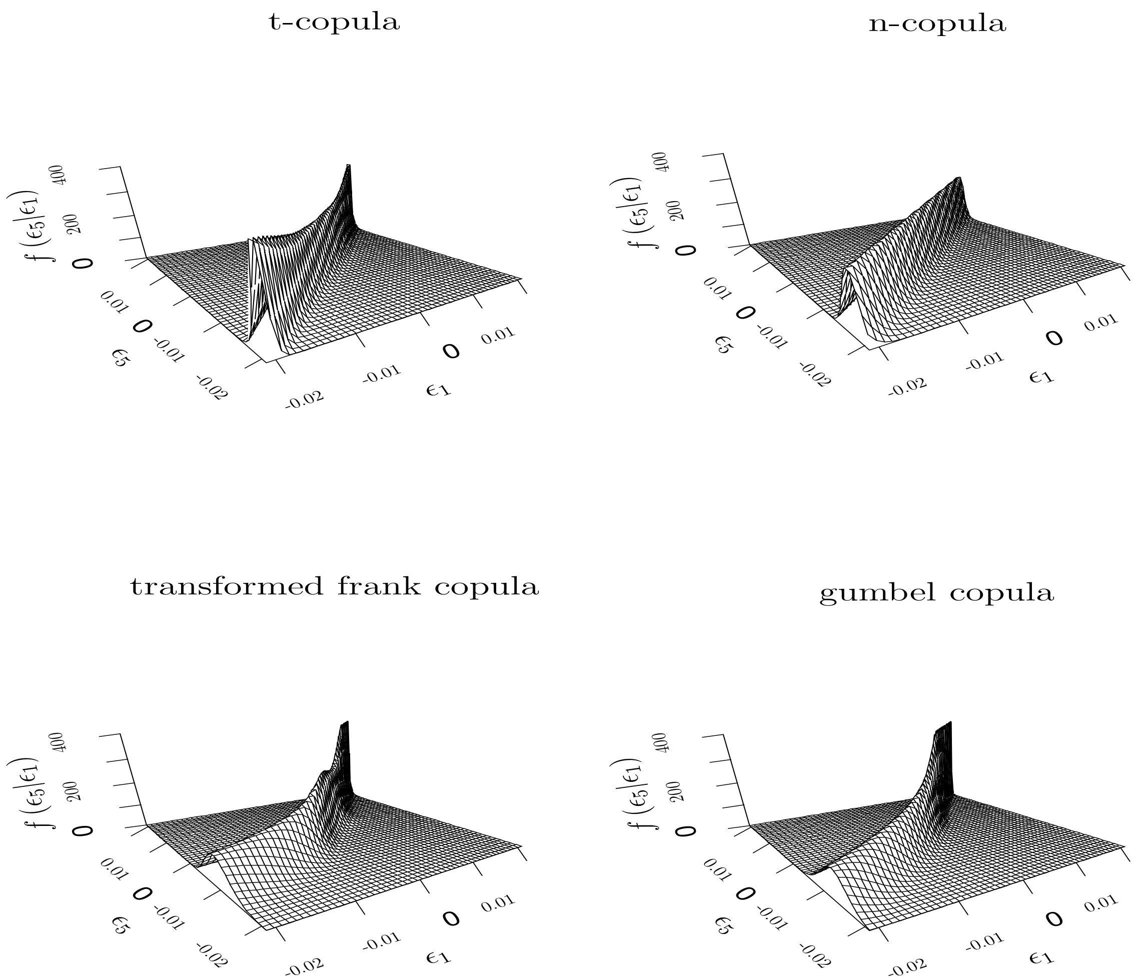

Figure 5: Conditional densities under the different fitted copula functions. Sample period 1982 to 2001.

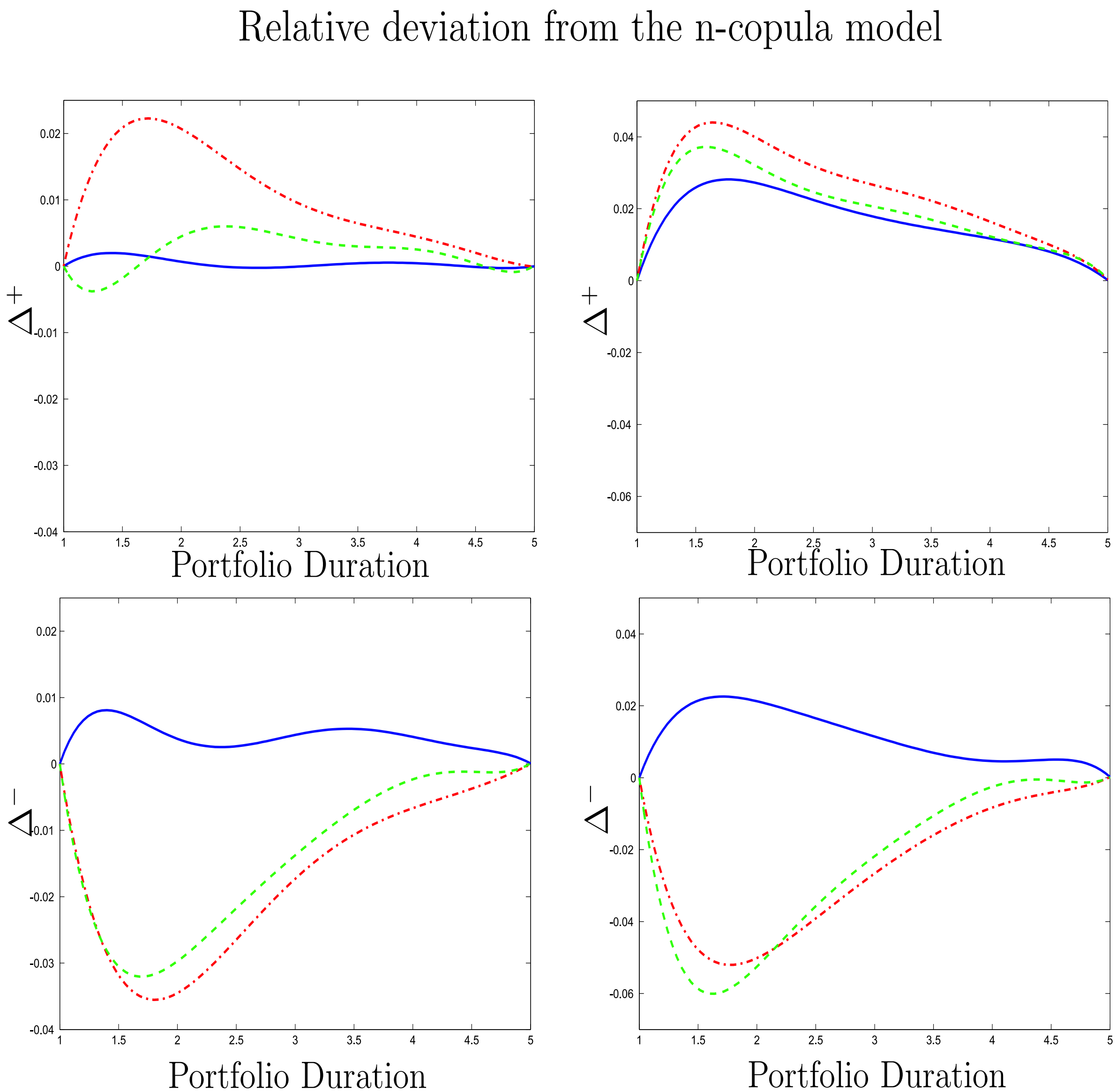

Figure 6: Relative deviations Δ^+ (top) and Δ^- (bottom) of the risk measures from the normal copula model for the $\alpha = 99\%$ (left) and $\alpha = 99.9\%$ (right) quantiles. The solid line belongs to the t-copula, the dashed-dotted line to the Gumbel copula, and the dashed line to the transformed Frank copula. Sample period 1982 to 2001.

Table 1

Summary statistics (sample mean, sample standard deviation, and first order sample autocorrelation) for the monthly U.S. Treasury zero-coupon yield data. Sample period 1982 to 2001 and subsamples 1982 to 1991 and 1992 to 2001.

October 1982 to December 2001			
Maturity	Mean	Std. Dev.	Autocorr.
1 year	0.0638	0.0201	0.9852
2 years	0.0675	0.0206	0.9859
3 years	0.0701	0.0204	0.9854
4 years	0.0721	0.0205	0.9851
5 years	0.0732	0.0205	0.9858

October 1982 to December 1991			
Maturity	Mean	Std. Dev.	Autocorr.
1 year	0.0794	0.0157	0.9641
2 years	0.0837	0.0165	0.9707
3 years	0.0863	0.0165	0.9709
4 years	0.0886	0.0166	0.9696
5 years	0.0897	0.0167	0.9708

January 1992 to December 2001			
Maturity	Mean	Std. Dev.	Autocorr.
1 year	0.0494	0.0109	0.9712
2 years	0.0526	0.0098	0.9581
3 years	0.0550	0.0088	0.9447
4 years	0.0569	0.0082	0.9371
5 years	0.0579	0.0080	0.9387

Table 2

Estimation results for the two-factor generalized Vasicek model with monthly observations on 1, 2, 3, 4 and 5 year maturity yields. Sample period 1982 to 2001 and subsamples 1982 to 1991 and 1992 to 2001. Kalman filter recursions are initialized with the values of the stationary mean and variance of the unobserved state variables. Maximization of the log-likelihood function is based on a sequential quadratic programming algorithm. White (1982) heteroskedasticity-consistent standard errors of the parameter estimates given in parenthesis.

	Kalman Filter Estimates		
Parameters	Oct. 82 to Dec. 01	Oct. 82 to Dec. 91	Jan. 92 to Dec. 01
R_0	0.0589	0.0637	0.0286
	(0.0291)	(0.0254)	(0.0176)
κ_1	0.0691	0.1225	0.1918
	(0.0227)	(0.0285)	(0.0143)
σ_1	0.0203	0.0260	0.2402
	(0.0034)	(0.0048)	(0.0151)
γ_1	-0.1850	-0.0465	-0.4359
	(0.1400)	(0.1696)	(0.1465)
κ_2	0.3719	0.4954	0.2131
	(0.0413)	(0.0610)	(0.0162)
σ_2	0.0188	0.0230	0.2385
	(0.0033)	(0.0047)	(0.0158)
γ_2	1.3358	1.4057	1.5395
	(0.1838)	(0.2509)	(0.2310)
ρ	-0.7807	-0.8199	-0.9991
	(0.0797)	(0.0774)	(0.0001)
σ_{ε_1}	0.0014	0.0008	0.0014
	(0.0001)	(0.0002)	(0.0001)
σ_{ε_2}	0.0004	0.0006	0.0002
	(0.0001)	(0.0001)	(0.0001)
σ_{ε_3}	0.0006	0.0007	0.0003
	(0.0001)	(0.0001)	(0.0001)
σ_{ε_4}	0.0006	0.0008	0.0002
	(0.0001)	(0.0001)	(0.0001)
σ_{ε_5}	0.0005	0.0006	0.0005
	(0.0001)	(0.0001)	(0.0001)

Table 3

Parametric ML-estimates of the joint distribution function G under the alternative copula models. Standard errors and t-values of the parameter estimates given in parenthesis. Sample period 1982 to 2001.

October 1982 to December 2001						
$T = 230$	$\widehat{\sigma}_{1,\cdot}$	$\widehat{\sigma}_{2,\cdot}$	$\widehat{\rho}$	$\widehat{\nu}$	$\widehat{\lambda}_L$	$\widehat{\lambda}_U$
C_t	3.4291E-3	3.4730E-3	0.8556	10.2957	0.3681	0.3681
s.e.	(0.1806E-3)	(0.1701E-3)	(0.0196)	(7.6051)	(0.1716)	(0.1716)
t-value	(18.99)	(20.42)	(43.76)	(1.35)	(2.14)	(2.14)
C_N	3.4517E-3	3.4510E-3	0.8537	-	0	0
s.e.	(0.1845E-3)	(0.1679E-3)	(0.0190)	-	-	-
t-value	(18.71)	(20.55)	(44.97)	-	-	-
	$\widehat{\sigma}_{1,\cdot}$	$\widehat{\sigma}_{2,\cdot}$	$\widehat{\vartheta}$	$\widehat{\delta}$	$\widehat{\lambda}_L$	$\widehat{\lambda}_U$
C_{TF}	3.4152E-3	3.4738E-3	4.1759	1.8101	0	0.5334
s.e.	(0.17785E-3)	(0.1665E-3)	(1.1523)	(0.2463)	-	(0.0764)
t-value	(19.20)	(20.86)	(3.62)	(3.29)	-	(6.98)
C_G	3.5237E-3	3.5433E-3	-	2.8805	0	0.7279
s.e.	(0.18045E-3)	(0.16912E-3)	-	(0.2047)	-	(0.0218)
t-value	(19.53)	(20.95)	-	(9.19)	-	(33.46)

Table 4

Parametric ML-estimates of the joint distribution function G under the alternative copula models. Standard errors and t-values of the parameter estimates given in parenthesis. Sample period 1982 to 1991.

October 1982 to December 1991						
$T = 110$	$\widehat{\sigma}_{1,\cdot}$	$\widehat{\sigma}_{2,\cdot}$	$\widehat{\rho}$	$\widehat{\nu}$	$\widehat{\lambda}_L$	$\widehat{\lambda}_U$
C_t	4.2012E-3	4.0293E-3	0.8871	27.6529	0.2008	0.2008
s.e.	(0.2740E-3)	(0.2684E-3)	(0.0206)	(77.8816)	(0.6241)	(0.6241)
t-value	(15.33)	(15.01)	(43.13)	(0.36)	(0.32)	(0.32)
C_N	4.2051E-3	4.0202E-3	0.8868	-	0	0
s.e.	(0.2746E-3)	(0.2627E-3)	(0.0202)	-	-	-
t-value	(15.31)	(15.31)	(43.79)	-	-	-
	$\widehat{\sigma}_{1,\cdot}$	$\widehat{\sigma}_{2,\cdot}$	$\widehat{\vartheta}$	$\widehat{\delta}$	$\widehat{\lambda}_L$	$\widehat{\lambda}_U$
C_{TF}	4.1678E-3	4.0233E-3	1.9274	2.6545	0	0.7016
s.e.	(0.2619E-3)	(0.2596E-3)	(1.8253)	(0.6329)	-	(0.0808)
t-value	(15.91)	(15.50)	(1.06)	(2.61)	-	(8.68)
C_G	4.2200E-3	4.0745E-3	-	3.2772	0	0.7645
s.e.	(0.2507E-3)	(0.2526E-3)	-	(0.3146)	-	(0.0251)
t-value	(16.83)	(16.13)	-	(7.24)	-	(30.48)

Table 5

Parametric ML-estimates of the joint distribution function G under the alternative copula models. Standard errors and t-values of the parameter estimates given in parenthesis. Sample period 1992 to 2001.

January 1992 to December 2001						
$T = 119$	$\widehat{\sigma}_{1,\cdot}$	$\widehat{\sigma}_{2,\cdot}$	$\widehat{\rho}$	$\widehat{\nu}$	$\widehat{\lambda}_L$	$\widehat{\lambda}_U$
C_t	2.5580E-3	2.9075E-3	0.8148	6.2034	0.4188	0.4188
s.e.	(1.8802E-3)	(0.1980E-3)	(0.0342)	(3.1757)	(0.1185)	(0.1185)
t-value	(13.60)	(14.68)	(23.83)	(1.95)	(3.53)	(3.53)
C_N	2.5634E-3	2.7780E-3	0.7886	-	0	0
s.e.	(0.1864E-3)	(0.1686E-3)	(0.0366)	-	-	-
t-value	(13.75)	(16.48)	(21.52)	-	-	-
	$\widehat{\sigma}_{1,\cdot}$	$\widehat{\sigma}_{2,\cdot}$	$\widehat{\vartheta}$	$\widehat{\delta}$	$\widehat{\lambda}_L$	$\widehat{\lambda}_U$
C_{TF}	2.5364E-3	2.8771E-3	5.8523	1.3362	0	0.3201
s.e.	(0.1847E-3)	(0.1764E-3)	(2.7736)	(0.3693)	-	(0.2408)
t-value	(13.73)	(16.31)	(2.11)	(0.91)	-	(1.33)
C_G	2.5802E-3	2.9629E-3	-	2.4207	0	0.6684
s.e.	(0.1895E-3)	(0.1963E-3)	-	(0.2360)	-	(0.0372)
t-value	(13.62)	(15.09)	-	(6.02)	-	(17.98)

Table 6

Goodness-of-fit statistics for the ML-estimates of the joint distribution function G under the alternative copula models. $^{+}$indicates best model fit for a given statistic, $^{-}$indicates worst model fit for a given statistic. Sample period 1982 to 2001 and subsamples 1982 to 1991 and 1992 to 2001.

October 1982 to December 2001				
$T = 230$	C_t	C_N	C_{TF}	C_G
$\chi^2_{31,32}$ (p-value)	0.34	0.25^{-}	0.90^{+}	0.30
AIC	-4204.30	-4204.24	-4215.34^{+}	-4203.88^{-}
BIC	-4188.54^{-}	-4191.93	-4199.58^{+}	-4191.57
EN	-9.15	-9.15	-9.19^{+}	-9.14^{-}
AD	21.61^{-}	0.69	0.52^{+}	0.63
$PD_{p=0.05}$	$22.58\%^{-}$	20.26%	$6.19\%^{+}$	9.67%
$\overline{PD}_{p=0.05}$	-4.08%	-7.21%	$-3.67\%^{+}$	$21.76\%^{-}$
October 1982 to December 1991				
$T = 110$	C_t	C_N	C_{TF}	C_G
$\chi^2_{11,12}$ (p-value)	0.65	0.73^{+}	0.47	0.37^{-}
AIC	-1954.91^{-}	-1956.80	-1960.44	-1961.27^{+}
BIC	-1942.11^{-}	-1946.70	-1947.64	-1951.17^{+}
EN	-8.92	-8.91^{-}	-8.96^{+}	-8.94
AD	14.6768^{-}	0.6751	0.6134^{+}	0.6489
$PD_{p=0.1}$	$7.25\%^{-}$	6.36%	$-0.91\%^{+}$	-1.86%
$\overline{PD}_{p=0.1}$	0.57%	$0.40\%^{+}$	7.66%	$14.29\%^{-}$
January 1992 to December 2001				
$T = 119$	C_t	C_N	C_{TF}	C_G
$\chi^2_{11,12}$ (p-value)	0.86^{+}	0.78	0.76	0.12^{-}
AIC	-2267.39^{+}	-2255.23	-2265.22	-2254.10^{-}
BIC	-2244.28	-2244.89	-2252.10^{+}	-2243.76^{-}
EN	-9.49	-9.51	-9.56^{+}	-9.47^{-}
AD	6.0714^{-}	0.2756^{+}	0.3732	0.2796
$PD_{p=0.1}$	$0.72\%^{+}$	-9.81%	-7.35%	$-12.81\%^{-}$
$\overline{PD}_{p=0.1}$	4.35%	-5.68%	$0.41\%^{+}$	$16.22\%^{-}$