Title
Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula

Author
Patton, Andrew J

Publication Date
2001-06-01

2001-09

UNIVERSITY OF CALIFORNIA, SAN DIEGO

DEPARTMENT OF ECONOMICS

MODELLING TIME-VARYING EXCHANGE RATE DEPENDENCE USING THE CONDITIONAL COPULA

BY

ANDREW J. PATTON

DISCUSSION PAPER 2001-09
JUNE 2001

Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula

Andrew J. Patton*
University of California, San Diego

First Draft: December 2000
This Draft: June 2001

*Contact information: Department of Economics, University of California, San Diego, 9500 Gilman Drive, La Jolla CA 92093-0508, U.S.A. Email: apatton@econ.ucsd.edu. The author would like to thank Graham Elliott, Rob Engle, Marjorie Flavin, Raffaella Giacomini, Tony Hall, Joshua Rosenberg, Kevin Sheppard, Alan Sorensen, Ross Starr, Allan Timmermann, and seminar participants at the University of Technology, Sydney School of Finance and Economics, Monash University Department of Econometrics and Business Statistics, the UCSD Econometrics seminar and the Commonwealth Scientific and Industrial Research Organisation Computing and Mathematical Sciences Division for their comments and suggestions. Financial support from the UCSD Project in Econometric Analysis Fellowship is gratefully acknowledged. This paper may be downloaded from http://www.econ.ucsd.edu/~apatton/research.html.

.

Modelling Time-Varying Exchange Rate Dependence

Using the Conditional Copula

Abstract

Linear correlation is only an adequate means of describing the dependence between two random variables when they are jointly elliptically distributed. When the joint distribution of two or more assets is not elliptical the linear correlation coefficient becomes just one of many possible ways of summarising the dependence structure between the variables. In this paper we make use of a theorem due to Sklar (1959), which shows that an n-dimensional joint distribution function may be decomposed into its n marginal distributions, and a *copula*, which completely describes the dependence between the n variables. We verify that Sklar's theorem may be extended to conditional distributions, and apply it to the modelling of the time-varying joint distribution of the Deutsche mark - U.S. dollar and Yen - U.S. dollar exchange rate returns. We find evidence that the conditional dependence between these exchange rates is time-varying, and that it is asymmetric: dependence is higher during appreciations of the U.S. dollar against the mark and the yen than during depreciations of the dollar. We also find strong evidence of a structural break in the conditional copula following the introduction of the euro.

KEYWORDS: time series, copulas, exchange rates, dependence.
J.E.L. Codes: C32, C51, C52, F31.

1 Introduction

The concept of dependence is of central interest to all economists, practitioners and academics alike. Linear correlation, the most widely used and quoted measure of dependence, is just one simple means of summarising the dependence between two random variables. For variables that have an elliptical joint distribution linear correlation is sufficient to describe their dependence, however if their joint distribution is not elliptical, then linear correlation becomes just one of many possible ways of summarising the dependence structure. In this paper we make use of a theorem due to Sklar (1959), which shows that an n-dimensional joint distribution function may be decomposed into its n marginal distributions, and a *copula*, which completely describes the dependence between the n variables[1].

By using an extension of Sklar's theorem we are able to exploit the success we have had in the modelling of univariate distributions by first specifying models for the marginal distributions of a multivariate distribution of interest, and then specifying a copula. As an example, let us consider the modelling of the joint distribution of two exchange rates: the Student's t distribution has been found to provide a reasonable fit to the conditional univariate distribution of daily exchange rate returns, see Bollerslev (1987) for example. A natural starting point in the modelling of the joint distribution of two exchange rates might then be a bivariate t distribution. However, the bivariate Student's t distribution has the restrictive property that both marginal distributions have the same degrees of freedom parameter. Studies such as Bollerslev (1987) have shown that different exchange rates have different degrees of freedom parameters. In Section 4.3 below we show that even the very flexible BEKK model for the conditional variance-covariance estimated assuming a bivariate t density for the standardised residuals fails goodness-of-fit tests of the specified density when estimated on the Deutsche mark - U.S. dollar and the Yen - U.S. dollar exchange rates. The condition that both exchange rate returns have the same degrees of freedom parameter is simply too restrictive. Note also that this is possibly the most ideal situation: where both assets have univariate distributions from the same family, the Student's t, and very similar degrees of freedom, 5.8 for the mark and 4.4 for the yen. We could imagine situations where the two variables of interest have quite different marginal distributions: a stock return and an exchange rate, for example, where no obvious choice for the bivariate density exists. Decomposing the multivariate distribution into the marginal distributions and the copula allows for the construction of better models of the individual variables than would be possible if we constrained ourselves to look only at existing multivariate distributions.

Despite the fact that copulas were introduced as a means of isolating the dependence structure

[1]The word *copula* comes from Latin for a 'link' or 'bond', and was coined by Sklar (1959), who first proved the theorem that a collection of marginal distributions can be 'coupled' together via a copula to form a multivariate distribution. It has been given various names, such as *dependence function* (Galambos, 1978 and Deheuvels, 1978), *uniform representation* (Kimeldorf and Sampson, 1975, and Hutchinson and Lai, 1990) or *standard form* (Cook and Johnson, 1981).

of a multivariate distribution over forty years ago, they have not been widely used in econometrics. One possible reason for this is that, in general, no analytical expressions exist for the parameters of a particular copula, and so maximum likelihood, or some other estimation method, is usually required. A further reason may lie in the fact that interest in modelling (and forecasting) the entire density of a random variable has only risen in the last few years. With the computational resources now available, and the growing interest in density and quantile models, the time seems ripe for copula theory to be applied to economic problems.

The aim of this paper is two-fold. Firstly, we show in Section 2 that the existing theory of (unconditional) copulas may be extended to the conditional case, thus allowing us to use copula theory in the modelling of time-varying conditional dependence. Time variation in the conditional first and second moments of economic time series has been widely reported, and so allowing for time variation in the conditional dependence between economic time series seems natural. The second aim of this paper is to show how we may apply the theory of conditional copulas in the modelling of time-varying exchange rate dependence. We examine daily Deutsche mark - U.S. dollar (DM-USD) and Japanese yen - U.S. dollar (Yen-USD) exchange rates over the period January 1991 to October 2000. The modelling of the entire conditional joint distribution of these exchange rates, rather than just, say, the conditional means, variances and linear correlation, has a number of attractive features: the first is that given the conditional joint distribution we can, of course, obtain the conditional means, variances and correlation, so this type of modelling nests solely modelling conditional moments. Secondly, we can obtain the time-paths of *any* other dependence measure of interest, such as rank correlation, which can capture non-linear dependence, or measures of dependence in the extremes, such as tail dependence[2]. Dependence during extreme events has been the subject of much recent analysis in the financial contagion literature. Further, there are economic situations where the entire conditional joint density is required, such as the pricing of financial options with multiple underlying assets, see Rosenberg (2000), or in the calculation of portfolio Value-at-Risk (VaR), see Hull and White (1998).

In our empirical application we find that the dependence between the DM-USD and Yen-USD exchange rates is asymmetric, in that they are more dependent during appreciations of the U.S. dollar (or alternatively, during depreciations of the mark and the yen) than they are during depreciations of the U.S. dollar. We also report evidence that the conditional dependence between these exchange rates is time-varying, though formally testing this is complicated by the presence of a nuisance parameter. Finally, we find strong evidence of a structural break in the conditional copula following the introduction of the euro in January 1999, with these two exchange rates exhibiting much weaker dependence after the break than before.

The structure of the remainder of this paper is as follows. Section 2 introduces the theory of the conditional copula, and Section 3 discusses some of the issues regarding the evaluation and

[2] This measure will be discussed in more detail in Section 4.

comparison of copula models. In Section 4 we apply the theory of conditional copulas to the modelling of the time-varying joint distribution of the Deutsche mark - U.S. dollar and Yen - U.S. dollar exchange rates. In that section we discuss how we allow for time variation in the conditional dependence, and how the competing copulas compare. In Section 5 we summarise our results, and discuss other potential applications of conditional copulas in economics. Finally, we provide the proofs of any theorems in Appendix A.

2 The Theory of the Conditional Copula

In this paper we will focus on bivariate distributions, but it should be noted that the theory of copulas is applicable to the more general multivariate case. The two random variables considered will be denoted X and Y, with distributions F and G respectively. Their joint distribution will be denoted H. We will assume in this paper that the marginal distribution functions, F and G, are continuous. The assumption of continuity is not required, but simplifies some of the presentation. Throughout this paper we will denote the distribution (or *c.d.f.*) of a random variable using an upper case letter, and the corresponding density (or *p.d.f.*) using the lower case letter. Also note that we will denote the extended real line as $\bar{\mathbb{R}} \equiv \mathbb{R} \cup \{\pm\infty\}$.

In Section 2.1 below, we introduce the copula via standard theory on the distribution of random variables. Following that, the more general theory of conditional copulas is presented.

2.1 The Copula and Transformations of Random Variables

Consider two random variables U and V, each a particular transformation of X and Y: let $U = F(X)$ and $V = G(Y)$. That is, U and V are the *probability integral transforms* of X and Y respectively. (The probability integral transform will be discussed below.) We will now attempt to find the joint density of U and V according to basic results in mathematical statistics on the distribution of transformations of random variables. One standard reference for this is Casella and Berger (1990). We will denote the joint density of U and V as c, which turns out to be the 'copula density'.

Since F and G are strictly increasing and continuous, we have that $X = F^{-1}(U)$ and $Y = G^{-1}(V)$, and $\frac{\partial X}{\partial U} = \left(\frac{\partial U}{\partial X}\right)^{-1} = \left(\frac{\partial F(X)}{\partial X}\right)^{-1} = f(X)^{-1}$ and $\frac{\partial Y}{\partial V} = \left(\frac{\partial V}{\partial Y}\right)^{-1} = \left(\frac{\partial G(Y)}{\partial Y}\right)^{-1} = g(Y)^{-1}$. Note also that $\frac{\partial X}{\partial V} = \frac{\partial Y}{\partial U} = 0$. Then,

$$\begin{aligned} c(u,v) &= h(X(u), Y(v)) \cdot \begin{vmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{vmatrix} \\ &= h\left(F^{-1}(u), G^{-1}(v)\right) \cdot \frac{\partial X}{\partial U} \cdot \frac{\partial Y}{\partial V} \\ c(u,v) &= \frac{h\left(F^{-1}(u), G^{-1}(v)\right)}{f\left(F^{-1}(u)\right) \cdot g\left(G^{-1}(v)\right)} \qquad (1) \end{aligned}$$

Equation (1) shows that the copula density of X and Y is equal to the ratio of the joint density,

h, to the product of the marginal densities, f and g. From this expression we can obtain a first result on the properties of copulas: if X and Y are independent, then the copula density takes the value 1 everywhere, since in that case the joint density is *equal to* the product of the marginal densities. We can also use the above equation to derive an expression for h as a function of x and y instead:

$$\begin{aligned} h\left(F^{-1}(u), G^{-1}(v)\right) &= f\left(F^{-1}(u)\right) \cdot g\left(G^{-1}(v)\right) \cdot c(u,v) \\ h(x,y) &= f(x) \cdot g(y) \cdot c(F(x), G(y)) \end{aligned} \tag{2}$$

Equation (2) is the 'density version' of Sklar's (1959) theorem: the joint density, h, can be decomposed into product of the marginal densities, f and g, and the copula density, c. Sklar's theorem holds under more general conditions than the ones we imposed for this illustration, and below we discuss the general proof.

2.2 The Theory of the Conditional Copula

For an introduction to the general theory of copulas the reader is referred to Nelsen (1999) or Chapter 6 of Schweizer and Sklar (1983). We will start with a few very basic, but very important, definitions based on those in Nelsen (1999). The seond condition below refers to the 'H-volume' of a rectangle $[x_1, x_2] \times [y_1, y_2]$ in $\bar{\mathbb{R}}^2$, denoted by V_H. This is simply the probability of observing a point in the region $[x_1, x_2] \times [y_1, y_2]$. It is expressed in the following way as it generalises more easily to the multivariate case.

Definition 1 *A conditional bivariate distribution function is a right continuous function* $H : \bar{R}^2 \to [0,1]$ *with the properties:*

1. $H(x, -\infty|\mathcal{F}) = H(-\infty, y|\mathcal{F}) = 0$, *and* $H(\infty, \infty|\mathcal{F}) = 1$
2. $V_H([x_1, x_2] \times [y_1, y_2]) \equiv H(x_2, y_2|\mathcal{F}) - H(x_1, y_2|\mathcal{F}) - H(x_2, y_1|\mathcal{F}) + H(x_1, y_1|\mathcal{F}) \geq 0$ *for all* $x_1, x_2, y_1, y_2 \in \bar{\mathbb{R}}$, *and* $x_1 \leq x_2$, $y_1 \leq y_2$.

   where $\mathcal{F}$ *is some conditioning set.*

The first condition simply provides the upper and lower bounds on the distribution function. The second condition ensures that the probability of observing a point in the region $[x_1, x_2] \times [y_1, y_2]$ is non-negative[3]. The conditional marginal distributions of X and Y are defined as $F(x|\mathcal{F}) \equiv H(x, \infty|\mathcal{F})$, and $G(y|\mathcal{F}) \equiv H(\infty, y|\mathcal{F})$. We now define the focus of this paper; the conditional copula.

[3] If we set $x_2 = x_1 + \varepsilon$ and $y_2 = y_1 + \varepsilon$ and let $\varepsilon \to 0^+$, then it becomes clear that this definition is just the generalisation of the condition that if the bivariate density exists, it must be non-negative on the domain of H.

Definition 2 *A two-dimensional conditional copula is a function* $C : [0,1] \times [0,1] \to [0,1]$ *with the following properties:*

1. $C(u,0|\mathcal{F}) = C(0,v|\mathcal{F}) = 0$*, and* $C(u,1|\mathcal{F}) = u$ *and* $C(1,v|\mathcal{F}) = v$*, for every* u, v *in* $[0,1]$
2. $V_C([u_1,u_2] \times [v_1,v_2] \,|\mathcal{F}) \equiv C(u_2,v_2|\mathcal{F}) - C(u_1,v_2|\mathcal{F}) - C(u_2,v_1|\mathcal{F}) + C(u_1,v_1|\mathcal{F}) \geq 0$ *for all* $u_1, u_2, v_1, v_2 \in [0,1]$*, such that* $u_1 \leq u_2$ *and* $v_1 \leq v_2$*.*

   where $\mathcal{F}$ *is some conditioning set.*

The first condition of Definition 2 provides the lower bound on the distribution function, and ensures that the marginal distributions, $C(u,1|\mathcal{F})$ and $C(1,v|\mathcal{F})$, are uniform. The condition that V_C is non-negative has the same interpretation as the second condition of Definition 1: it simply ensures that the probability of observing a point in the region $[u_1,u_2] \times [v_1,v_2]$ is non-negative.

By drawing on the above conditions for the conditional copula, and extending its domain to $\bar{\mathbb{R}}^2$, we may alternatively define a conditional copula as the conditional bivariate distribution of a pair of random variables (U,V) having margins that are $Unif\,(0,1)$. The extension of the domain to $\bar{\mathbb{R}}^2$ is accomplished as follows:

$$\text{Let} \quad C^*(u,v|\mathcal{F}) = \left\{ \begin{array}{cl} 0 & \text{for } u < 0 \text{ or } v < 0, \\ C(u,v|\mathcal{F}) & \text{for } (u,v) \in [0,1] \times [0,1]\,, \\ u & \text{for } u \in [0,1]\,, v > 1, \\ v & \text{for } u > 1, v \in [0,1]\,, \\ 1 & \text{for } u > 1, v > 1. \end{array} \right\}$$

This alternative definition of the conditional copula as the conditional bivariate distribution of a pair of random variables (U,V) having margins that are $Unif\,(0,1)$ becomes even more intuitive when we consider a transformation known as the *probability integral transformation.* The random variable $U_t = F(X_t|\mathcal{F})$ is the probability integral transform of X_t, and is known to have the $Unif\,(0,1)$ distribution, regardless of the original distribution, F. This result was first introduced by Fisher (1932), see Casella and Berger (1990) for more details[4].

The link between this transformation and the theory of copulas now becomes clear: the copula is the joint distribution function of the probability integral transforms of each of the variables X and Y with respect to their marginal distributions, F and G. We now move on to an extension of the the key result in the theory of copulas: Sklar's (1959) theorem for conditional distributions:

[4]The probability integral transform has also been use in the context of goodness-of-fit tests as far back as the 1930s, see K. Pearson (1933) for example. More recently Diebold, *et al.* (1998) extended the probability integral transform theory to the time series case, and proposed using it in the evaluation of density forecasts. We will discuss this further in Section 3 below.

Theorem 3 (Sklar's Theorem for Continuous Conditional Distributions) *Let H be a conditional bivariate distribution function with continuous margins F and G, and let $\mathcal{F}$ be some conditioning set. Then there exists a unique conditional copula $C : [0,1] \times [0,1] \to [0,1]$ such that*

$$H(x,y|\mathcal{F}) = C(F(x|\mathcal{F}), G(y|\mathcal{F})|\mathcal{F}), \quad \forall x, y \in \bar{\mathbb{R}} \tag{3}$$

Conversely, if C is a conditional copula and F and G are the conditional distribution functions of two random variables X and Y, then the function H defined by equation (3) is a bivariate conditional distribution function with margins F and G.

The density function equivalent of (3) is useful for maximum likelihood analysis, and is obtained quite easily, provided that F and G are differentiable, and H and C are twice differentiable.

$$\begin{aligned}
h(x,y|\mathcal{F}) &\equiv \frac{\partial^2 H(x,y|\mathcal{F})}{\partial x \partial y} \\
&= \frac{\partial^2 C(F(x|\mathcal{F}), G(y|\mathcal{F})|\mathcal{F})}{\partial (F(x|\mathcal{F})\, \partial (G(y|\mathcal{F}))} \cdot \frac{\partial F(x|\mathcal{F})}{\partial x} \cdot \frac{\partial G(y|\mathcal{F})}{\partial y} \\
&= \frac{\partial^2 C(u,v|\mathcal{F})}{\partial u \partial v} \cdot f(x|\mathcal{F}) \cdot g(y|\mathcal{F}) \\
h(x,y|\mathcal{F}) &= c(u,v|\mathcal{F}) \cdot f(x|\mathcal{F}) \cdot g(y|\mathcal{F}), \quad \forall x, y \in \bar{\mathbb{R}}
\end{aligned} \tag{4}$$

where $u \equiv F(x|\mathcal{F})$, and $v \equiv G(y|\mathcal{F})$.

We can also obtain a corollary to Theorem 3, analogous to that of Nelson's (1999) corollary to Sklar's Theorem, which enables us to extract the conditional copula from any conditional bivariate distribution function, but first we need the definition of the 'quasi-inverse' of a function.

Definition 4 *The quasi-inverse, $F^{(-1)}$, of a distribution function F is defined as:*

$$F^{(-1)}(u) = \inf\{x : F(x) \geq u\}, \text{ for } u \in [0,1]. \tag{5}$$

If F is strictly increasing then the above definition returns the usual functional inverse of F, but more importantly it allows us to consider inverses of non-strictly increasing functions.

Corollary 5 *Let H be any conditional bivariate distribution with continuous marginal distributions, F and G, and let $F^{(-1)}$ and $G^{(-1)}$ denote the (quasi-) inverses of the marginal distributions. Finally, let $\mathcal{F}$ be some conditioning set. Then there exists a unique conditional copula $C : [0,1] \times [0,1] \to [0,1]$ such that*

$$C(u,v|\mathcal{F}) = H\left(F^{(-1)}(u|\mathcal{F}), G^{(-1)}(v|\mathcal{F})|\mathcal{F}\right), \quad \forall u, v \in [0,1] \tag{6}$$

This corollary completes the idea that a bivariate distribution function may be decomposed into three parts. Given any two marginal distributions and any copula we have a joint distribution, and from any given joint distribution we can extract the implied marginal distributions and copula.

To provide some idea as to the flexibility that the above framework gives us, we now consider various joint distributions, all with standard normal marginal distributions and all implying a linear correlation coefficient, ρ, of 0.5. The contour plots of these distributions are presented in Figure 1. In the upper left corner of this figure is the standard bivariate normal distribution with $\rho = 0.5$. The other elements of this figure show the dependence structures implied by other copulas, with each copula calibrated so as to also yield $\rho = 0.5$. It is quite clear that knowing the marginal distributions and linear correlation is not sufficient to describe a joint distribution: Clayton's copula, for example, has contours that are quite peaked in the joint lower tail, implying greater dependence there than in the joint upper tail. Gumbel's copula implies just the opposite. The Joe-Clayton copula, which we will discuss in more detail below, is slightly peaked in both joint tails, though more so in the upper than the lower. The functional forms of the copulas presented in Figure 1 may be found in Joe (1997).

3 Evaluation of conditional density models

Before moving on to developing models for the conditional copula, we must first establish a means of evaluating their goodness-of-fit. Measures of goodness-of-fit are not only of importance for evaluating the fit of a proposed copula, but for testing the specification of the marginal distributions. Modelling of the conditional copula requires that the models for the marginal distributions be indistinguishable from the true marginal distributions.

As discussed above, a copula may be viewed as the joint distribution of two uniform random variables, thus the evaluation of copula models is a special case of the more general problem of evaluating (multivariate) density models. The density model (or forecast) evaluation literature is relatively young, and no single method has emerged as best. Studies by Diebold, *et al.* (1998) and Diebold, *et al.* (1999) focus on the probability integral transforms of the data in the evaluation of the density model, and so are clearly relevant in evaluating copula models.

As mentioned in footnote 4, the probability integral transform has been used in goodness-of-fit testing as far back as K. Pearson (1933), and since then in Neyman (1937), E. S. Pearson (1938), Dawid (1984), Kling and Bessler (1989) and Diebold, *et al.* (1998). Diebold, *et al.* (1998) showed that for the time series framework the sequence of probability integral transforms will be *i.i.d.* $Unif\,(0,1)$ if the sequence of densities is correct, and proposed testing the specification of a density model by testing whether or not the transformed series was *i.i.d.*, and $Unif\,(0,1)$ in two separate stages[5,6]. Let us denote the two transformed series as $\{u_t\}_{t=1}^T$ and $\{v_t\}_{t=1}^T$, where $u_t \equiv F_t(x_t|\mathcal{F}_{t-1})$ and $v_t \equiv G_t(y_t|\mathcal{F}_{t-1})$, for $t = 1, 2, ..., T$ and $\mathcal{F}_t = \sigma(x_t, y_t, x_{t-1}, y_{t-1}, ...)$. Diebold, *et al.* (1998) propose firstly testing the independence of the first four moments of U_t and V_t, by regressing $(u_t - \bar{u})^k$ and $(v_t - \bar{v})^k$ on 20 lags of both $(u_t - \bar{u})^k$ and $(v_t - \bar{v})^k$, for $k = 1, 2, 3, 4$. Under the null that both $\{u_t\}_{t=1}^T$ and $\{v_t\}_{t=1}^T$ represent $i.i.d.$ samples all coefficients in these regressions should be zero. If the two series $\{u_t\}_{t=1}^T$ and $\{v_t\}_{t=1}^T$ pass the tests for serial dependence, then we may test the hypothesis that the transformed series are $Unif(0,1)$ via the Kolmogorov-Smirnov (K-S) test[7]. Diebold, *et al.* (1999) extend the results of Diebold, *et al.* (1998) to the evaluation of multivariate density models/forecasts. They propose testing the 'conditional' probability integral transform: a bivariate distribution may be decomposed into a conditional and a marginal distribution, $H_t(x_t, y_t|\mathcal{F}_{t-1}) = H_{Y,t}(y_t|x_t, \mathcal{F}_{t-1}) \cdot F(x_t|\mathcal{F}_{t-1})$, and $H_t(x_t, y_t|\mathcal{F}_{t-1}) = H_{X,t}(x_t|y_t, \mathcal{F}_{t-1}) \cdot G(y_t|\mathcal{F}_{t-1})$, where $H^i_{X,t}(x_t|y_t, \mathcal{F}_{t-1})$ and $H^i_{Y,t}(y_t|x_t, \mathcal{F}_{t-1})$ are the conditional $c.d.f.$s of X and Y respectively[8]. The variables U_t and V_t are defined in the same way as above, and two new variables are defined: $U^c_t \equiv H_{X,t}(X_t|Y_t, \mathcal{F}_{t-1})$ and $V^c_t \equiv H_{Y,t}(Y_t|X_t, \mathcal{F}_{t-1})$. The method of Diebold, *et al.* (1999) involves checking that each of $\{u_t\}_{t=1}^T$, $\{v_t\}_{t=1}^T$, $\{u^c_t\}_{t=1}^T$ and $\{v^c_t\}_{t=1}^T$ are distributed as $i.i.d.$ $Unif(0,1)$ random variables, using the same tests as for the univariate case. If this holds, then all marginal and conditional distributions of the bivariate distribution are correctly specified, implying that the bivariate distribution is correctly specified.

There are two drawbacks of the above approach to evaluating a density model: the main drawback is that we must test the correctness of the density model separately from testing for serial dependence in the transformed variables[9]. The second drawback is that the fact that the

[5]Ideally we would like to test the *joint* hypothesis, however no such test is currently available, and so two separate tests are used.

[6]It should be noted that these tests were developed for the case when the parameters of the proposed model are *known*, and not estimated from the sample. Constructing the variables u_t and v_t using parameter estimates is not innocuous. Indeed, it was known as far back as David and Johnson (1948) that when the probability integral transform is taken with respect to the correct distribution but using estimated parameters the resulting random variable does *not* have the $Unif(0,1)$ distribution; instead it has a distribution that depends on the distribution of the original (untransformed) random variable. The implications for these specification tests are that we need, as some authors in the past have done, see Engle and Manganelli (1999) and Diebold *et al.* (1998), to interpret the tests as being *conditional* on the estimated parameters. These tests, then, ignore any estimation error in the parameters. The best we can hope for is that for large sample sizes the magnitude of the estimation uncertainty is small. Some of the implications of marginal parameter estimation uncertainty for the application of the theory of copulas to economics are addressed in Patton (2001).

[7]See Shao (1999) for the theory underlying this test. In the implementation of this test we first sort the sequences of transformed variables, $\{u_t\}_{t=1}^T$ and $\{v_t\}_{t=1}^T$, into ascending order, denoted by $\{z^u_t\}_{t=1}^T$ and $\{z^v_t\}_{t=1}^T$. The K-S test statistics are then calculated as $D^u_T = \max_t \left|\frac{t}{T} - z^u_t\right|$ and $D^v_T = \max_t \left|\frac{t}{T} - z^v_t\right|$. We may employ a numerical approximation due to Press, *et al.* (1989) to obtain the p-values corresponding to the test statistics.

[8]The conditional $c.d.f.$ s of X and Y are given by $H^i_{X,t}(x_t|y_t, \mathcal{F}_{t-1}) = \frac{\partial H^i_t(x_t, y_t|\mathcal{F}_{t-1})}{\partial y}$ and $H^i_{Y,t}(y_t|x_t, \mathcal{F}_{t-1}) = \frac{\partial H^i_t(x_t, y_t|\mathcal{F}_{t-1})}{\partial x}$.

[9]Berkowitz (1999) proposed one solution to this problem. He suggested that instead of testing the $\{u_t\}_{t=1}^T$ series,

Kolmogorov-Smirnov test has lower power in the tails of the distribution than in the centre, see Stephens (1986), and this is a critical region for the construction of Value-at-Risk estimates for a portfolio assets, for example. We propose here an alternative test, which draws on the interval forecasting literature and quantifies the intuition that Diebold, *et al.* (1998) suggest can be gained by looking at the empirical histograms of the transformed data. Diebold, *et al.* suggest that by comparing the number of observations in each bin, otherwise known as a 'hit' in that bin, with what would be expected under the null hypothesis we may gain some insight as to where the model fails, if at all. For example, too many observations in the bins near zero or one would suggest that the density model has tails that are too thin. This form of evaluation obviously has its roots in K. Pearson's (1900) χ^2 test, see D'Agostino and Stephens (1986) for more details.

In the following test we decompose the density model into a set of 'region' models ('interval' models in the univariate case), each of which should be correctly specified under the null hypothesis that the entire density is correctly specified. The specification introduced below is an extension of the 'hit' regressions of Christoffersen (1998) and Engle and Manganelli (1999), proposed to evaluate interval forecasts, such as Value-at-Risk forecasts. We will describe the test below in a general setting, and discuss the details of implementation in Section 4.5.

Let W_t be the (possibly multivariate) random variable under analysis, and define the support of W_t as $\mathcal{S}$. Let $\{R_j\}_{j=0}^{K}$ be regions in $\mathcal{S}$ such that $R_i \cap R_j = \emptyset$ if $i \neq j$, and $\cup_{j=0}^{K} R_j = \mathcal{S}$. Let π_{jt} be the true probability that $W_t \in R_j$ and let p_{jt} be the probability suggested by the model[10]. Finally, let $\Pi_t \equiv [\pi_{0t}, \pi_{1t}, ..., \pi_{Kt}]'$ and $P_t \equiv [p_{0t}, p_{1t}, ..., p_{Kt}]'$. Under the null hypothesis that the model is correctly specified we have that $P_t = \Pi_t$ for $t = 1, 2, ..., T$. Let us define the variables to be analysed in the tests as $Hit_t^j \equiv \mathbf{1}\{X_t \in R_j\}$, where $\mathbf{1}\{A\}$ takes the value 1 if the argument, A, is true and zero elsewhere, and $M_t \equiv \sum_{j=0}^{K} j \cdot \mathbf{1}\{X_t \in R_j\}$.

We may test that the model is adequately specified in each of the $K+1$ regions individually via tests of the hypothesis $H_0 : Hit_t^j \sim i.n.i.d.Bernoulli\,(p_{jt})$ versus $H_1 : Hit_t^j \sim i.n.i.d.Bernoulli\,(\pi_{jt})$, where π_{jt} is a function of both p_{jt}, and other elements of the time $t-1$ information set thought to possibly have explanatory power for the probability of a hit. This is where our test differs from those presented in Christoffersen (1998) and Engle and Manganelli (1999): the former proposed modelling π_{jt} as a first-order Markov chain to check for first-order serial dependence of the hits, while the latter proposed using a linear probability model to determine if other variables, such as lagged hits and also lagged levels of the Value-at-Risk, had significant predictive value. The

say, we may define a new series: $\left\{z_t \equiv \Phi^{-1}(u_t)\right\}_{t=1}^{T}$, where Φ^{-1} is the inverse cdf of a standard normal distribution. The null hypothesis that $\{u_t\}_{t=1}^{T}$ is *i.i.d.* $Unif\,(0,1)$ may be tested by testing that $\{z_t\}_{t=1}^{T}$ is *i.i.d.* $N(0,1)$, which is possibly easier due to the large number of tests of normality available.

[10] Given the similarity between this test and Pearson's χ^2 test it would not be surprising to find that the power of the test is maximised when the probability mass in each region is equal. For a univariate density model this is a simple task, however it may be a more difficult task in the more general multivariate case. Also, it may be that the researcher has a particular interest in certain regions of the support (the lower 10% square, representing the 10% Value-at-Risk, for example) being correctly specified. For these reasons we consider the case where the probability mass in each region is possibly unequal.

markov chain approach suffers from the drawback that it is difficult to check for the influence of other variables or longer lags, while Engle and Manganelli's (1999) model may be improved relatively easily by using a better model for the hits than a linear probability model, which assumes normally distributed errors. We propose using a logit model for the hits, which yields more efficient parameter estimates, and thus hopefully a more powerful test[11]. For more details on the logit model, see Davidson and MacKinnon (1993) or Greene (1997). The model we propose for π_{jt} is:

$$\pi_{jt} = \pi_j\left(Z_{jt}, \beta_j, p_{jt}\right) = \Lambda\left(\lambda_j\left(Z_{jt}, \beta_j\right) - \ln\left[\frac{1-p_{jt}}{p_{jt}}\right]\right) \tag{7}$$

where $\Lambda(x) \equiv \frac{1}{1+e^{-x}}$ is the logistic transformation, Z_{jt} is a matrix containing variables thought to influence the probability of a hit, β_j is a $(k_j \times 1)$ vector of parameters to be estimated, and λ_j is any function of regressors and parameters such that $\lambda_j(Z,0) = 0$ for all Z. The condition on λ_j is imposed so that when $\beta_j = 0$ we have that $\pi_{jt} = \pi_j(Z_{jt}, 0, p_{jt}) = p_{jt}$, and thus the competing hypotheses may be expressed as $\beta_j = 0$ versus $\beta_j \neq 0$. The parameter β_j may be found via maximum likelihood, where the likelihood function to be maximised is: $\mathcal{L}\left(\pi_j\left(Z_j, \beta_j, p_j\right)|Hit^j\right) = \Sigma_{t=1}^{T} Hit_t^j \cdot \ln \pi_j\left(Z_{jt}, \beta_j, p_{jt}\right) + \left(1 - Hit_t^j\right) \cdot \ln\left(1 - \pi_j\left(Z_{jt}, \beta_j, p_{jt}\right)\right)$. The test may then quite easily be conducted as a likelihood ratio test, where $LR_j \equiv -2 \cdot \left(\mathcal{L}\left(p_j|Hit^j\right) - \mathcal{L}\left(\pi_j\left(Z_j, \hat{\beta}_j, p_j\right)|Hit^j\right)\right)$ $\sim \chi^2_{k_j}$ under the null hypothesis that the model is correctly specified in region R_j.

We may test whether the proposed density model is correctly specified in all $K+1$ regions simultaneously by testing the hypothesis $H_0 : M_t \sim Multinomial\left(P_t\right)$ versus $H_1 : M_t \sim Multinomial\left(\Pi_t\right)$, where again we specify Π_t to be a function of both P_t and variables in the time $t-1$ information set thought to possibly influence the probability of a hit in one of the regions. We propose the following setup for the elements of Π_t:

$$\pi_t^1 = \pi^1\left(Z_t, \boldsymbol{\beta}, P_t\right) = \Lambda\left(\lambda_1\left(Z_{1t}, \beta_1\right) - \ln\left[\frac{1-p_{1t}}{p_{1t}}\right]\right) \tag{8}$$

$$\begin{aligned}\pi_t^j &= \pi^j\left(Z_t, \boldsymbol{\beta}, P_t\right) \\ &= \left(1 - \sum_{i=1}^{j-1} \pi_{it}\right) \cdot \Lambda\left(\lambda_j\left(Z_{jt}, \beta_j\right) - \ln\left[\frac{1-\sum_{i=1}^{j} p_{it}}{p_{jt}}\right]\right), \quad for\ j = 2, ..., K\end{aligned} \tag{9}$$

$$\pi_t^0 = 1 - \sum_{j=1}^{K} \pi_{jt} \tag{10}$$

[11]If we wished instead to retain the simplicity of the test of Engle and Manganelli (1999) we could employ an alternative extension: If we define $Hit_t^* \equiv \left(p_t\left(1-p_t\right)\right)^{-1/2} \cdot \left(Hit_t - p_t\right)$, then we may use OLS to regress Hit_t^* on a constant and variables in the time-t information set in the same manner as Engle and Manganelli (1999). The test that all of the parameters in the Hit^* regression are zero would also be conducted in the same fashion. Standardising the variance of the dependent variable in the hit regression, in addition to standardising the mean as in Engle and Manganelli (1999), is necessary as the conditional variance of Hit_t under the null is $p_t\left(1-p_t\right)$, and thus if p_t is time-varying this causes Hit_t to be heteroscedastic. In the case that p_t is constant this concern obviously does not arise.

where $\Lambda(x) \equiv \frac{1}{1+e^{-x}}$ is the logistic transformation, $Z_t \equiv [Z_1, ..., Z_K]'$ and $\boldsymbol{\beta} \equiv [\beta_1, ..., \beta_K]'$. (Let the length of $\boldsymbol{\beta}$ be denoted K_β.) This rather complicated-looking expression for Π_t is specified in such a way that $\Pi_t(Z_t, \mathbf{0}, P_t) = P_t$ for $t = 1, 2, ..., T$ for all Z_t. Further, it allows each of the elements of Π_t to be a function of a set of regressors, Z_{jt}, while ensuring that each $\pi_{jt} \geq 0$ and that $\Sigma_{j=0}^{K} \pi_{jt} = 1$. Again the competing hypotheses may be expressed as $\boldsymbol{\beta} = 0$ versus $\boldsymbol{\beta} \neq 0$. The likelihood function to be maximised to obtain the parameter $\boldsymbol{\beta}$ is $\mathcal{L}(\Pi(Z, \boldsymbol{\beta}, P) | Hit) = \Sigma_{t=1}^{T} \Sigma_{j=0}^{K} \ln \pi_{jt} \cdot \mathbf{1}\{M_t = j\}$. The joint test may also be conducted as a likelihood ratio test: $LR_{ALL} \equiv -2 \cdot \left(\mathcal{L}(P|Hit) - \mathcal{L}\left(\Pi\left(Z, \hat{\boldsymbol{\beta}}, P\right) |Hit\right)\right) \backsim \chi^2_{K_\beta}$ under the null hypothesis that the model is correctly specified in all K regions.

4 An Application of the Conditional Copula

The theory of copulas has been in use in the applied statistical literature for over twenty years. One of the earliest applications of copulas was Clayton's (1978) study of familial tendency in chronic disease. Other studies have included Cook and Johnson's (1981) and Genest and Rivest's (1993) analysis of hydrogeochemical data used in the exploration for uranium and Oakes' (1989) analysis of Fox River flood data. The application of copula theory to economic problems is a much more recent phenomenon.

Perhaps the earliest paper to propose the use of the theory of copulas in the analysis of economic problems was Embrechts, McNeil and Straumann (1999), or the more technical version of the same paper: Embrechts, McNeil and Straumann (2000). In these papers the authors outline the 'properties and pitfalls' of correlation as a measure of dependence. Rosenberg (2000) uses copula theory in the pricing of a financial option having two underlying assets, in his case these assets were the S&P500 and the DAX 30 (the latter being a price index of thirty blue chip German stocks). Joe (1997) provides a detailed look at how the use of ARMA processes to model serial dependence may be generalised to utilise the theory of copulas. Bouyé, *et al.* (2000) extend this theory by using copulas to describe serial dependence in continuous time stochastic processes such as Brownian motion and the Ornstein-Uhlenbeck process. Finally, Costinot, *et al.* (2000) apply (unconditional) copulas to the study of dependence between financial markets during extreme events, one example of which is the 1997-1998 Asian financial crisis. No paper, to our knowledge, has considered applying to the modelling of time-varying conditional distributions.

Having discussed the extension of copula theory to the conditional case, we now apply the theory of conditional copulas to the modelling of the conditional bivariate distribution of the daily Deutsche mark - U.S. dollar (DM-USD) and Japanese yen - U.S. dollar (Yen-USD) exchange rate returns over the period January 2, 1991 to October 12, 2000. This represents the post-unification era in Germany (the countries were united in November of 1989, and some financial integration was still being carried out during 1990) and includes the first twenty-two months of the euro's reign as the official currency of Germany[12]. These two exchange rates are of interest as they are the two most heavily traded currency pairs, representing close to 50% of total foreign exchange trading volume (see Melvin, 2000). Given their status, the DM-USD and Yen-USD exchange rates have been relatively widely studied, see Andersen and Bollerslev (1998), Diebold *et al.* (1999), Andersen *et al.* (2000), amongst others.

In addition to the economic interest in these series, they also represent a statistically interesting pair of series. While there is generally very little time variation in the conditional means of exchange rates, it is well documented that the conditional variances of exchange rates vary systematically over time, see the papers mentioned above, and also Bollerslev (1987), Bollerslev (1990), Engle *et al.* (1990), and Kearney and Patton (2000), amongst many others. Thus we have an abundance of evidence that the conditional marginal distributions vary over time. Evidence also exists that the dependence between these two assets varies over time: the multivariate GARCH literature, of which many of the previously mentioned papers are a part, provides ample evidence that the conditional correlation between these two exchange rates varies over time, suggesting that perhaps the entire dependence structure may be time-varying, though as mentioned above, there may be better ways of capturing this time variation in dependence than through conditional correlation. The specification of a conditional joint distribution via the combination of two time-varying conditional marginal distributions and a potentially time-varying conditional copula, then, seems perfectly suited to the problem at hand.

Up until this point we have not much discussed the conditioning set, $\mathcal{F}$, we have merely shown that the existing results in the statistics literature hold when there exists such a conditioning set. For time series applications the natural conditioning set to consider is the sigma algebra generated by all previous observations, i.e., $\mathcal{F}_t = \sigma(x_t, y_t,\ x_{t-1}, y_{t-1}, ...,\ x_1, y_1)$. With $\mathcal{F}_t$ defined in this way we then re-write equation (3) as

$$H_t(x_t, y_t|\mathcal{F}_{t-1}) = C_t(F_t(x_t|\mathcal{F}_{t-1}), G_t(y_t|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}), \quad x, y \in \bar{R},\ t = 1, 2, ..., T \tag{11}$$

The above equation makes it clear that the joint distribution of (X_t, Y_t) may differ from the joint distribution of (X_{t-1}, Y_{t-1}). Thus a sample of pairs of observations, $\{(x_t, y_t)\}_{t=1}^{T}$, may not represent T observations of the same joint distribution, but T observations from T different joint distributions. Obviously, without assuming some structure we cannot attempt to estimate the form of H, or that of C, F, or G. What is often assumed is that the functional form of the distribution remains constant over time, while the parameters of the distribution vary according to some equation. In modelling the marginal distributions, for example, we will assume that the conditional means evolve according to an autoregressive process, and that the conditional variances

[12]The mark is still to be used for transactions in Germany until the end of 2001, but the mark/Euro exchange rate was fixed on January 1, 1999, and all international transactions are denominated in Euros. See the European Central Bank web site (http://www.ecb.int) for more information.

evolve according to a GARCH(1,1) process. The distributions of the standardised innovations in both margins are assumed to be Student's t distributions for the entire sample period.

We can similarly think about the evolution of C_t. There are three possible time paths for C_t: The first is the degenerate case that it does not vary at all. The second case is that the functional form of the conditional copula remains fixed while the parameters of the conditional copula evolve through time; an analogous form of time-variation to that assumed for the marginal distributions in this paper. The third form of time-variation involves changes in both the form of the conditional copula and the parameters of the conditional copula. Nelsen (1999) shows that any convex linear combination of copulas is also a copula[13], and so one possible way of modelling the latter form of time-variation would be to set the conditional copula to a weighted sum of various types of copulas, each with parameters that may or may not vary through time, and allow the weights to vary over time. In this paper we will consider only the first and second types of variation in the conditional copula.

4.1 Description of the Data

As mentioned above, the data set used for this analysis comprises daily Deutsche mark - U.S. dollar and Japanese yen - U.S. dollar exchange rates over the period 2 January 1991 to 12 October, 2000, giving us 2513 observations. The data were taken from the database of Datastream International. As usual, we take the log-difference of each exchange rate, and multiply by 100. Table 1 below presents some summary statistics of the data.

[INSERT TABLE 1 HERE]

The above table shows that neither exchange rate had a significant trend over the sample period, both means being very small relative to the standard deviation of each series. Both series also exhibit slight negative skewness, and substantial excess kurtosis. The Jarque-Bera test of the normality of the unconditional distribution of each exchange rate strongly rejects the null, suggesting that neither exchange rate return series is unconditionally normal. We also test for the presence of serial correlation up to the 20^{th} lag in the squared returns, an indication of ARCH-type heteroscedasticity, via the ARCH LM test of Engle (1982). As expected, for both series there exists strong evidence of serial correlation in the squared returns, providing evidence that both F_t and G_t are time-varying.

4.2 The Model

In specifying a model of the bivariate density of DM-USD and Yen-USD exchange rates we must specify three models: the models for the marginal distributions of each exchange rate, and the model for the conditional copula. The models for the marginal distributions must be close enough to the

[13] We used this result to construct the 'Mixture of Normals' copula presented in Figure 1.

unknown true distributions that we cannot reject a test that the probability integral transformations of each of the marginals is *i.i.d.* $Unif(0,1)$. Of course, we do not expect that we will be able to find the actual data generating process; we merely want to get as close as possible to it. In addition to the transforms of each marginal being independent of their own lags, they must also be independent of each other's lags. Recall that the conditioning set for each marginal distribution (and the conditional copula) must be the same, thus each transformed variable must be independent of the information in the conditioning set of its marginal distribution. Tests of these conditions are discussed in Section 4.3 below.

4.2.1 The models for the marginal distributions

The models employed for the marginal distributions are presented below. We will denote the log-difference of the DM-USD exchange rate as the variable X_t, and the log-difference of the Yen-USD exchange rate as the variable Y_t.

$$X_t = \mu_x + \phi_{1x} X_{t-1} + \varepsilon_t \tag{12}$$

$$h_t^x = \omega_x + \beta_x h_{t-1}^x + \alpha_x \varepsilon_{t-1}^2 \tag{13}$$

$$\sqrt{\frac{\upsilon_x}{h_t^x(\upsilon_x - 2)}} \cdot \varepsilon_t \overset{\mathfrak{D}}{=} t_{\upsilon_x} \tag{14}$$

$$Y_t = \mu_y + \phi_{1y} Y_{t-1} + \phi_{10y} Y_{t-10} + \eta_t \tag{15}$$

$$h_t^y = \omega_y + \beta_y h_{t-1}^y + \alpha_y \eta_{t-1}^2 \tag{16}$$

$$\sqrt{\frac{\upsilon_y}{h_t^y(\upsilon_y - 2)}} \cdot \eta_t \overset{\mathfrak{D}}{=} t_{\upsilon_y} \tag{17}$$

That is, the marginal distribution for the DM-USD exchange rate is assumed to be completely characterised by an AR(1), t-GARCH(1,1) specification, while the marginal distribution for the Yen-USD exchange rate is assumed to be characterised by an AR(1,10)-t-GARCH(1,1) specification[14]. In our particular case it happened that we only needed univariate models for these two marginal distributions (no lags of the 'other' variable appear in either variable's model). This will not always be so.

[14] The marginal distribution specification tests, described in Section 4.3, suggested that the model for the conditional mean of the Yen-dollar exchange rate return needed the tenth lag. This lag was not required for the DM-dollar exchange rates.

4.2.2 The models for the copula

In selecting a copula to use, we must have a clear idea of the properties of the data under analysis. Many of the copulas presented in the statistics literature are best suited to variables that take on extreme values in only one direction: survival times (Clayton, 1978), concentrations of particular chemicals (Cook and Johnson, 1981 and Genest and Rivest, 1993), flood data (Oakes, 1989). However, exchange rates have extremes in both directions: large positive *and* negative returns.

For the purposes of comparison, we will specify and estimate two alternative copulas, the Gaussian copula and the Joe-Clayton copula, both with and without time variation. The first copula considered, the Gaussian (or Normal) copula is the dependence function associated with bivariate normality, and is extracted via Corollary 5. It is given in equation (18) below.

$$C(u,v|\rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left\{\frac{-(r^2 - 2\rho rs + s^2)}{2(1-\rho^2)}\right\} dr\ ds,\ -1 < \rho < 1 \tag{18}$$

where Φ^{-1} is the inverse of the standard normal *c.d.f.*

The transformations $\Phi^{-1}(u) = \Phi^{-1} \circ F(x)$ and $\Phi^{-1}(v) = \Phi^{-1} \circ G(y)$ transform the variables X and Y, which are distributed according to F and G, into standard normal random variables. The normal copula takes as arguments the standard normal transforms of X and Y, and assumes that they are *jointly* normally distributed. This is how we are able to back out the dependence implied by bivariate normality. We estimate two forms of the normal copula: one assuming a constant correlation parameter, ρ, and the other allowing ρ_t to vary over time. We propose the following evolution equation for ρ_t:

$$\rho_t = \tilde{\Lambda}\left(\omega_\rho + \beta_\rho \cdot \rho_{t-1} + \alpha \cdot \frac{1}{10}\sum_{j=1}^{10} \Phi^{-1}(u_{t-j}) \cdot \Phi^{-1}(v_{t-j})\right) \tag{19}$$

where $\tilde{\Lambda}(x) \equiv \frac{1-e^{-x}}{1+e^{-x}}$ is the modified logistic transformation, designed to keep ρ_t in $(-1,1)$ at all times.

Equation (19) reveals that we assume ρ_t follows something akin to a restricted ARMA(1,10) process: we include ρ_{t-1} as a regressor to capture any persistence in the dependence parameter, and the mean of the product of the last ten observations of the transformed variables $\Phi^{-1}(u_{t-j})$ and $\Phi^{-1}(v_{t-j})$, to capture any variation in dependence.

The second copula that will be used is the 'BB7' copula of Joe (1997), which we will refer to as the Joe-Clayton copula, as it is constructed by taking a particular Laplace transformation of

Clayton's copula[15]. The unconditional version of this copula is

$$C(u,v|\kappa,\gamma) = 1 - \left(\left\{[1-(1-u)^{\kappa}]^{-\gamma} + [1-(1-v)^{\kappa}]^{-\gamma} - 1\right\}^{-1/\gamma}\right)^{1/\kappa}, \ \kappa \geq 1, \ \gamma > 0 \tag{20}$$

Joe (1997) asserts that this copula has a number of nice properties: it collapses to Clayton's copula when $\kappa = 1$, and the Fréchet-Hoeffding upper bound[16] is approached when either parameter approaches infinity. One of the most useful properties of the Joe-Clayton copula, for our purposes, is the way in which the parameters of the copula relate to a particular measure of dependence known as *tail dependence.* This measure of dependence is defined below.

Definition 6 *If the limit*

$$\lim_{\varepsilon\to 0} \Pr[U \leq \varepsilon | V \leq \varepsilon, \mathcal{F}] = \lim_{\varepsilon\to 0} \Pr[V \leq \varepsilon | U \leq \varepsilon, \mathcal{F}] = \lim_{\varepsilon\to 0} C(\varepsilon, \varepsilon|\mathcal{F})/\varepsilon = \tau^L$$

exists, then the conditional copula C *exhibits lower tail dependence if* $\tau^L \in (0,1]$ *and no lower tail dependence if* $\tau^L = 0$. *Similarly, if the limit*

$$\lim_{\delta\to 1} \Pr[U > \delta | V > \delta, \mathcal{F}] = \lim_{\delta\to 1} \Pr[V > \delta | U > \delta, \mathcal{F}] = \lim_{\delta\to 1} (1 - 2\delta + C(\delta,\delta|\mathcal{F}))/(1-\delta) = \tau^U$$

exists, then the conditional copula C *exhibits upper tail dependence if* $\tau^U \in (0,1]$ *and no upper tail dependence if* $\tau^U = 0$.

Tail dependence is an interesting measure of dependence as it captures the behaviour of the random variables during extreme events. Informally, it measures the probability that we will observe an extremely large positive (negative) realisation of one variable, given that we have observed that the other variable also took on an extremely large positive (negative) value. As an example, the bivariate normal distribution (and thus the normal copula) has $\tau^U = \tau^L = 0$ for correlation not equal to one, meaning that in the extreme tails of the distribution the variables are independent. The bivariate Student's t distribution, on the other hand, has both $\tau^U \neq 0$ and $\tau^L \neq 0$ for correlation not equal to one, implying that even at the most extreme tail of the distribution (indeed, the *limit* of the distribution) the variables are dependent. The Joe-Clayton copula allows upper and lower tail dependence to range anywhere from zero to one.

Joe (1997) writes that the tail dependence parameters of the Joe-Clayton copula are given by: $\tau^L(\kappa,\gamma) = 2^{-1/\gamma}$ and $\tau^U(\kappa,\gamma) = 2 - 2^{1/\kappa}$. Notice that the lower tail dependence is defined completely by γ, and the upper tail dependence is defined completely by κ. We will use the one-to-one mapping of each of the parameters of the Joe-Clayton copula to a tail dependence measure

[15] For more details on the construction of this copula or on Laplace transformations in copula theory, the reader is referred to Joe (1997).

[16] The Fréchet-Hoeffding upper bound is a theoretical upper bound on the value that a joint distribution can take at any given point. This upper bound corresponds to perfect positive dependence between the two random variables.

to assist us in defining an evolution equation for the parameters. The difficulty in specifying how the parameters evolve over time lies in defining the forcing variable for the equation. Unless the parameter has some interpretation, as ρ does in the normal copula, it is very difficult to know what might (should) influence it to change. In view of this, we will define the time variation in the Joe-Clayton copula in terms of time variation in the upper and lower tail dependence measures, and then find the parameters of the copula that correspond to the given upper and lower tail dependence measures at each point in time. The evolution equations for the Joe-Clayton copula are:

$$\tau_t^U = \Lambda\left(\omega_U + \beta_U \tau_{t-1}^U + \alpha_U \cdot \frac{1}{10}\sum_{j=1}^{10} |u_{t-j} - v_{t-j}|\right) \tag{21}$$

$$\tau_t^L = \Lambda\left(\omega_L + \beta_L \tau_{t-1}^L + \alpha_L \cdot \frac{1}{10}\sum_{j=1}^{10} |u_{t-j} - v_{t-j}|\right) \tag{22}$$

where $\Lambda(x) \equiv \frac{1}{1+e^{-x}}$ is the logistic transformation, used to keep τ^U and τ^L in $(0,1)$ at all times.

In the above equations we propose that the upper and lower tail dependence parameters each follow again an ARMA(1,10)-type model. The right hand side of the model for the tail dependence evolution equation contains an autoregressive term, $\beta_U \tau_{t-1}^U$ and $\beta_L \tau_{t-1}^L$, and a forcing variable. Identifying a forcing variable for a time-varying limit probability is somewhat difficult: limits are not really an empirical concept. We propose using the mean absolute difference between u_t and v_t over the previous ten observations as a forcing variable[17]. The intuition behind this can be explained with the aid of Figure 2. If X and Y are perfectly positively dependent (otherwise known as 'comonotonic') then the transformed variables U and V will all lie on the main diagonal of the unit square. The absolute value of the difference between u_t and v_t is proportional to the minimum distance from the point (u_t, v_t) to the main diagonal, and we thus use the mean absolute difference between u_t and v_t over the previous ten observations as an indication of how far from comonotonicity the data were.

As the upper and lower tail dependence measures are one-to-one functions of the two conditional copula parameters, we can compute the γ_t and κ_t implied by a particular τ_t^L and τ_t^U as follows: $\gamma_t = \gamma(\tau_t^L) = -\left[\log_2(\tau_t^L)\right]^{-1}$ and $\kappa_t = \kappa(\tau_t^U) = \left[\log_2\left(2 - \tau_t^U\right)\right]^{-1}$. Thus, in addition to specifying the evolution of the tail dependence parameters over the sample, equations (21) and (22) also specify the evolution of the parameters of the copula.

[17]A few variations on this particular forcing variable were used, such as weighting the observations by how close they are to the extremes, or by using an indicator variable for whether the observation was in the first, second, third or fourth quadrant. No significant improvement was found, and so we have elected to use the simplest model.

4.2.3 Estimating the model

Maximum likelihood is the natural estimation procedure to use in this context: in specifying models for the two marginal distributions and the copula, we have defined a joint distribution function for the two exchange rates, and thus a joint likelihood. Further, the procedure employed to develop the joint distribution lends itself naturally to multi-stage estimation of the model. Although estimating all of the coefficients simultaneously yields the most efficient estimates, the large number of parameters can make numerical maximisation of the likelihood function difficult. In this paper we make use of two-stage maximum likelihood results, see White (1994) for the general theory and Patton (2001) for the details on applying the theory to the estimation of copula models. Under the usual conditions the estimates obtained via maximum likelihood are consistent and asymptotically normal. We adjust the standard errors of the copula parameter estimates to reflect the fact that they are based on marginal distribution parameters estimated in an earlier stage[18].

The Broyden, Fletcher, Goldfarb and Shanno (BFGS) algorithm was used to maximise the likelihood. The Matlab code written to carry out the computations for this paper, as well as the data set used, will be available on the author's web site in the near future. The results of the estimations are presented below.

4.2.4 For Comparison: Normal and Student's t BEKK models

For the purposes of comparison we also estimate two alternative models using existing techniques (the results are not presented in the interests of parsimony, but are available from the author upon request). For both of the additional models we first model the conditional means of the two exchange rate returns series, using the models in equations (12) and (15). We then estimate one of the more flexible multivariate GARCH models on the residuals: the BEKK(1,1) model introduced by Engle and Kroner (1995). This model is written as:

$$H_t = C'C + B'H_{t-1}B + A'e'_{t-1}e_{t-1}A \tag{23}$$

where $H_t = \begin{bmatrix} h_t^x & h_t^{xy} \\ h_t^{xy} & h_t^y \end{bmatrix}$, $C = \begin{bmatrix} c_{11} & 0 \\ c_{12} & c_{22} \end{bmatrix}$, $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, h_t^x is the conditional variance of X at time t, and h_t^{xy} is conditional covariance between X and Y at time t.

The two models differ in their assumption regarding the joint distribution of the residuals: the first model assumed bivariate normality, while the second assumes a bivariate Student's t distribution.

[18] We have estimated these models both via one-stage (simultaneous) maximum likelihood and the two-stage method described in the text. The parameter estimates and standard errors obtained from the two proceedures were not very different, and so we have elected to present the more concise two-stage estimation results.

4.3 Results for the Marginal Distributions

The parameter estimates and standard errors for marginal distribution models are presented in Table 2. From this table we see that the degrees of freedom parameter for the two exchange rates are indeed different: 5.8073 for the DM-USD margin and 4.3817 for the Yen-USD margin, implying that the distribution of Yen-USD exchange rate returns has fatter tails than the corresponding distribution for the DM-USD exchange rate. The t-statistic for the significance of difference in the degrees of freedom parameters is 1.9192, indicating that they are not significant different at the 5% alpha level[19] (though they would be at the 5.5% alpha level).

[INSERT TABLE 2 HERE]

In Table 3 we present the results of the LM tests for independence of the probability integral transforms, U and V, of the original data and the Kolmogorov-Smirnov (K-S) tests for the correctness of the density specification. Both of these tests were introduced in Section 3. This table indicates no evidence of serial correlation in the first four moments of U and V, suggesting that the the AR and $GARCH$ models proposed for the conditional means and variances of the two exchange rate returns are adequate[20]. The Kolmogorov-Smirnov test for the correctness of the density specification yields p-values of 0.8706 and 0.9114, suggesting that the density specifications are also adequate.

[INSERT TABLE 3 HERE]

In light of the possible low power of the K-S test, we employ the hit tests discussed in Section 3 to check for the correctness of the dynamic and the density specifications in particular regions of the support. We chose to use the five following regions: the lower 10% tail, the interval from the 10^{th} to the 25^{th} quantile, the interval from the 25^{th} to the 75^{th} quantile, the interval from the 75^{th} to the 95^{th} quantile, and the upper 10% tail. These regions represent economically interesting subsets of the support - the upper and lower tails are notoriously difficult to fit, and so checking for correct specification there is important, while the middle 50% of the support contains the 'average' observations. We use as regressors ('Z_{jt}' in the notation of Section 3) a constant, to check that the model implies the correct proportion of hits, and three variables that count the number of hits in that region in the last day, week and month, to check that the model dynamics are correctly specified[21]. The λ_j functions are set to simple linear functions of the parameters and the regressors: $\lambda_j\left(Z_{jt},\beta_j\right) = Z_{jt}\cdot\beta_j$. The results of the tests in the individual regions and the joint test for all

[19] All tests in this paper will be conducted at the 5% alpha level.

[20] We also checked for serial correlation in the first four moments of the (un-transformed) standardised residuals of the two exchange rate models. These also indicated that no serial dependence was present.

[21] We also conducted this test including as additional regressors three variables that counted the number of hits in the corresponding region of the other variable's support over the last day, week and month. The results did not change significantly.

five regions are presented in Table 4 below. For comparison, we also present the results of the hit tests for the Normal BEKK and t-BEKK models discussed in the previous section.

[INSERT TABLE 4 HERE]

Table 4 shows how clearly the assumption of conditional normality for these exchange rate returns is rejected. The joint density model found by estimating the AR models for the mean and a BEKK(1,1) model for the variance, with a bivariate normal distribution assumed for the standardised residuals fails in 3 out of 5 regions for the DM-USD margin and in 4 out of 5 regions for the Yen-USD margin. Both margins fail the joint test that all regions are correctly specified. The DM-USD margin of the t-BEKK model passes the hit test in all regions, and passes the joint test. The Yen-USD margin, on the other hand, fails the hit test in the lower tail region, and fails the joint test. Thus we conclude that even the very flexible BEKK model is not appropriate when used with the assumption of a bivariate Student's t distribution - a more flexible specification of the joint distribution is required.

The model for the DM-USD marginal distribution to be used with the copula models (denoted 'Copula DM') passes both the individual and the joint tests that all regions are correctly specified. The 'copula Yen' model fails in the lower tail region, but passes the joint test. We take the finding that both marginal models for the copula distributions pass the joint test as evidence that they are near enough to the $Unif\,(0,1)$ distribution for us to move on to modelling the copula. That the t-BEKK model for the Yen-USD margin fails the joint test, while the more flexible 'copula Yen' model passes suggests that although the difference in the degrees of freedom parameters of the DM-USD margin and the Yen-USD margin was not significant at the 5% alpha level, it is still important.

4.4 The Four Models' Results

We now present the results of the estimation of the four models described in Section 4.2. Recall that the four models proposed all assumed that the marginal distributions of the exchange rates are described by an AR-tGARCH models, and differ only in their specification of the conditional copula. The first two models assume that the conditional copula is constant, while the second two allow for time variation in the parameter(s) of the conditional copula. Note that even with a constant copula the resulting conditional bivariate density is time varying due to the variation in the marginal distributions. The results for the various copula models are collected and presented in Table 5.

4.4.1 The constant normal copula

The first conditional copula model estimated was the normal copula with constant correlation. The parameter of this copula, ρ, is equal to 0.4560, indicating a relatively high amount of association between these two exchange rates. We can also find the implied unconditional correlation between the standardised residuals of the variables. The calculation of this correlation involves the evaluation of a double integral, and we approximate this quantity numerically using quadrature[22]. The implied unconditional residual correlation is 0.4141, a figure quite close to the unconditional correlation of the standardised residuals of the two models, and also to that between the raw exchange rate return series; those two figures being 0.4468 and 0.4335 respectively.

[INSERT TABLE 5 HERE]

4.4.2 The constant Joe-Clayton copula

The second conditional copula model we discuss is the model formed by joining the marginal distributions using a Joe-Clayton copula with constant tail dependence. The parameters κ and γ imply unconditional tail dependence measures of $\tau^U = 0.3197$ and $\tau^L = 0.1921$, suggesting that there is greater dependence between these two series in large 'up' days than in large 'down' days. A large 'up' day corresponds to one where the US dollar appreciated heavily against the mark and the yen, and similarly a large 'down' day corresponds to a day when the US dollar depreciated heavily against the mark and the yen. Taken together, these results suggest that the limiting probability of the dollar appreciating heavily against the mark, given that it has appreciated heavily against the yen, is about one-third. (As tail dependence is a symmetric concept, it does not matter on which of the two currencies one conditions on the dollar having appreciated against.) The corresponding depreciation probability is about one-fifth, meaning that the exchange rates are less dependent in bad markets (for the US dollar) than in good markets. We can test for the significance of the asymmetry in the dependence between these two exchange rates, by testing whether τ^U is significantly different from τ^L. The test statistic for this is simply the difference, 0.1275, divided by the standard deviation of this difference, which is approximately[23] 0.0437. This leads to a t-statistic (*p-value*) of 2.9197 (*0.0035*), indicating that the asymmetry in the dependence between these two exchange rates is indeed significant.

The unconditional correlation between the standardised residuals implied by the Joe-Clayton copula parameters is 0.4267, quite close to the correlation implied by the constant Normal copula. Thus both copulas imply approximately the same *linear* dependence between the two exchange rates, while differing on the other forms of dependence that may exist.

4.4.3 The time-varying normal copula

The third model estimated was the normal copula with time-varying correlation, as described in equations (18) and (19). The positive sign of β_ρ indicates, as expected, a positive relationship

[22] We use Gauss-Legendre quadrature, with ten nodes for each margin, leading to a total of 100 nodes. See Judd (1998) for more on this technique.

[23] As this is a test of a nonlinear restriction of the estimated parameters, we approximate the variance of the restriction using a Taylor series expansion of the nonlinear function of the estimated parameters about the function evaluated at the true parameters. For more details, see Chapter 7 of Greene (1997).

between the correlation parameter at time t and that at time $t+1$. A plot of the implied conditional correlation over the sample period, presented in the top panel of Figure 4, confirms this.

From this graph we can see that the first two and a half years was a period of higher correlation than the average (the conditional correlation is generally above the unconditional correlation drawn with a dashed line). This is followed by about a year of lower than average correlation, and then another increase in correlation that lasts about two years. We can see that towards the end of the sample the conditional correlation volatility increases, and finally in February 2000 the conditional correlation went negative. The period of higher volatility in the conditional correlation corresponds to the lead up to the introduction of the euro in Germany, while the last 22 months of data represent the first months of the euro's life. It might be reasonable to expect that a structural break in the dependence relation between the DM-USD and Yen-USD exchange rates occurred upon the introduction of the euro, and we investigate this possbility in Section 4.6. From the current results it certainly does appear that the dependence has changed, though it is not clear from the sample whether the dependence between these exchange rates will remain negative or return to its historical level.

[INSERT FIGURE 4 HERE]

4.4.4 The time-varying Joe-Clayton copula

The fourth and final model estimated is that using the time varying Joe-Clayton copula described in equations (20) to (22). As expected, the coefficients on the mean absolute difference between u_t and v_t for the previous ten periods, α_U and α_L, are negative, indicating that a smaller mean difference leads to an increase in tail dependence. Both autoregressive parameters are positive, as expected.

An interesting finding from these results is that the upper tail dependence measure appears more persistent that the lower tail dependence measure: the coefficient β_U is greater than β_L, and the coefficient on the forcing variable in the lower tail dependence equation, α_L, is greater than α_U. In Figures 6 and 7 we present the time path of the tail dependence measures and the time path of the implied copula parameters.

[INSERT FIGURE 6 HERE]

[INSERT FIGURE 7 HERE]

In Figure 7 we immediately see that the upper tail dependence is consistently higher than the lower tail dependence, in fact, on over 99% of the days in the sample estimated conditional upper tail dependence was greater than conditional lower tail dependence. The greatest difference in the tail dependence measures was over 0.50; a very large amount for a probability. As with the conditional correlation from the time varying normal copula, we see that there were two episodes of increased dependence between these exchange rates, each about two years long. Also similar to the results from the normal copula is the reduced dependence between these exchange rates since the introduction of the euro: both the upper and lower tail dependence measures were below their unconditional levels for the final year of the sample. As the Joe-Clayton copula is parameterised by the upper and lower tail dependence measures, it cannot capture negative dependence: tail dependence is a *positive* dependence concept. The lowest possible tail dependence is zero, corresponding to independence. Thus while the normal copula suggests that the dependence relation became negative in the last ten to twelve months of the sample, the Joe-Clayton copula results imply that the dependence has been reduced, but is still positive. This constraint on the Joe-Clayton copula is a drawback, and clearly if one thought that there was a significant chance of the conditional dependence being negative one would have to choose a copula that could capture this change in the direction of the dependence. We will determine in the next section whether this constraint of the Joe-Clayton copula lead to it performing worse than the normal copula in tests of goodness-of-fit.

In Figure 5 we present the time series of conditional correlation implied by the results for the time-varying Joe-Clayton copula. The conditional correlation at each point in the sample is computed via numerical integration, and we present for the purposes of comparison the unconditional correlation implied by the constant Joe-Clayton copula in this figure also. We can see that the shape of the time path of the conditional correlation is similar to that observed for the conditional tail dependence, with two pronounced periods of increased dependence, and declining dependence towards the end of the sample. We present this particular measure of dependence to show how the two copula functional forms differ in their predictions of a common dependence measure. This shows just how dependent the implied conditional correlation is on the copula functional form assumption - most previous studies that estimate equations for the conditional correlation assume bivariate normality, and so would obtain a result similar to the top panel of Figure 4.

[INSERT FIGURE 5 HERE]

4.4.5 Summary of results

In summary, these results show that we have substantial evidence of asymmetric depedence, and qualitative evidence of time-varying dependence. Asymmetry in the constant Joe-Clayton copula was found to be significant, and indicates that dependence is greater during appreciations of the U.S. dollar than during depreciations of the U.S. dollar. Further evidence in support of asymmetry was found in the results from the time-varying Joe-Clayton copula model: on over 99% of days in the sample the upper tail dependence measure was greater than the lower tail dependence measure. Time variation in the conditional copula seems significant: when the parameters of the conditional copulas are allowed to vary through time, they deviate quite substantially from the parameter found when conditional dependence is assumed to be constant.

We compared the normal and the Joe-Clayton copula results by looking at the residual conditional correlation implied by each of the models. We found that the estimate of the conditional correlation varied somewhat with the choice of copula model, indicating that which copula is chosen is not innocuous, even for simple measures of dependence like linear correlation.

It should be noted, indeed it is one of the main benefits of using copula theory, that we could have chosen *any* other dependence measure to compare the copula models. Knowledge of the conditional copula and marginal distributions is sufficient for one to compute any measure of conditional dependence that may be of interest.

4.5 Comparing the Alternative Models

The evaluation of these models is more involved than usual, as we wish to evaluate the entire conditional bivariate density model, not merely a particular set of conditional moments of the bivariate density. Further, the fact that most of the models considered are not nested in other models as special cases means that standard hypothesis testing theory is not available to us. A naïve estimate of the likelihood ratio statistic comparing the time-varying Joe-Clayton to the time-varying normal copula yields a value of 21.61, giving a p-value of 0.0001 in favour of the Joe-Clayton copula, however, these figures are strictly speaking not valid. Vuong (1989) presents a means of conducting likelihood ratio tests of non-nested hypotheses, but only for *i.i.d.* data, and so we cannot directly compare the likelihoods.

Additionally, testing for the significance of time-varying dependence versus constant dependence is complicated by the fact that, like testing for homoscedasticity versus a GARCH(1,1) alternative, at least one parameter is unidentified under the null hypothesis[24]. Here the naïve likelihood ratio statistics (*p-values*) for the significance of time variation in the normal and Joe-Clayton copulas are 85.16 (*0.0000*) and 96.99 (*0.0000*) respectively. These figures, though, cannot be interpreted in the standard way, due to the presence of the unidentified nuisance parameter; we present the figures merely for illustrative purposes. We now turn to the tests described in Section 3.

It turned out that all four copula models pass the LM and Kolmogorov-Smirnov tests (described in Section 4.3) that the sequences $\{u_t^c\}_{t=1}^T$ and $\{v_t^c\}_{t=1}^T$ are *i.i.d.* $Unif\,(0,1)$, implying that we cannot reject any of the null hypotheses that a model is correctly specified[25]. This is a surprising result, as the four models are clearly quite different. The result that *all* models pass the K-S test may indicate the possible low power of this test rather than that all four models are adequately specified. We thus turn to the hit tests for (hopefully) a more powerful test of goodness-of-fit.

We divide the support of the copula into seven rectangular regions[26], each of which with an economic interpretation, and one 'remnant' region. The regions are presented graphically in Figure

[24] Andrews and Ploberger (1994) propose a means of overcoming the presence of a nuisance parameter that is unidentified under the null. This test, however, does not lend itself easily to models that require more difficult maximum likelihood estimation, rather than simple OLS.

[25] The results of these tests are omitted due to space constraints, but are available from the author on request.

[26] Using rectangular regions makes computing the probability of a hit in that region implied by the copula model, $\hat{C}$, particularly simple: it is just the $\hat{C}$-volume of the region, defined in Section 2.

3. Regions 1 and 2 correspond to the lower and upper 10% Value-at-Risk for each variable. The ability to correctly capture the probability of both exchange rates taking on extreme values simultaneously is of critical importance to portfolio managers and macroeconomists, amongst many others. Regions 3 and 4 represent moderately large up and down days: days in which both exchange rates were somewhere between their 10^{th} and 25^{th}, or 75^{th} and 90^{th}, quantiles. Region 5 is the 'median' region: days when both exchange rates were in the middle 50% of their distributions. Regions 6 and 7 are the extremely asymmetric days, those days when one exchange rate was in the upper 25% of its distribution while the other was in the lower 25% of its distribution. This part of the support is important for diversification reasons: if we can correctly model these areas we may better diversify risk.

[INSERT FIGURE 3]

We again specify a simple linear function for λ_j, that is: $\lambda_j\left(Z_{jt}, \beta_j\right) = Z_{jt} \cdot \beta_j$, and we include in Z_{jt} a constant term, to capture any over- or under-estimation of the unconditional probability of a hit in region j, and three variables that count the number of hits that occurred in the past day, one week and one month, to capture any violations of the assumption that the hits are serially independent. The results for each of the seven regions, for the four models considered are presented below. For the joint test we define the zeroth region as that part of the support not covered by regions one to seven.

[INSERT TABLE 6 HERE]

Table 6 reveals that both the constant the time-varying normal copulas pass the hit tests for all regions and pass the joint test. Both the constant and the time-varying Joe-Clayton copulas pass the hits tests in all but region 2, the upper tail region, and both fail the joint test. This result is somewhat surprising, as the Joe-Clayton copula was selected for inclusion in this study because it offered the ability to flexibly model the upper and lower tails of the joint distribution, whereas the normal copula is more restrictive. From these results it would appear that although the Joe-Clayton copula yielded a substantially higher likelihood, it does not perform as well as the model using the normal copula. One possible explanation for this lies in the fact that the Joe-Clayton copula, as mentioned above, cannot capture negative dependence. The time path of the conditional correlation implied by the time-varying normal copula suggested that the dependence between these two exchange rate returns became negative following the introduction of the euro, and thus the Joe-Clayton would be inappropriate. In the next section, we investigate the behaviour of these exchange rates before and after the introduction of the euro, allowing for a structural break.

4.6 A Structural Break - The Euro

On the 1^{st} of January, 1999 the euro was introduced. Eleven European countries[27] agreed to an irrevocable conversion rate between their currencies and the new euro, the conversion rate for the Deutsche mark is 1 euro = 1.95583 marks[28]. In this final section we examine the impact that the introduction of the euro had on the conditional joint distribution of these exchange rates.

The data used in this study is comprised of daily Deutsche mark - U.S. dollar and Japanese yen - U.S. dollar exchange rates over the period 2 January 1991 to 12 October, 2000, which includes 461 observations in the period following the introduction of the euro. There are substantial reasons to believe that both the marginal distribution of the mark and the joint distribution of the DM-USD and Yen-USD exchange rate returns underwent a structural break on January 1, 1999. The DM-USD exchange rate is now 'pegged' to the euro-USD exchange rate, with the euro essentially being a portfolio of eleven currencies, and the percentage return, obtained by taking the log-difference of the level of the exchange rate, on the DM-USD exchange rate is by definition exactly equal to that on the euro-USD exchange rate. Thus what drives the dynamics in the observed DM-USD returns in the post-euro sample are the determinants of the euro-USD exchange rate. Factors relating to the mark will still have a large impact on the euro-USD exchange rate, due to the relative importance of the German economy within the eleven countries, but they will obviously not be the sole determinants.

We can examine the impact on the introduction of the euro on the joint distribution of the DM-USD and Yen-USD exchange rates by allowing the parameters of the joint distribution to change between the pre- and post-euro subsamples. Note that allowing the parameters to change pre- and post-euro is equivalent to increasing the information set, previously defined as $\mathcal{F}_t = \sigma\left(x_t, y_t, x_{t-1}, y_{t-1}, ..., x_1, y_1\right)$, to include an exogenous indicator variable, W_t, which takes the value 0 in the pre-euro sample and 1 in the post-euro sample. Thus we now define the information set as $\mathcal{F}_t = \sigma(x_t, y_t, w_{t+1}, x_{t-1}, y_{t-1}, ..., x_1, y_1)$. One of the minor complications that arise when moving from the standard (unconditional) copula case to the conditional copula case is that all components of the joint distribution (the two marginal distributions and the copula) must be based on the same information set. This implies that although we have no reason to expect that the Yen-USD marginal distribution also underwent a structural break upon the introduction of the euro, we must allow for it.

The model allowing for structural break was constructed by assuming the same functional forms of the marginal distributions and copula, but allowing each parameter to change in the post-euro period. To minimise the number of additional parameters in the new models, we conducted tests for the significance of the change in the parameter, and imposed constancy on those parameters

[27]The eleven participating nations are: Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, The Netherlands, Portugal and Spain.

[28]The complete list of conversion rates for all eleven currencies now linked to the Euro may be found at http://www.ecb.int/press/pr981231_2.htm.

that were not significantly different in the two periods. We tested for dependence of the new transformed variables, denoted U_t^b and V_t^b, on the euro indicator variable by regressing each of the first four moments on a constant (to remove the mean) and the indicator variable. If the new model is correctly specified with respect to the new, larger information set, then the coefficient on the indicator variable in this regression should be zero.

The only parameter that was significantly different in the DM margin in the pre-euro and post-euro samples, according to LR tests, was the parameter for the unconditional mean return, μ_x. It turned out, however, that although the constant in the GARCH model for the DM margin was not significantly different from the first period to the second, it was needed in order for the indicator variable to have no explanatory power for U_t^b.[29] When allowing both μ_x and ω_x to change from the first period to the second U_t^b passed the regression tests of independence of the euro indicator variable. The LR test for the significance of the structural break yielded a test statistic (*p-value*) of 9.5354 (*0.0085*), indicating that the break is significant. It should be pointed out that we do not include monetary policy variables in our models. A change in interest rate policy in Germany following the introduction of the euro, for example, may explain at least a part of the change in DM dynamics. We leave the investigation of this as a possible explanation for future work.

We conducted tests for the significance of the changes in the parameters of the Yen margin between the pre- and post-euro periods, and found none to be significant. Further, the original transformed variable, V_t, passed the regression tests for independence of the euro indicator variable. The results of the estimation of the models allowing for a structural break are presented in Table 7 below. We have included the results for the Yen margin even though they did not change from those presented in Table 2 for the purposes of comparison with the new DM margin.

[INSERT TABLE 7 HERE.]

Table 7 reveals that the average rate of depreciation of the DM against the USD rose from 0.0128% per day during the pre-euro sample to 0.0981% per day, a substantial increase. The unconditional daily variance of the DM-USD exchange rate implied by the GARCH parameters rose over 40% from 0.5316 to 0.7674, while the annualised unconditional standard deviation[30] rose from 11.57% to 13.91%. We again used the LM and K-S tests to check the new proposed marginal distributions, in the pre-euro, post-euro and joint samples. Both margins passed all tests in all three time periods[31]. In Table 8 below, we present the results of the joint test that each margin is correctly specified in the five regions used in the previous tests. (Both marginal distributions passed all tests for the individual regions.) This table shows that both margins pass the joint test in all three samples.

[29] Specifically, the first moment of the transformed variable obtained from the distribution only allowing for a break in μ_x had a significant coefficient on the indicator variable in the simple regression test.

[30] This is approximated by taking the square root of the unconditional daily variance multiplied by 252, the average number of trading days in a year.

[31] The complete results are omitted due to space constraints, but are available from the author on request.

[INSERT TABLE 8 HERE.]

In Table 9 we present the results of allowing the parameters of the four copula models previously discussed to change following the introduction of the euro. The likelihood ratio test statistics (*p-values*) for the significance of the structural break for the four models are: 134.0548 (*0.0000*), 107.7615 (*0.0000*), 69.3381 (*0.0000*) and 59.2717 (*0.0000*) - clearly the structural break is very significant in all copula models.

The results for the two constant copulas clearly indicate that the dependence between the DM-USD and the Yen-USD exchange rate returns decreased substantially following the break. The constant normal copula results suggest that residual correlation fell from 0.4929 before the break to 0.0773 after the break. The constant Joe-Clayton copula parameters suggest that upper and lower tail dependence fell from 0.3964 and 0.2542 before the break to 0.0001 and 0.0018 after the break. The implied correlation from the constant Joe-Clayton copula fell from 0.4983 to 0.0781 - very similar results to those from the constant normal copula. As for the previous results, we can test the significance of the asymmetry implied by the constant Joe-Clayton copula, both before and after the break. The test statistic (*p-value*) for the pre-euro period is 3.1303 (*0.0017*), rejecting the null hypothesis of symmetry in the pre-euro period. In the post-euro period the test statistic (*p-value*) is -0.0269 (*0.9785*), indicating that no evidence of asymmetry is present in the post-euro data.

[INSERT TABLE 9 HERE.]

The time paths of the parameters and implied correlations in the models of the previous section suggested that the conditional dependence between these exchange rates changed following the introduction of the euro. By allowing the time paths to depend explicitly on which sub-period the data are from we obtain very strong evidence that the dependence changed. We initially allowed all parameters of the time-varying copulas to change following the break, but found that the results suggested that the conditional dependence was constant in the post-euro sample[32]. Specifically, the time path of the parameters of both the time-varying normal copula and the time-varying Joe-Clayton matched very closely the dependence implied by the constant version of the same copula in the post-euro sample. For both the normal and the Joe-Clayton copulas, then, we imposed the condition that dependence be constant following the break, though it was allowed to vary before the break. The implied correlation from the two copulas can be seen in Figures 8 and 9 below. The time paths of the parameters of the Joe-Clayton copula and the conditional tail dependence are presented in Figures 10 and 11 below.

[INSERT FIGURES 8 TO 11 HERE.]

[32] Again, we face the problem of an unidentified nuisance parameter in testing for the significance of time variation in the conditional dependence, and so we must rely on qualitative methods to assess its importance.

In Figure 11 we see that the conditional upper tail dependence is consistently greater than the conditional lower tail dependence in the pre-euro sample. In fact, on no day in the pre-euro sample was conditional lower tail dependence greater than conditional upper tail dependence. The mean difference between the two tail dependence measures was 0.1857, implying that these exchange rates were much more dependent during appreciations of the US dollar than they were during depreciations of the US dollar.

Figures 8 through 11 make it very clear that a large change in the dependence structure took place when the mark became linked to the euro. The change in the marginal distribution of the mark is difficult to detect graphically: the change in μ_x, while substantial, is drowned out by the large amount of noise in the returns, and the change in ω_x is not significant. Thus the structural break could possibly be overlooked if only considering the marginal distribution of the DM-USD exchange rate returns. The structural break in the conditional copula, however, is unmistakeable.

In Table 10 below we present a summary of the results of the hit tests on the copula models with an indicator variable for the euro. All models passed all tests for the individual regions, and as Table 10 below reports, all models passed the joint test that the copula is well specified for all eight regions.

[INSERT TABLE 10 HERE.]

A formal means of selecting which of the copulas is better given that all pass tests of goodness-of-fit is not currently available. As the time-varying Joe-Clayton copula has the highest log-likelihood, however, one could justify its selection as the model of choice on information theoretic grounds: a higher log-likelihood means that the time-varying Joe-Clayton copula is closer to the *true* copula than the next best alternative, as measured by the Kullback-Leibler information criterion[33]. It is acknowledged that the fact that the normal copulas pass the specification tests may indicate that the asymmetry implied by the Joe-Clayton copula is not significant, and thus that the normal copula is a good approximation. However, as the Joe-Clayton copula has symmetry as a special case, and symmetry appears to be strongly rejected, we instead infer that the passing of the specification tests by the normal copula models is evidence of the difficulty the tests considered have in rejecting models close to the true distribution.

5 Conclusion

The theory of copulas provides a means of thinking more generally about the dependence between random variables. The linear correlation coefficient provides a very convenient summary of the association between two variables, but it is by no means the only measure of interest. The use of copulas in the analysis of economic data is a quite recent phenomeon, however the growing interest

[33] See White (1994) for more on this interpretation of the log-likelihood.

in density forecasting and in quantile-based measures of risk suggest that the theory of copulas may enjoy wider use in the future.

In this paper we showed that the existing theory of copulas may be extended to the conditional case, allowing us to use it in the analysis of time-varying conditional dependence. We applied the theory of conditional copulas to model the time-varying conditional joint distribution of the daily Deutsche mark - U.S. dollar and Yen - U.S. dollar exchange rates, over the period from January 1991 to October 2000. AR - tGARCH models were employed for the marginal distributions of each exchange rate, and two different copulas were estimated: the copula associated with the bivariate normal distribution, and the Joe-Clayton copula, which allows for asymmetric dependence in the joint distribution. We allowed for time-variation in the dependence structure between the two exchange rates by allowing the parameters of the two copulas to evolve over the sample period, employing an evolution equation similar to the GARCH model for conditional variances. For comparison, we also estimated constant versions of each of these copulas.

Some attention was paid to tests of the relative goodness-of-fit of the copulas analysed. Goodness-of-fit testing in our study was complicated by the fact that we wished to test the adequacy of the entire density, rather than just a set of moments from this density, and by the fact that many of our models were non-nested. We employed an extension of the 'hit' tests of Christoffersen (1998) and Engle and Manganelli (1999) to test for the goodness-of-fit of the four models considered, and proposed a new test for evaluating the performance of multiple interval forecasts simultaneously.

Our results indicate the presence of time variation in the conditional dependence between these two exchange rates, though formally testing for its significance is complicated by the presence of a nuisance parameter unidentified under the null of constant dependence. We find substantial evidence that the dependence function is asymmetric; specifically, dependence is greater during appreciations of the U.S. dollar (or alternatively, during depreciations of the mark and the yen) than during depreciations of the U.S. dollar. Finally, we report strong evidence of a structural break in the conditional copula following the introduction of the euro in January 1999. The dependence between these exchange rates falls dramatically following the break.

This paper has presented just one example of an economic question that copula theory may assist us in answering. Many further applications or extensions are possible. For example, to estimate the Value-at-Risk of a portfolio, one needs a model for the entire joint density of the assets in the portfolio. Constructing such a model is made much simpler using the conditional copula framework. Further, copulas may be used to construct models for multivariate density forecasting, an area gaining interest in finance and econometrics. The use of conditional copulas in the more general multivariate case is possible, though some care may be required to keep the model tractable. Also, other forms of time variation in the dependence between two or more assets may be explored: in this paper we considered allowing the parameter of the copula to vary through time, holding the form of the copula fixed. An alternative to this may be to consider conditional copulas that vary in functional form, perhaps in a Markov switching model.

6 Bibliography

Akaike, H., 1973, Information Theory and an Extension of the Likelihood Ratio Principle, *Proceedings of the Second International Symposium of Information Theory*, B. N. Petrov and F. Caski, eds., 257-281.

Andersen, Torben G., and Bollerslev, Tim, 1998, Answering the Skeptics:Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review*, 39(4), 885-905.

Andersen, Torben G., Bollerslev, Tim, Diebold, Francis X., and Labys, Paul, 2001, The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association*, 96, 42-55.

Andrews, Donald W. K., and Ploberger, Werner, 1994, Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative, *Econometrica*, 62(6), 1383-1414.

Berkowitz, Jeremy, 1999, Evaluating the Forecasts of Risk Models, Working Paper, Federal Reserve Board, Washington D.C.

Bollerslev, Tim, 1987, A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *Review of Economics and Statistics*, 69, 542-547.

Bollerslev, Tim, 1990, Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach, *Review of Economics and Statistics*, 72, 498-505.

Bouyè, Eric, Durrleman, Valdo, Nikeghbali, Ashkan, Riboulet, Gael and Roncalli, Thierry, 2000, Copulas for Finance: A Reading Guide and Some Applications, Working Paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais. France.

Casella, George, and Berger, Roger L., 1990, *Statistical Inference*, Duxbury Press, U.S.A.

Christoffersen, Peter F., 1998, Evaluating Interval Forecasts, *International Economic Review*, 39(4), 841-862.

Clayton, D. G., 1978, A Model for Association in Bivariate Life Tables and its Application in Epidemiological Studies of Familial Tendeny in Chronic Disease Incidence, *Biometrika*, 65(1), 141-151.

Cook, R. Dennis, and Johnson, Mark E., 1981, A Family of Distributions for Modelling Non-elliptically Symmetric Multivariate Data, *Journal of the Royal Statistical Society*, 43(2), 210-218.

Costont, Arnaud, Roncalli, Thierry and Teïletche, Jérôme, 2000, Revisiting the Dependence between Financial Markets with Copulas, Work in progress, Groupe de Recherche Opérationnelle, Crédit Lyonnais, France.

D'Agostino, Ralph B., and Stephens, Michael A., eds., 1986, *Goodness-of-Fit Techniques*, Marcel Dekker Inc., New York.

David, F. N., and Johnson, N. L., 1948, The Probability Integral Transformation when Parameters are Estimated from the Sample, *Biometrika*, 35(1), 182-190.

Dawid, A. P., 1984, Statistical Theory: The Prequential Approach, *Journal of the Royal Statistical Society*, 92, 147-278.

Davidson, Russell, and MacKinnon, James G., 1993, *Estimation and Inference in Econometrics*, Oxford University Press, New York.

Deheuvels, P., 1978, La fonction de dépendence empirique et ses propriétés. Un test non paramétrique d'indépendence, *Acad. Roy. Belg. Bull. Cl. Sci.*, 65(5), 274-292.

Diebold, Francis X., Gunther, T., and Tay, Anthony S., 1998, Evaluating Density Forecasts, with Applications to Financial Risk Management, *International Economic Review*, 39, 863-883.

Diebold, Francis X., Hahn, Jinyong, and Tay, Anthony S., 1999, Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns on Foreign Exchange, *The Review of Economics and Statistics*, 81(4), 661-673.

Embrechts, Paul, McNeil, Alexander and Straumann, Daniel, 1999, Correlation: Pitfalls and Alternatives, *RISK*, 12(5), 11-21.

Embrechts, Paul, McNeil, Alexander and Straumann, Daniel, 2000, Correlation and Dependence Properties in Risk Management: Properties and Pitfalls, Working Paper, Departement Mathematik, ETHZ, Zürich, Switzerland.

Engle, Robert F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation, *Econometrica,* 50, 987-1007.

Engle, Robert F., Ito, Takatoshi, and Lin, Wen-Ling, 1990, Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market*, Econometrica,* 58(3), 525-542.

Engle, Robert F., and Kroner, Kenneth F., 1995, Multivariate Simultaneous Generalized ARCH, *Econometric Theory*, 11, 122-150.

Engle, Robert F., and Manganelli, Simone, 1999, CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles, Working Paper, Department of Economics, University of California, San Diego.

Fisher, R. A., 1932, *Statistical Methods for Research Workers.*

Forbes, Kristin, and Rigobon, Roberto, 1999, No Contagion, Only Interdependence: Measuring Stock Market Comovements, *NBER* Working Paper 7267.

Galambos, J., 1978, *The Asymptotic Theory of Extreme Order Statistics*, John Wiley & Sons, New York.

Genest, Christian, and Rivest, Louis-Paul, 1993, Statistical Inference Procedures for Bivariate Archimedean Copulas, *Journal of the American Statistical Association*, 88(423), 1034-1043.

Greene, William H., 1997, *Econometric Analysis*, Prentice-Hall, New Jersey.

Hull, John, and White, Alan, 1998, Value at Risk When Daily Changes in Market Variables are not Normally Distributed, *Journal of Derivatives*, 5(3), 9-19.

Hutchinson, T. P. and Lai, C. D., 1990, *Continuous Bivariate Distributions, Emphasising Applications*, Rumsby, Sydney.

Joe, Harry, 1993, A Parametric Family of Multivariate Distributions with Given Margins, *Journal of Multivariate Analysis*, 46, 262-282.

Joe, Harry, 1997, *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.

Judd, Kenneth L., 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, Massachusetts.

Kearney, Colm, and Patton, Andrew J., 2000, Multivariate GARCH Modeling of Exchange Rate Volatility Transmission in the European Monetary System, *The Financial Review*, 41, 29-48.

Kimeldorf, G. and Sampson, A. R., 1975, Uniform representations of bivariate distributions, *Communications in Statistics*, 4, 617-627.

Kling, J. L., and Bessler, D. A., 1989, Calibration-based Predictive Distributions: An Application of Prequential Analysis to Interest Rates, Money, Prices and Output, *Journal of Business*, 62, 477-499.

Melvin, Michael, 2000, *International Money and Finance*, 6th Ed., Addison-Wesley Inc., New York.

Nelsen, Roger B., 1999, *An Introduction to Copulas*, Springer-Verlag, New York.

Neyman, J., 1937, Smooth Tests for Goodness of Fit, *Skand. Aktuariedtidskrift*, 20, 149-199.

Oakes, David, 1989, Bivariate Survival Models Induced by Frailties, *Journal of the American Statistical Association*, 84(406), 487-493.

Patton, Andrew J., 2001, Estimation of Copula Models for Time Series, manuscript in progress.

Pearson, K., 1900, On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be Reasonably Supposed to have Arisen from Random Sampling, *Philosphical Magazine Series*, 50, 157-175.

Pearson, K., 1933, On a Method of Determining Whether a Sample of Size n Supposed to Have Been Drawn from a Parent Population Having a Known Probability Integral has Probably Been Drawn at Random, *Biometrika*, 25, 3/4, 379-410.

Pearson, E. S., 1938, The Probability Integral Transformation for Testing Goodness of Fit and Combining Independent Tests of Significance, *Biometrika*, 30, 1/2, 134-148.

Press, William H., Flannery, Brian P., Teukolsky, Saul A., and Vetterling, William T., 1989, *Numerical Recipes*, Cambridge University Press, New York.

Rosenberg, Joshua V., 2000, Nonparametric Pricing of Multivariate Contingent Claims, Working Paper, Stern School of Business, New York University.

Schweizer, B., and Sklar, A., 1983, *Probabilistic Metric Spaces*, Elsevier Science, New York.

Shao, Jun, 1999, *Mathematical Statistics,* Springer-Verlag, New York.

Sin, Chor-yiu, and Granger, Clive W. J., 1996, Estimating and Forecasting Quantiles with Asymmetric Least Squares, Working Paper, Department of Economics and Finance, City University of Hong Kong.

Sklar, A., 1959, Fonctions de répartition à n dimensionset leurs marges, *Publ. Inst. Statis. Univ. Paris*, 8, 229-231.

Stephens, Michael A., 1986, Tests Based on EDF Statistics, in D'Agostino, Ralph B., and Stephens, Michael A., eds. *Goodness-of-Fit Techniques*, Marcel Dekker Inc., New York.

White, Halbert, 1994, *Estimation, Inference and Specification Analysis*, Econometric Society Monographs No. 22, Cambridge University Press, New York.

7 Appendix A: Proofs (draft).

The proof of Sklar's (1959) theorem is simplified with the following two lemmas.

Lemma 7 *Let* $x_1 \le x_2$ *and* $y_1 \le y_2$*, then*

$$\gamma^x(x) \equiv H(x, y_2|\mathcal{F}) - H(x, y_1|\mathcal{F}) \text{ is a non-decreasing function of } x\text{, and}$$
$$\gamma^y(y) \equiv H(x_2, y|\mathcal{F}) - H(x_1, y|\mathcal{F}) \text{ is a non-decreasing function of } y\text{.}$$

Proof of Lemma 7. From the definition of a bivariate distribution function we know that $H(x_2, y_2|\mathcal{F}) - H(x_2, y_1|\mathcal{F}) - H(x_1, y_2|\mathcal{F}) + H(x_1, y_1|\mathcal{F}) \ge 0$, so $H(x_2, y_2|\mathcal{F}) - H(x_2, y_1|\mathcal{F}) \ge H(x_1, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})$ for all $x_1 \le x_2, y_1 \le y_2 \in R$. Similarly for $\gamma^y(y)$. ■

Lemma 8 $|H(x_2, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})| \le |F(x_2|\mathcal{F}) - F(x_1|\mathcal{F})| + |G(y_2|\mathcal{F}) - G(y_1|\mathcal{F})|$

Proof of Lemma 8. By the triangle inequality we have:

$$|H(x_2, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})| \le |H(x_2, y_2|\mathcal{F}) - H(x_1, y_2|\mathcal{F})| + |H(x_1, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})|$$

Assume $x_1 \le x_2$ and $y_1 \le y_2$, then by the above lemma we have

$$\begin{aligned} H(x_2, y_1|\mathcal{F}) - H(x_1, y_1|\mathcal{F}) &\le H(x_2, y_2|\mathcal{F}) - H(x_1, y_2|\mathcal{F}) \\ &\le H(x_2, \infty|\mathcal{F}) - H(x_1, \infty|\mathcal{F}) \\ &= F(x_2|\mathcal{F}) - F(x_1|\mathcal{F}) \end{aligned}$$

Considering the case when $x_1 \ge x_2$ and applying the same logic leads us to

$$H(x_1, y_1|\mathcal{F}) - H(x_2, y_1|\mathcal{F}) \le F(x_1|\mathcal{F}) - F(x_2|\mathcal{F})$$

So we have

$$|H(x_2, y_1|\mathcal{F}) - H(x_1, y_1|\mathcal{F})| \le |F(x_2|\mathcal{F}) - F(x_1|\mathcal{F})|$$

Similarly for y we find that

$$|H(x_1, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})| \le |G(y_2|\mathcal{F}) - G(y_1|\mathcal{F})|$$

and so

$$\begin{aligned} |H(x_2, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})| &\le |H(x_2, y_2|\mathcal{F}) - H(x_1, y_2|\mathcal{F})| + |H(x_1, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})| \\ &\le |F(x_2|\mathcal{F}) - F(x_1|\mathcal{F})| + |G(y_2|\mathcal{F}) - G(y_1|\mathcal{F})| \quad \blacksquare \end{aligned}$$

Proof of Theorem 3. From Lemma 8 we know that: $|H(x_2, y_2|\mathcal{F}) - H(x_1, y_1|\mathcal{F})| \le |F(x_2|\mathcal{F}) - F(x_1|\mathcal{F})| + |G(y_2|\mathcal{F}) - G(y_1|\mathcal{F})|$. Thus, if $x_2 = x_1$ and $y_2 = y_1$, then $H(x_2, y_2|\mathcal{F}) = H(x_1, y_1|\mathcal{F})$. The function C is defined by the set of ordered pairs:

$\{((F(x|\mathcal{F}), G(y|\mathcal{F}))\,,\ H(x, y|\mathcal{F}))\,|x, y \in \bar{R}\}$.

That C is a copula must be verified: the domain of C is clearly $[0, 1] \times [0, 1]$, as this is the range of F and G. The range of C is similarly determined to be $[0, 1]$ as this is the range of H. We now check the two conditions for C to be a copula, as given in Definition 2.

[1] Since F and G are continuous, we know that the inverse functions F^{-1} and G^{-1} are well-defined. So,

$$C(u, 0|\mathcal{F}) = H(F^{-1}(u|\mathcal{F}), G^{-1}(0|\mathcal{F})|\mathcal{F})$$

$$
\begin{aligned}
&= H(F^{-1}(u|\mathcal{F}), -\infty|\mathcal{F}) \\
&= 0, \text{ by the definition of a distribution function, and,} \\
C(0, v|\mathcal{F}) &= H(F^{-1}(0|\mathcal{F}), G^{-1}(v|\mathcal{F})|\mathcal{F}) \\
&= H(-\infty, G^{-1}(v|\mathcal{F})|\mathcal{F}) \\
&= 0, \text{ also by the definition of a distribution function.}
\end{aligned}
$$

Further,

$$
\begin{aligned}
C(u, 1|\mathcal{F}) &= H(F^{-1}(u|\mathcal{F}), G^{-1}(1|\mathcal{F})|\mathcal{F}) \\
&= H(F^{-1}(u|\mathcal{F}), \infty|\mathcal{F}) \\
&= F(F^{-1}(u|\mathcal{F})|\mathcal{F}) \\
&= u
\end{aligned}
$$

and

$$
\begin{aligned}
C(1, v|\mathcal{F}) &= H(F^{-1}(1|\mathcal{F}), G^{-1}(v|\mathcal{F})|\mathcal{F}) \\
&= H(\infty, G^{-1}(v|\mathcal{F})|\mathcal{F}) \\
&= G(G^{-1}(v|\mathcal{F})|\mathcal{F}) \\
&= v
\end{aligned}
$$

[2] For this part, let $u_i = F\left(x_i|\mathcal{F}\right)$, $v_i = G\left(y_i|\mathcal{F}\right)$, and consider the points $x_1, x_2, y_1, y_2 \in R$ s.t. $x_1 \leq x_2$ and $y_1 \leq y_2$. Then,

$$
\begin{aligned}
V_C\left([u_1, u_2] \times [v_1, v_2]\right) &= C(u_2, v_2|\mathcal{F}) - C(u_1, v_2|\mathcal{F}) - C(u_2, v_1|\mathcal{F}) + C(u_1, v_1|\mathcal{F}) \\
&= H\left(x_2, y_2|\mathcal{F}\right) - H\left(x_1, y_2|\mathcal{F}\right) - H\left(x_2, y_1|\mathcal{F}\right) + H\left(x_1, y_1|\mathcal{F}\right) \\
&= V_H\left([x_1, x_2] \times [y_1, y_2]\right) \\
&\geq 0, \text{ by the fact that } H \text{ is a conditional distribution function.}
\end{aligned}
$$

Thus the function C defined above is a conditional copula.

The proof of the converse requires us to verify the conditions that make H a distribution function, given F and G are distribution functions, and C is a copula.

$$
\begin{aligned}
H(x, -\infty|\mathcal{F}) &= C(F(x|\mathcal{F}), G(-\infty|\mathcal{F})|\mathcal{F}) \\
&= C(F(x|\mathcal{F}), 0|\mathcal{F}) \\
&= 0
\end{aligned}
$$

and

$$
\begin{aligned}
H(-\infty, y|\mathcal{F}) &= C(F(-\infty|\mathcal{F}), G(y|\mathcal{F})|\mathcal{F}) \\
&= C(0, G(y|\mathcal{F})|\mathcal{F}) \\
&= 0
\end{aligned}
$$

Further,

$$
\begin{aligned}
H(x, \infty|\mathcal{F}) &= C(F(x|\mathcal{F}), G(\infty|\mathcal{F})|\mathcal{F}) \\
&= C(F(x|\mathcal{F}), 1|\mathcal{F}) \\
&= F(x|\mathcal{F})
\end{aligned}
$$

and

$$
\begin{aligned}
H(\infty, y|\mathcal{F}) &= C(F(\infty|\mathcal{F}), G(y|\mathcal{F})|\mathcal{F}) \\
&= C(1, G(y|\mathcal{F})|\mathcal{F})
\end{aligned}
$$

$= G(y|\mathcal{F})$

We now need to show that $V_H \geq 0$.

$$\begin{aligned} V_H([x_1, x_2] \times [y_1, y_2]) &= H(x_2, y_2|\mathcal{F}) - H(x1, y_2|\mathcal{F}) - H(x_2, y_1|\mathcal{F}) + H(x_1, y_1|\mathcal{F}) \\ &= C(F(x_2|\mathcal{F}), G(y_2|\mathcal{F})|\mathcal{F}) - C(F(x_1|\mathcal{F}), G(y_2|\mathcal{F})|\mathcal{F}) \\ &\quad - C(F(x_2|\mathcal{F}), G(y_1|\mathcal{F})|\mathcal{F}) + C(F(x_1|\mathcal{F}), G(y_1|\mathcal{F})|\mathcal{F}) \\ &= C(u_2, v_2|\mathcal{F}) - C(u_1, v_2|\mathcal{F}) - C(u_2, v_1|\mathcal{F}) + C(u_1, v_1|\mathcal{F}) \\ &= V_C([u_1, u_2] \times [v_1, v_2]) \\ &\geq 0 \text{ by the fact that } C \text{ is a conditional copula.} \end{aligned}$$

This completes the proof of the converse. ■

Proof of Corollary 5. This proof follows directly from that of Theorem 3, letting $x \equiv F^{(-1)}(u|\mathcal{F})$ and $y \equiv G^{(-1)}(v|\mathcal{F})$, and noting that $u = F\left(F^{(-1)}(u|\mathcal{F})|\mathcal{F}\right)$ and $v = G\left(G^{(-1)}(v|\mathcal{F})|\mathcal{F}\right)$ $\forall u, v \in [0,1]$. ■

8 Tables and Figures

Table 1: *Summary Statistics*

	DM-USD	Yen-USD
Mean	0.0166	−0.0091
Std. Deviation	0.6682	0.7320
Skewness	−0.1353	−0.6465
Kurtosis	5.0041	8.4321
Jarque-Bera stat	428.22	3264.81
Jarque-Bera p-val	0.0000	0.0000
Median	0.0331	0.0078
Maximum	3.1874	3.5711
Minimum	−3.3195	−6.5818
ARCH LM stat	8.5868	12.9933
ARCH LM p-val	0.0000	0.0000

Note: This table presents some summary statistics of the data used in this paper. The data are 100 times the log-differences of the daily Deutsche mark - U.S. dollar and Japanese yen - U.S. dollar exchange rates. The sample period runs from January 1991 to October 2000, yielding 2513 observations.

Table 2: *Results for the Marginal Distributions*

DM Margin			Yen Margin		
	Coeff	**Std Err**		**Coeff**	**Std Err**
μ_x	0.0276*	0.0111	μ_y	0.0144	0.0111
ϕ_{1x}	0.0142	0.0200	ϕ_{1y}	0.0043	0.0195
			ϕ_{10y}	0.0664*	0.0183
ω_x	0.0039	0.0030	ω_y	0.0059	0.0034
β_x	0.9485*	0.0161	β_y	0.9453*	0.0161
α_x	0.0448*	0.0126	α_y	0.0458*	0.0125
ν_x	5.8073*	0.6383	ν_y	4.3817*	0.3800

$\mathcal{LL}_X = -2,379.4945$, $\mathcal{LL}_Y = -2,469.7595$

*An ' * ' indicates that the parameter is significant at the 5% level.*

Note: An asterix indicates that the parameter is significant at the 5% level.

.

Table 3: *LM Tests of independence and Kolmogorov-Smirnov Tests of the Density*		
	$(u_t - \bar{u})$	$(v_t - \bar{v})$
First moment	34.8434	30.9334
p-value	0.7131	0.8553
Second moment	33.2698	36.4737
p-value	0.7757	0.6429
Third moment	33.5627	41.0963
p-value	0.7645	0.4366
Fourth moment	33.8072	42.9562
p-value	0.7550	0.3595
K-S Stat	0.0119	0.0112
K-S p-value	0.8796	0.9114

Note: This table presents the results of LM tests of the independence of the first four moments of the variables U_t and V_t, described in the text. We regress $(u_t - \bar{u})^k$ and $(v_t - \bar{v})^k$ on twenty lags of both variables, for $k = 1, 2, 3, 4$. The test statistic is $(T - 40) \cdot R^2$ for each regression, and is distributed under the null as χ^2_{40}.

Table 5: *Results for the Copula Models*				
		Coeff	**Std Err**	$\mathcal{CL}$
Constant Normal	ρ	0.4560*	0.0167	*291.9811*
Constant	κ	1.3356*	0.0348	*296.8704*
Joe-Clayton	γ	0.4202*	0.0384	
Time-Varying	ω_ρ	0.0015	0.0052	*334.5621*
Normal	α_ρ	0.1212*	0.0160	
	β_ρ	2.0684*	0.0250	
Time-Varying	ω_U	−2.0621*	0.2056	*345.3665*
Joe-Clayton	α_U	−0.9192	0.8966	
	β_U	4.4548*	0.2938	
	ω_L	−1.3444*	0.5876	
	α_L	−6.5119*	3.1394	
	β_L	4.1406*	0.5880	

Note: An asterix indicates that the parameter is significant at the 5% level. '$\mathcal{CL}$' stands for the copula likelihood at the optimum.

.

Table 4: *Hit test results for the marginal distributions*

	Normal BEKK DM	Normal BEKK Yen	t BEKK DM	t BEKK Yen	Copula DM	Copula Yen
Test stat 1	5.1095	14.5768	8.8698	9.9004	7.3911	10.0053
p-value 1	*0.2762*	*0.0057*	*0.0644*	*0.0421*	*0.1166*	*0.0403*
Test stat 2	30.0152	16.5288	7.2563	7.8860	4.3805	5.4210
p-value 2	*0.0000*	*0.0024*	*0.1229*	*0.0958*	*0.3570*	*0.2468*
Test stat 3	53.5698	79.2013	6.7396	4.5611	3.7318	2.7020
p-value 3	*0.0000*	*0.0000*	*0.1503*	*0.3354*	*0.4435*	*0.6089*
Test stat 4	2.6915	5.2890	1.1491	1.8842	1.5011	3.7060
p-value 4	*0.6107*	*0.2589*	*0.8864*	*0.7570*	*0.8264*	*0.4473*
Test stat 5	18.9513	19.9437	6.6647	5.1809	2.8474	3.8913
p-value 5	*0.0008*	*0.0005*	*0.1547*	*0.2692*	*0.5837*	*0.4209*
Test stat ALL	73.8875	86.9531	22.4174	26.3902	15.0387	24.6226
p-value ALL	*0.0000*	*0.0000*	*0.1302*	*0.0488*	*0.5218*	*0.0768*

Note: 'Test stat' refers to the likelihood ratio statistic testing the null hypothesis that the model is correctly specified. 'P-value' refers to the area in the right tail of the distribution of the test statistic, a χ^2_4 random variable for the individual region tests and a χ^2_{16} random variable for the joint test.The numbers 1 through 7 refer to the regions of the marginal distribution support described in the text. 'ALL' refers to the joint test of all regions simultaneously.

.

Table 6: *Hit test results for the copula models*

	Constant Normal	Constant Joe-Clayton	Time-varying Normal	Time-varying Joe-Clayton
Test stat 1	9.4482	7.3304	7.4209	6.6815
p-value 1	*0.0508*	*0.1194*	*0.1152*	*0.1537*
Test stat 2	4.1361	9.8002	2.2563	16.0750
p-value 2	*0.3879*	*0.0439*	*0.6887*	*0.0029*
Test stat 3	6.4374	6.1412	6.4108	6.5137
p-value 3	*0.1688*	*0.1888*	*0.1705*	*0.1639*
Test stat 4	1.9063	1.8627	1.4813	1.4595
p-value 4	*0.7530*	*0.7610*	*0.8300*	*0.8338*
Test stat 5	8.2495	6.7740	7.9588	3.3560
p-value 5	*0.0829*	*0.1483*	*0.0931*	*0.5001*
Test stat 6	2.8800	6.9902	0.9284	3.1363
p-value 6	*0.5781*	*0.1364*	*0.9205*	*0.5353*
Test stat 7	6.5945	8.4697	0.6086	3.2932
p-value 7	*0.1589*	*0.0758*	*0.9621*	*0.5100*
Test stat ALL	37.8357	46.1668	26.9924	41.4195
p-value ALL	*0.1015*	*0.0167*	*0.5187*	*0.0491*

Note: 'Test stat' refers to the likelihood ratio statistic testing the null hypothesis that the model is correctly specified. 'P-value' refers to the area in the right tail of the distribution of the test statistic, a χ^2_4 random variable for the individual region tests and a χ^2_{28} random variable for the joint test.The numbers 1 through 7 refer to the regions of the copula support depicted in Figure 3. 'ALL' refers to the joint test of all regions simultaneously.

.

Table 7: *Results for the Marginal Distributions Pre- and Post-Euro periods*

	DM Margin			Yen Margin	
	Coeff	**Std Err**		**Coeff**	**Std Err**
μ_x^1	0.0128	0.0120	μ_y	0.0144	0.0111
μ_x^2	0.0982*	0.0268			
ϕ_{1x}	0.0111	0.0200	ϕ_{1y}	0.0043	0.0195
			ϕ_{10y}	0.0664*	0.0183
ω_x^1	0.0037	0.0029	ω_y	0.0059	0.0034
ω_x^2	0.0053	0.0038			
β_x	0.9485*	0.0160	β_y	0.9453*	0.0161
α_x	0.0446*	0.0125	α_y	0.0458*	0.0125
ν_x	5.6860*	0.6180	ν_y	4.3817*	0.3800
$\mathcal{LL}_X = -2,374.7268,$			$\mathcal{LL}_Y = -2,469.7595$		

Note: An asterix indicates that the parameter is significant at the 5% level. '$\mathcal{CL}$' stands for the copula likelihood at the optimum. The superscripts on the parameters refer to the period before or after January 1, 1999. Note that the Yen margin did not change between these two periods.

Table 8: *Multinomial test results for the marginal distributions,with structural break*

	Full Sample		Pre-Euro		Post-Euro	
	DM	Yen	DM	Yen	DM	Yen
Test stat ALL	14.2945	35.0488	16.2624	25.8905	13.1024	21.4322
p-value ALL	*0.5768*	*0.1685*	*0.4348*	*0.0556*	*0.6653*	*0.1625*

Note: 'Test stat' refers to the likelihood ratio statistic testing the null hypothesis that the model is correctly specified. 'P-value' refers to the area in the right tail of the distribution of the test statistic, a χ^2_{16} random variable, for the joint test. 'ALL' refers to the joint test of all regions simultaneously.

.

Table 9: *Results for the Copula Models Pre- and Post-Euro Periods*

		Coeff	**Std Err**	$\mathcal{CL}$
Constant Normal	ρ^1	0.5435*	0.0146	*359.0085*
	ρ^2	0.0855	0.0508	
Constant	κ^1	1.4678*	0.0428	*350.7512*
Joe-Clayton	γ^1	0.5061	0.0451	
	κ^2	1.0001	0.0442	
	γ^2	0.1101	0.0580	
Time-Varying	ω_ρ^1	−0.1834*	0.0038	*369.2312*
Normal	α_ρ^1	0.0483*	0.0159	
	β_ρ^1	2.5410*	0.0121	
	$\bar{\rho}^2$	0.0866	0.0492	
Time-Varying	ω_U^1	−2.1136*	0.0688	*374.5024*
Joe-Clayton	α_U^1	−0.5691*	0.2211	
	β_U^1	4.4315*	0.0788	
	ω_L^1	−1.7832*	0.3468	
	α_L^1	−2.9472	1.9843	
	β_L^1	4.2067*	0.2631	
	$\bar{\kappa}^2$	1.0001	0.0519	
	$\bar{\gamma}^2$	0.1098	0.0575	

Note: An asterix indicates that the parameter is significant at the 5% level. '$\mathcal{CL}$' stands for the copula likelihood at the optimum. The superscripts on the parameters refer to the period before or after January 1, 1999.

Table 10: *Multinomial test results for the copula models,with structural break*

	Full Sample	**Pre-Euro**	**Post-Euro**
Constant	24.9467	28.3683	32.3445
Normal	*0.6307*	*0.4451*	*0.2607*
Constant	36.0588	38.3891	30.0025
Joe-Clayton	*0.1411*	*0.0913*	*0.3631*
Time-varying	24.3990	25.0714	32.3137
Normal	*0.6603*	*0.6239*	*0.2619*
Time-varying	34.9416	33.7263	30.4127
Joe-Clayton	*0.1716*	*0.2100*	*0.3438*

Note: Above we report the test statistic and *p-value* for the multinomial test for the goodness-of-fit of each copula model.

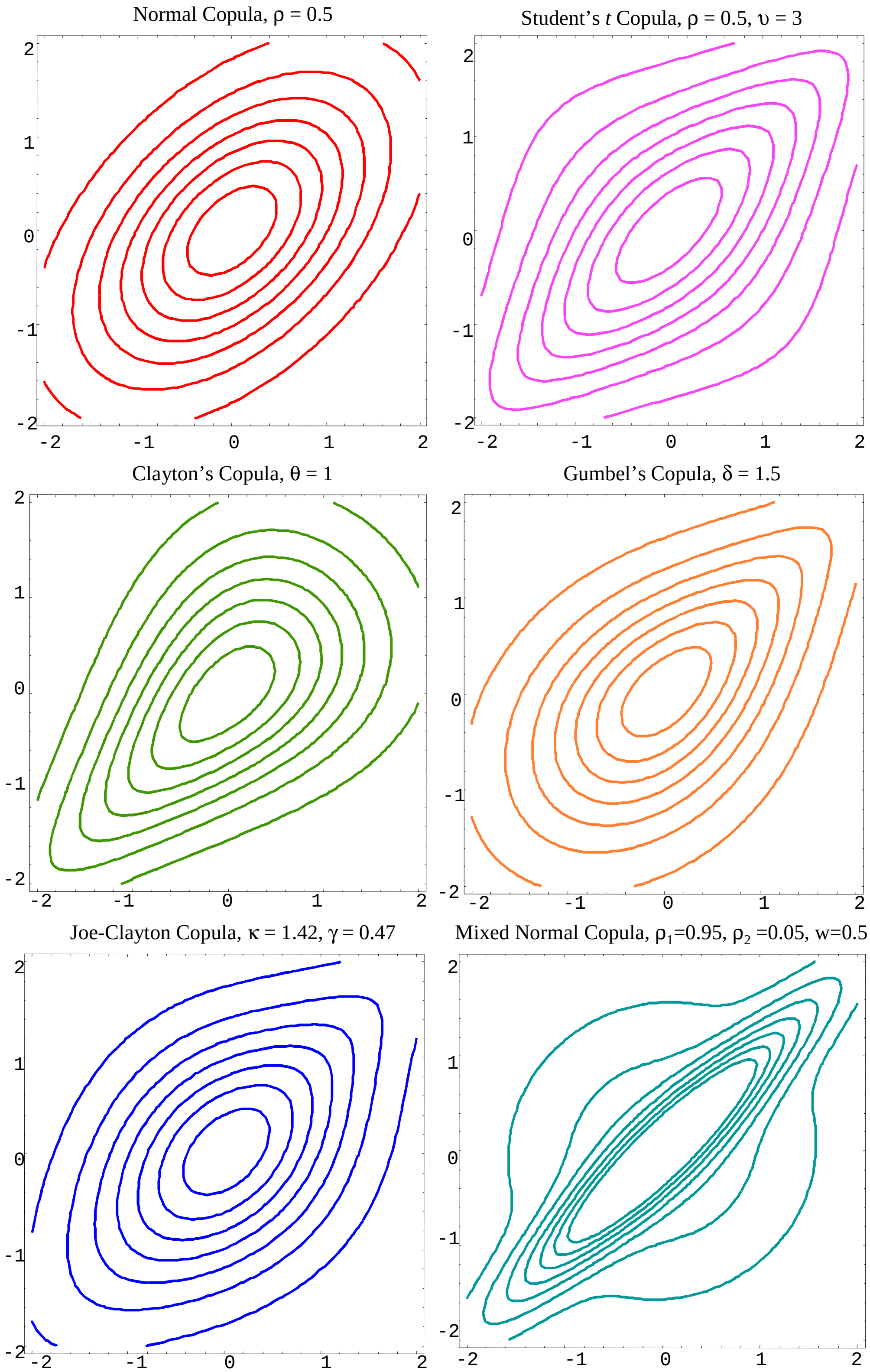

Figure 1: *Contour plots of various distributions all with standard normal marginal distributions and linear correlation coefficients of 0.5.*

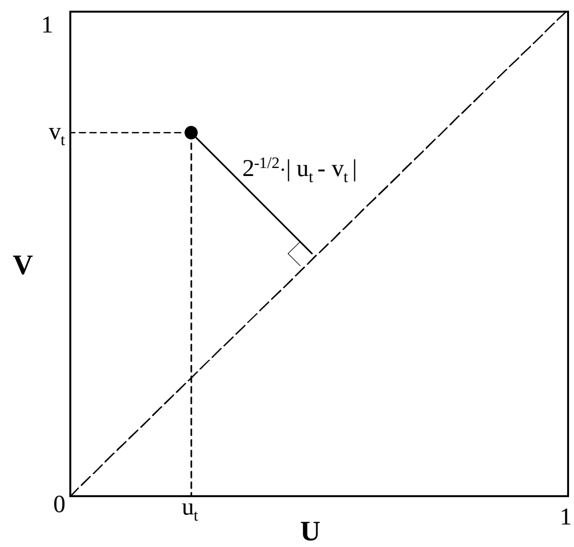

Figure 2: *Motivation for the choice of forcing variable in the specification of the time-varying Joe-Clayton copula.*

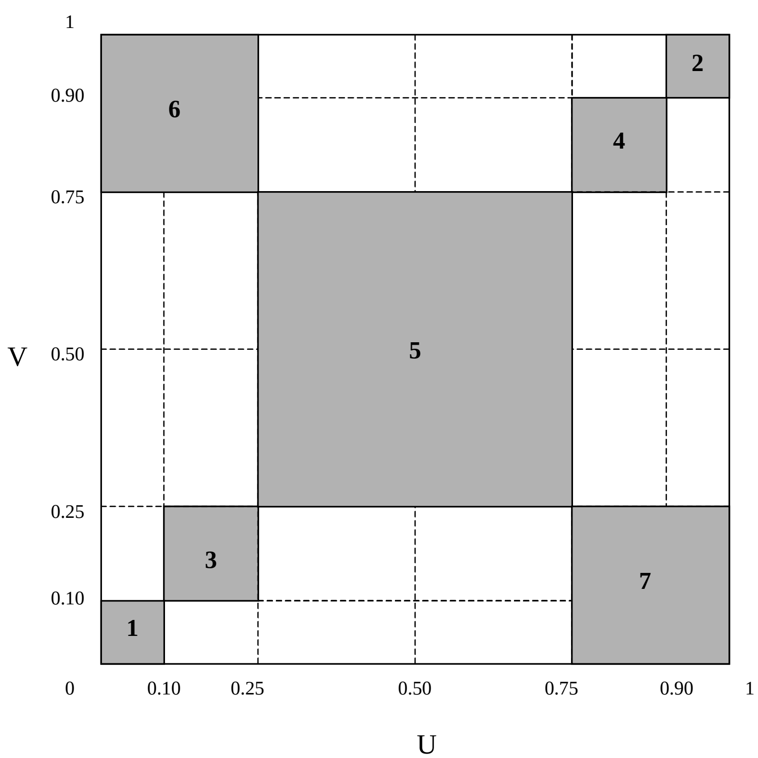

Figure 3: *Regions used in the hit tests*

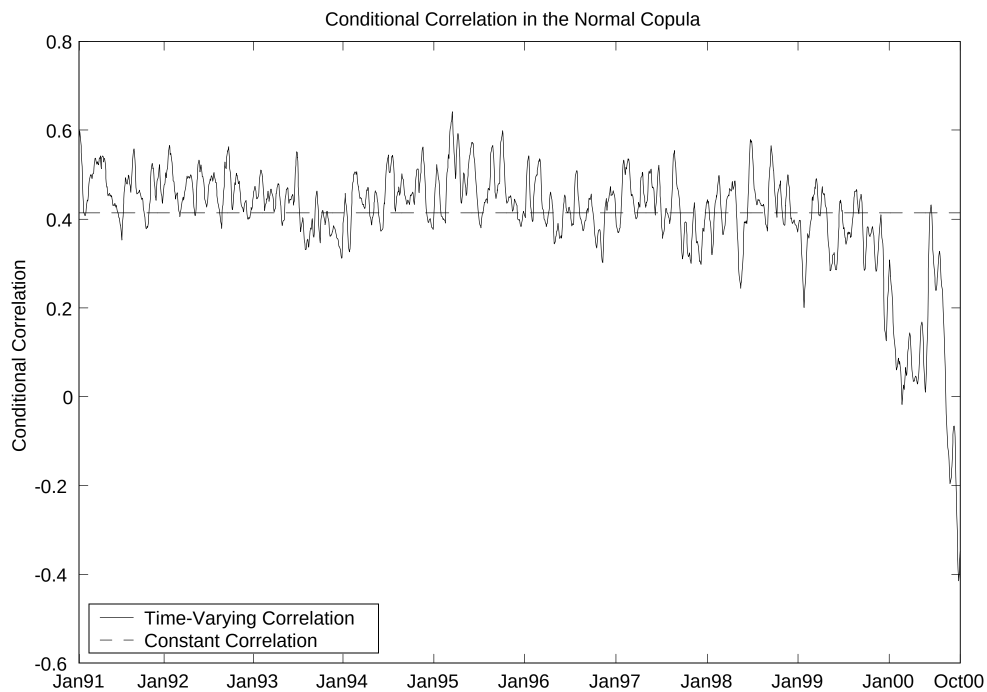

Figure 4: *Time-varying correlation estimates from the two Normal copulas.*

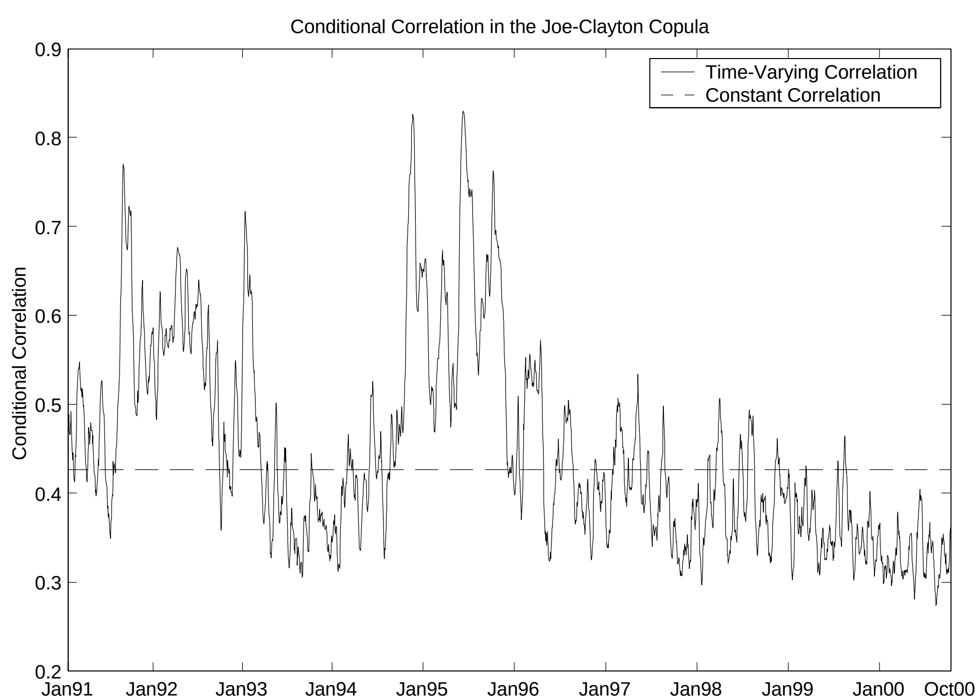

Figure 5: *Time-varying correlation estimates from the two Joe-Clayton copulas.*

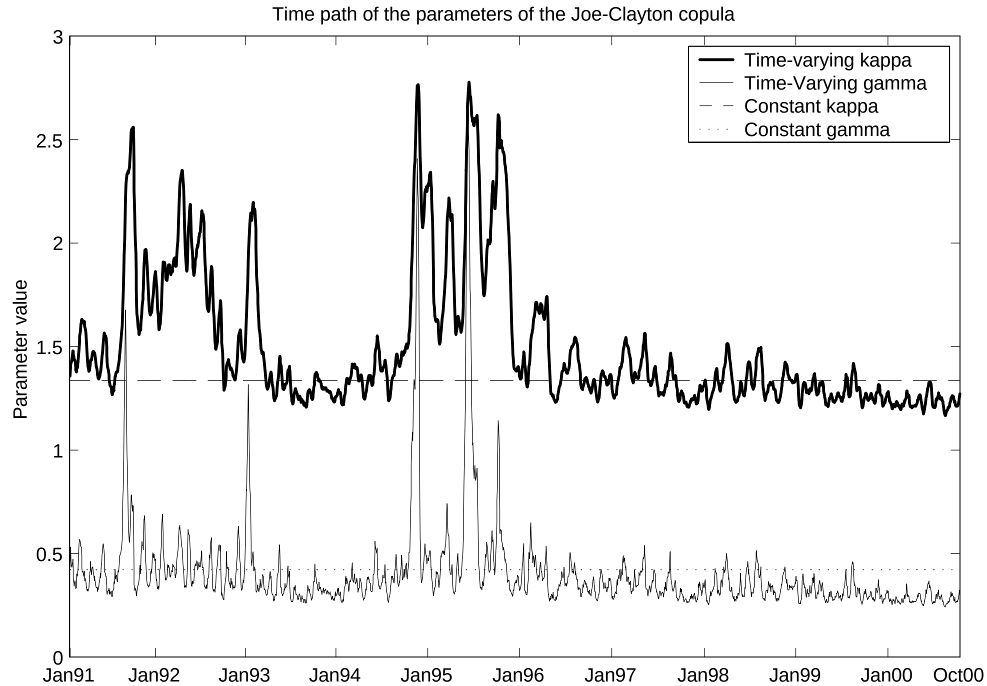

Figure 6: *Time-varying parameters of the Joe-Clayton copulas.*

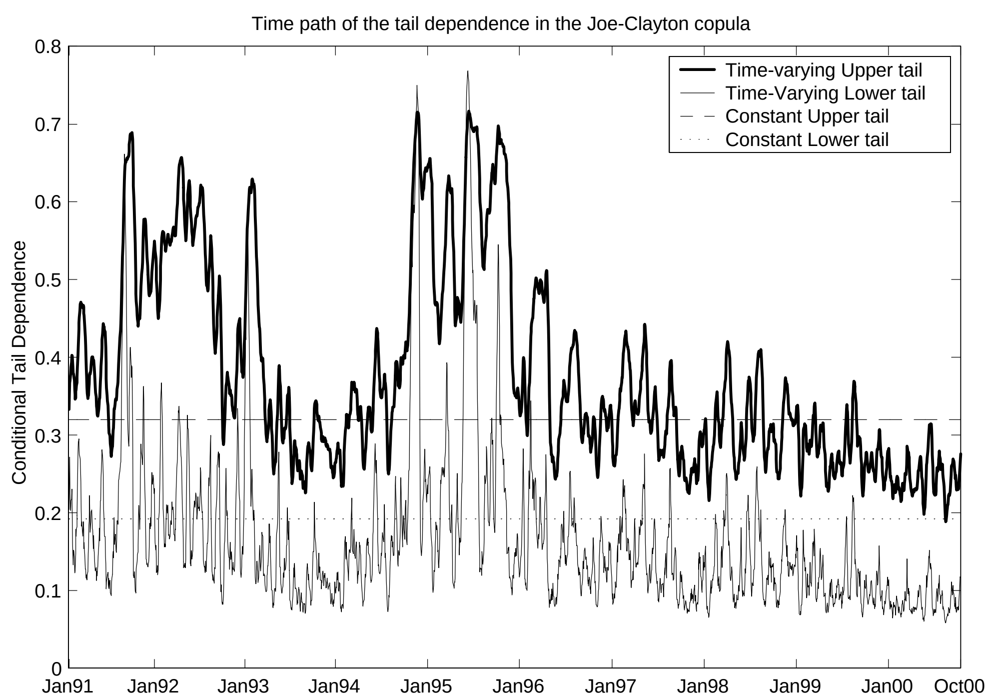

Figure 7: *Time-varying tail dependence in the Joe-Clayton copulas.*

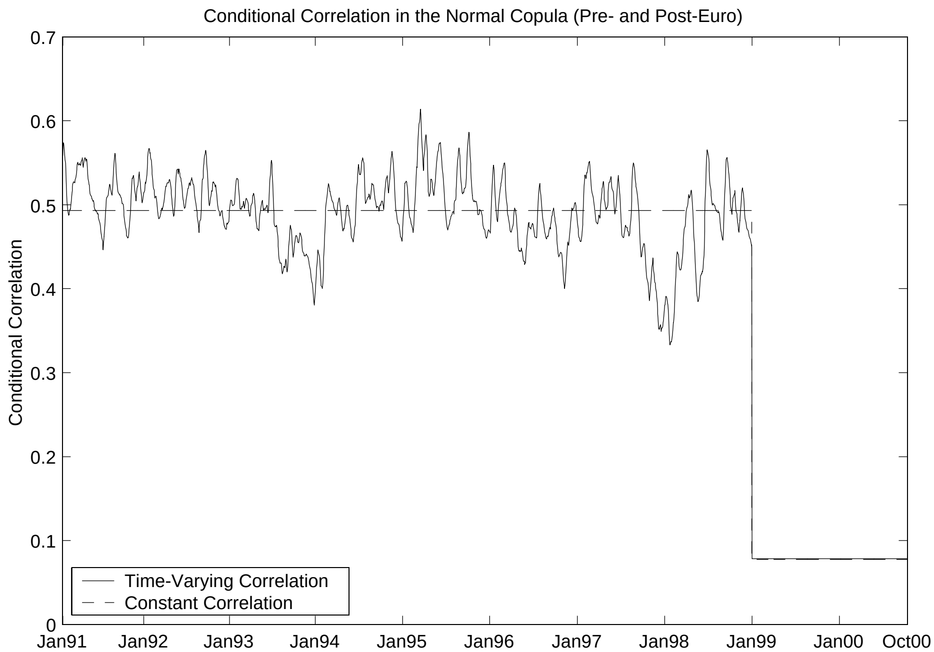

Figure 8: *Time-varying correlation estimates from the Normal copulas allowing for a structural break at the introduction of the Euro on Jan 1, 1999.*

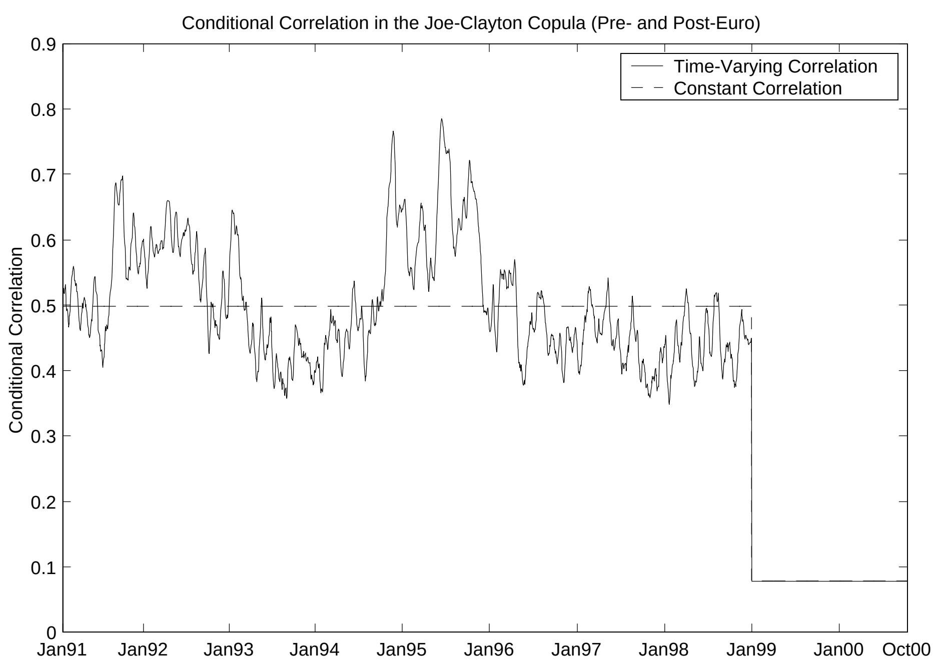

Figure 9: *Time-varying correlation estimates from the Joe-Clayton copulas allowing for a structural break at the introduction of the Euro on Jan 1, 1999.*

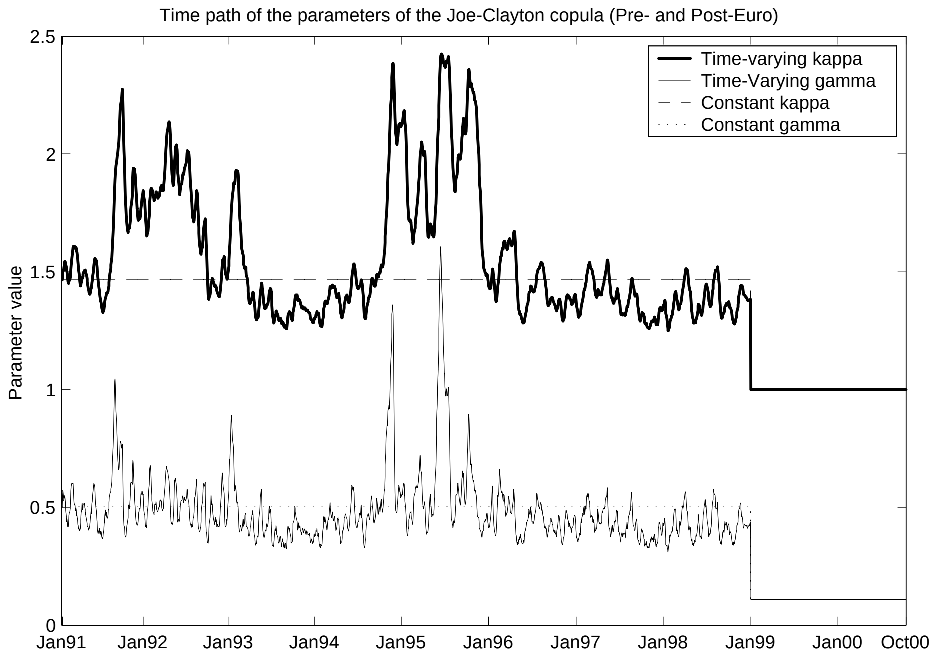

Figure 10: *Time-varying parameters of the Joe-Clayton copulas allowing for a structural break at the introduction of the Euro on Jan 1, 1999.*

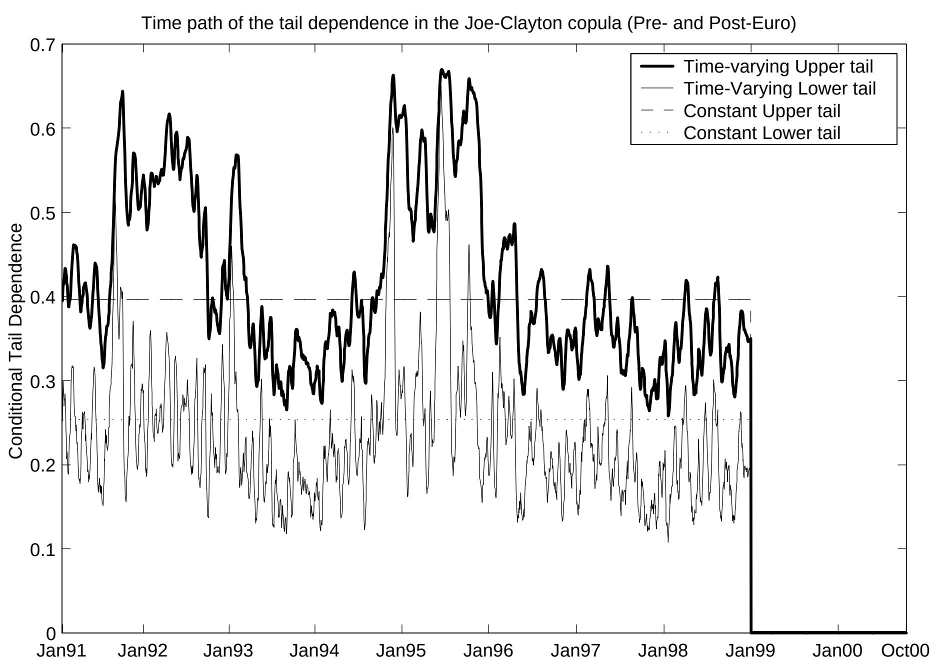

Figure 11: *Time-varying tail dependence from the Joe-Clayton copulas allowing for a structural break at the introduction of the Euro on Jan 1, 1999.*