A Comparison of Dynamic Copula-GARCH Models

Lucas C. Dowiak

March 11, 2024

PhD Program in Economics, City University of New York, Graduate Center, New York, NY, 10016, *Email: ldowiak@gradcenter.cuny.edu*

Abstract: This essay uses three separate tools to model the evolution of dependence structure between major global currencies from 2000 to 2018. It uses the ARMA-GARCH approach to time series modeling with Copula theory and multi-regime time series models to provide a flexible framework to model the return series of financial assets. The framework is applied to the log returns of the exchange rates between the US Dollar and the European Euro, British Pound, and Japanese Yen from 2000 to 2018.

Keywords: Copulas, financial crisis, foreign exchange, time-varying dependence **JEL Classification**: C51, G01, G15, F31

1 Introduction

The field of international finance has a long-running concern regarding the distribution of asset returns. Harry Markowitz (Markowitz 1952; Markowitz 1959) provided the launching pad for modern portfolio theory by using a Gaussian framework to model the returns for a portfolio composed of an arbitrary number of assets. Over the past few decades though, a consensus has conformed on the belief that assuming normality, or joint-normality, of asset returns is not always wise. Empirical evidence strongly supports the fact that returns for certain asset classes frequently exhibit traits that deviate from normality. Two famous articles Mandelbrot 1963 and Fama 1965 present evidence that the tails of the Gaussian distribution are far too narrow to account for the number of observations occurring at least two standard deviations from the mean. More recently, F. M. Longin 1996 abandons the use of elliptical distributions to describe extreme returns altogether. Ang and Chen 2002 and Kroner and Ng 1998 study asymmetric co-movements in the returns of certain equity assets. Ang and Chen 2002 document stronger downside risk between a range of portfolios and the aggregate U.S. market than upside risk. Kroner and Ng 1998 find "significant asymmetric effects in both the variances and covariances" for large-firm and small-firm returns.

The assumption of a constant correlation has also been challenged. In a popular article, F. Longin and Solnik 1995 review the correlation of excess returns on equities for the countries of France, U.S., U.K., Sweden, Japan, Canada, and Germany using monthly data from 1960-1990. Their results suggest a rejection of a constant international conditional correlation. In a similar vein, Ramchand and Susmel 1998 use a markov switching ARCH model and find that the correlation between the U.S. and other foreign markets is significantly higher in times of elevated volatility. Rodríguez 2007 studies financial contagion during the Asian financial crisis (circa 1997) and the Mexican peso devaluation (circa 1994). By combining a markov switching ARCH model with copula-based distributions, Rodríguez 2007 documents a change in dependence between a high variance regime and a low variance regime.

In this article, we employ a framework that will allow asset returns to deviate from the normality assumption by distribution and by time-varying dependence. In particular, this article studies evolution of three exchange rate pairs from 2000 to 2018, a period of time that includes the effect the financial crisis of 2007-2008 had on the foreign exchange market. To incorporate a degree of flexibility regarding the distribution of returns, copula theory will be applied to daily exchange rate data for the world's most influential currencies: the US Dollar, Japanese Yen, European Euro, and the British Pound. According to the 2010 Triennial Central Bank Survey conducted by the Bank of International Settlements, these currencies account for 78.0% of all foreign exchange turnover in 2010. To capture potential breaks in the dependence over time, we estimate and compare three popular multi-regime models: a markov transition model, a smooth transition model, and a sequential break point model.

Daily data is required since one focus of the article concentrates on estimating the dependence between observations in the extremes of the joint distribution. Weekly or monthly data will lead to a sample size that is too small to populate the tails, leading to imprecise estimates of the dependence and distributional parameters. The source for these series will be the Bank of England, whose database on historical exchange rates includes many of the world's currencies and is easy to access. The timeframe under review will be restricted to January 4, 2000 to October 4, 2018.

The layout for the rest of the paper is as follows. Section 2 discusses all relevant copula theory and important definitions. Section 3 provides the details of three popular multi-regime time series models while section 4 focuses on estimating an ARMA-GARCH model for the exchange rates of the world's four most used currencies, by volume. Section 5 reviews the evolution of dependence between the major foreign exchange rates during the timeframe under review. Section 6 offers some concluding remarks.

2 Copula Theory

In this section, we briefly discuss the necessary copula theory and definitions that are put to use in this paper. Since our analysis pertains exclusively to bivariate distributions, all theory will be presented in the bivariate case. An accessibly treatment of basic copula theory can be found in Nelsen 2007 while additional topics on multivariate dependence can be found in Joe 1997. For the individual in need of a concise introduction to the topic, the articles by Embrechts, Lindskog, and A. McNeil 2003 and Genest and Favre 2007 are a great entry point.

2.1 Sklar's Theorem

The term 'copula' has already been used various times in the introduction. To start, we provide a formal definition:

Definition 1 *A bivariate copula is a function C from $[0,1]^2$ to $[0,1]$ that satisfies (1) For every u, w in $[0,1]$: $C(u,0) = C(0,w) = 0$; (2) For every u, w in $[0,1]$: $C(u,1) = u$ and $C(1,w) = w$; and (3) For every u_1, u_2, w_1, w_2 in $[0,1]$ such that $u_1 \leq u_2$ and $w_1 \leq w_2 : C(u_2, w_2) - C(u_1, w_2) - C(u_2, w_1) + C(u_1, w_1) \geq 0$.*

Taken together, properties (1) - (3) imply that a copula is a nondecreasing function in each argument, a quality shared by all distribution functions. It is, in fact, helpful to think of copulas as distribution functions whose domain is restricted to the unit square—or the unit hypercube in higher dimensions. With this in mind, we present Sklar's theorem, the foundation for empirical applications of copulas.

Theorem 1 (Sklar, 1959) *Let F_{XY} be a joint distribution function with margins F_X and F_Y. Then there exists a copula C such that for all x, y in $\bar{\mathbb{R}}$,*

$$F_{XY}(x,y) = C(F_X(x), F_Y(y)) = C(u,w) \tag{1}$$

If F_X and F_Y are continuous, then C is unique; otherwise, C is uniquely determined on $RanX \times RanY$. Conversely, if C is a copula and F_X and F_Y are distribution functions, then the function F_{XY} defined by equation 1 is a joint distribution function with margins F_X and F_Y.

The copula approach to multivariate distributions first transforms each marginal distribution from the extended real line $\left(\bar{\mathbb{R}}\right)$ to the unit interval through the *probability integral transform* (PIT). First investigated by Rosenblatt 1952, this transform states that if X is a continuous random variable with density f_X, then the random variable U defined by

$$u = \int_{-\infty}^{x} f_X(t)\, dt = F_X(x) \tag{2}$$

is identically and independently distributed (iid) uniformly on the unit interval. In the bivariate case, this transformation shifts the domain of analysis from the extended real plane to the unit square. A copula is then chosen to install a dependence structure between the two marginal distributions. Sklar's theorem confirms the existence of a copula C whose value $C(u,w) = C(F_X(x), F_Y(y))$ provides the joint probability $P[U < u,\ W < w] = P[X < x,\ Y < y]$. In application, the copula approach allows us to separate a joint distribution F_{XY} into marginal distributions F_X, F_Y and a dependence function C. The dependence between the marginal distributions F_X, F_Y is completely governed by the copula, and just as Sklar's theorem states, we can model this dependence separately.

2.2 Patton's Extension

Taking inspiration from Hansen 1994 work on time-varying density estimation, Patton 2006 broadens Sklar's original theorem to conditional distributions. This extension paves the way for conditional copulas with time-varying parameters. Consider a bivariate series $(X_t, Y_t)'$ on some information set $\Im_{t-1}$. In the same manner as

Theorem 1, we can decompose the conditional distribution of (X_t, Y_t) given $\Im_{t-1}$ into its marginal distributions and the conditional copula. Given

$$X_t \mid \Im_{t-1} \sim F_X(x|\Im_{t-1}), \quad Y_t \mid \Im_{t-1} \sim F_Y(y|\Im_{t-1})$$
$$(X_t, Y_t)|\Im_{t-1} \sim F_{XY}(x, y|\Im_{t-1})$$

Then

$$F_{XY}(x, y|\Im_{t-1}) = C(F_X(x|\Im_{t-1}), F_Y(y|\Im_{t-1}) \mid \Im_{t-1}) \tag{3}$$

As discussed in Patton 2006, the same information set must be used for each margin and the copula. Conditioning the marginal models and the copula on different information sets will yield a function $\hat{F}_{XY}(\cdot|\Im_{t-1})$ that is not the joint distribution of $(X, Y)|\Im_{t-1}$. In practice, this complication has the potential to offset the most attractive quality of the copula approach: modeling each marginal distribution and the copula separately.

Say our information set consists of two conditioning variables $Z_1, Z_2 \in \Im_{t-1}$, and we condition X on Z_1 and Y on Z_2. The only way $\hat{F}_{XY}(\cdot|\Im_{t-1})$ will be the joint distribution function of $(X, Y)|\Im_{t-1}$ is the special case when $F_X(x|Z_1) = F_X(x|Z_1, Z_2)$ and $F_Y(y|Z_2) = F_Y(y|Z_1, Z_2)$. This can occur when Z_1 is orthogonal to $Y|Z_2$ and Z_2 is orthogonal to $X|Z_1$. To provide an example of this special case in the context of exchange rates, define the Euro-Dollar return as X_t with a lagged value as its conditioning variable $X_{t-k} = Z_1$. Likewise, define the Yen-Dollar return as Y_t with a lagged value as its conditioning variable $Y_{t-k} = Z_2$. If the lagged Euro-Dollar return X_{t-k} has no predictive power for the conditional Yen-Dollar return $Y_t|Y_{t-k}$, and conversely the lagged Yen-Dollar return Y_{t-k} has no predictive power for the conditional Euro-Dollar return $X_t|X_{t-k}$, then our orthogonality conditions are met and we can correctly interpret the function $\hat{F}_{XY}(\cdot|\Im_{t-1})$ as the joint distribution function of $(X, Y)|\Im_{t-1}$. These conditions must be evaluated and motivates the regression tests discussed at the end of Section 4.

2.3 Methods for Constructing Copulas

There are several methods for constructing copulas. The three most popular methods are used in this study and subsequently discussed.

2.3.1 The Inversion Method

The Gaussian and Student-t Copula are created using the inversion method: $C(u,w) = F_{XY}\left(F_X^{-1}(u), F_Y^{-1}(w)\right)$, where F_X^{-1} and F_Y^{-1} are quantile functions. A more exact description of the bivariate Student-t copula distribution is written as:

$$\begin{aligned} C_{\boldsymbol{t}}(u,w;\rho,\nu) &= \boldsymbol{t}_v\left(t_\nu^{-1}(u), t_\nu^{-1}(w)\right) \\ &= \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(w)} q(1-\rho)^{-\frac{1}{2}} \left[1+\nu^{-1}\left(x^2-2\rho xy+y^2\right)\right]^{-(\nu+2)/2} dxdy \end{aligned} \tag{4}$$

where $\boldsymbol{t}_v$ is the bivariate t-distribution function, t_ν^{-1} is the univariate quantile function, $x = t_\nu^{-1}(u)$, $y = t_\nu^{-1}(w)$, and $q = \Gamma\left(\frac{1}{2}\right)^{-1}\Gamma\left(\frac{\nu+2}{2}\right)(\nu\pi)^{-1}$. Differentiation of equation 4 with respect to u and w produces the Student-t copula density:

$$\begin{aligned} c_{\boldsymbol{t}}(u,w;\rho,\nu) &= Q(1-\rho)^{-\frac{1}{2}}\left[1+\nu^{-1}(1-\rho)^{-1}\left(x^2-2\rho xy+y^2\right)\right]^{-(\nu+2)/2} \\ &\times \left[\left(1+\nu^{-1}x^2\right)\left(1+\nu^{-1}y^2\right)\right]^{(v+1)/2} \end{aligned} \tag{5}$$

where $Q = \Gamma\left(\frac{v}{2}\right)^{-1}\Gamma\left(\frac{\nu+1}{2}\right)^{-2}\Gamma\left(\frac{\nu+2}{2}\right)$. The Gaussian copula is obtained in a similar fashion. If $\boldsymbol{N}_\rho$ denotes the standard bivariate normal distribution and Φ^{-1} denotes the quantile function for the univariate standard normal distribution, the Gaussian copula is summarized by $C_{\boldsymbol{N}}(u,w;\rho) = \boldsymbol{N}_\rho\left(\Phi^{-1}(u), \Phi^{-1}(w)\right)$.

2.3.2 Archimedean Copulas

The second class we consider are referred to as Archimedean copulas. Families of these copulas are defined by their generator function φ, which usually contain one or two parameters. These parameters play a fundamental role in specifying the dependence structure between the marginal distributions. Consider the following definitions:

Definition 2 *The generator* φ *of an Archimedean copula is a function from* $[0,1]$ *to* $[0,\infty]$ *that is continuous, strictly decreasing, convex, and satisfies* $\varphi(1) = 0$.

Definition 3 *The pseudo-inverse of φ is the function $\varphi^{[-1]}$ from $[0,\infty]$ to $[0,1]$ given by $\varphi^{[-1]}(t) = \varphi^{[-1]}(t)$ for t in $[0,\varphi(0)]$ and $\varphi^{[-1]}(t) = 0$ for t in $[\varphi(0),\infty]$. If $\varphi(0)=\infty$, the pseudo-inverse is simply the inverse $\varphi^{-1}(t)$.*

Given definition 2 and definition 3, an Archimedean copula is constructed in the following manner:

$$C(u,w) = \varphi^{[-1]}(\varphi(u) + \varphi(w)) \tag{6}$$

Again, when viewed through the context of Theorem 1, one can interpret $C(u,w)$ as a cumulative distribution function. To provide a concrete example, consider the Gumbel copula. The generator function is given by $\varphi(t) = (-\ln t)^{\theta}$. Using equation 6, it is straight forward to construct the copula that corresponds to this particular generator:

$$C_G(u,w;\theta) = \exp\left\{-\left[(-\ln u)^{\theta} + (-\ln w)^{\theta}\right]^{\frac{1}{\theta}}\right\} \tag{7}$$

Differentiating equation 7 with respect to u and w yields the copula density:

$$\begin{aligned} c_G(u,w;\theta) \quad = \quad & C_G(u,w;\theta)(\ln u \cdot \ln w)^{\theta-1}(uw)^{-1}\left[(-\ln u)^{\theta} + (-\ln w)^{\theta}\right]^{\frac{1-2\theta}{\theta}} \times \\ & \left[\theta - 1 + \left[(-\ln u)^{\theta} + (-\ln w)^{\theta}\right]^{\frac{1}{\theta}}\right] \end{aligned} \tag{8}$$

2.3.3 Mixture Copulas

A third method for copula construction exists if one takes an amalgam of copulas and combines them in a certain manner. Just as we can create new distributions by taking a mixture of Gaussians, we can construct mixture copulas. Specifically, if we have a finite collection of N copulas, we can show that the copula C^* in equation 9 satisfies definition 1.

$$C^*(u,w;\boldsymbol{\theta}) = \sum_{k=1}^{N} \pi_k C_k(u,w;\theta_k) \tag{9}$$

with $\pi_k \geq 0; \pi_1 + \cdots + \pi_N = 1$ and $\boldsymbol{\theta} = (\theta_1, ..., \theta_N, \pi_1, ..., \pi_N)$.

The mixture concept can be extended to include a collection of infinite copulas. Suppose we have a copula family governed by a single parameter $C(\cdot;\theta)$. If we consider the parameter θ as random variable drawn from a continuous distribution $\theta \sim H$, then integrating over the support of H produces a copula:

$$C^*(u, w) = \int_{\mathbb{R}} C(u, w; \theta)\, dH(\theta) \tag{10}$$

The distribution H is referred to as the 'mixing distribution'. Notice that in the infinite case (2.10), the copula family is consistent while the parameter of that family varies. This is in contrast to the finite mixture case 9, where the component copulas C_k are free to differ from one another.

2.4 Survival Copulas and Symmetry

It was mentioned that Theorem 1 confirms the existence of a copula C whose value $C(u, w) = C(F_X(x), F_Y(y))$ provides the probability $P[X < x, Y < y] = P[U < u, W < w]$. At times, the joint probability $P[X > x, Y > y] = 1 - F_{XY}(x, y) \equiv S_{XY}(x, y)$ is also of interest. As econometricians, we refer to the function S_{XY} as a survival or duration function. For an arbitrary copula C, here is a corresponding survival copula $\hat{C}$ that models the joint survival function:

$$\begin{aligned}
S_{XY}(x, y) &= P[X > x, Y > y] \\
&= 1 - F_X(x) - F_Y(y) + F_{XY}(x, y) \\
&= S_X(x) + S_Y(y) - 1 + C(1 - S_X(x), 1 - S_Y(y)) \\
&= \hat{C}(S_X(x), S_Y(y))
\end{aligned}$$

The relationship between a copula C and its survival copula $\hat{C}$ is given by:

$$\hat{C}(u, w) = u + w - 1 + C(1 - u, 1 - v) \tag{11}$$

Equation 11 has the effect of rotating a copula 180 degrees along the main diagonal of the unit square. Many papers in the literature discuss the process of producing 'symmetric' copulas, especially in the bivariate case. There is nothing intimidating about this process. It is a special case of the mixture method defined in section (2.3.3). For $N = 2$, one can take an arbitrary copula C, its corresponding survival copula $\hat{C}$, and apply equal weights so that $C^* (u, w) = 0.5 \left[C (u, w) + \hat{C} (u, w)\right]$. At times, restricting the weights in mixture copula can aid in identification. If the model is simple enough, this restriction can be jettisoned, and the weights can be estimated along with the parameters of the component copulas.

2.5 Dependence

There are many forms of dependence when discussing multivariate distributions. This section outlines two specific forms: rank correlation described by Kendall's tau and 'tail–dependence'. Noticeably, our analysis avoids using Pearson's correlation coefficient $\rho = cov (X, Y) [var (X) var (Y)]^{-1/2}$, which provides a measure of *linear* association. Although appropriate for elliptical distributions, it is less appropriate for discussing the association of random variables whose joint distribution is non-elliptical, or whose dependence is non-linear. A practical discussion of this issue is covered in Embrechts, A. J. McNeil, and Straumann 2002.

2.5.1 Kendall's Tau

A global measure of association, Kendall's tau is a popular measure of the degree of monotonic dependence between two random variates. Invariant to strictly increasing transformations, Kendall's tau provides a measure of correlation between the ranks of two marginal distributions. This short section begins with the definition of concordant and discordant observations.

Definition 4 *For a vector of continuous random variables (X, Y), let (x_i, y_i) and (x_j, y_j) represent two observations. We say (x_i, y_i) and (x_j, y_j) are concordant if $(x_i - x_j) (y_i - y_j) > 0$ and discordant if $(x_i - x_j) (y_i - y_j) < 0$.*

The population version of Kendall's tau is defined as the probability of concordance minus the probability of discordance:

$$\tau = P [(X_1 - X_2) (Y_1 - Y_2) > 0] - P [(X_1 - X_2) (Y_1 - Y_2) < 0] \qquad (12)$$

For an arbitrary copula C, Nelson (2006, page 159) demonstrates that Kendall's tau can be obtained by integrating over the unit square in the following manner:

$$\tau = 4 \int \int_{[0,1]^2} C\left(u, w\right) dC\left(u, w\right) - 1 \tag{13}$$

By choosing to represent global dependence through Kendall's tau, we are creating an appropriate standard to compare estimates of dependence produced by different copulas, elliptical or otherwise. Equation (13) is especially helpful when calculating Kendall's tau for mixture copulas, where in most cases, no closed form relationship exists between the parameters of the mixture copula and the measure of correlation. Brute force numerical integration of (13) is a useful fallback in these situations.

2.5.2 Tail Dependence

Tail dependence is a localized measure of association and allows us to make probabilistic statements about observations in the upper-right corner and lower-left corner of the unit square. In our present context on exchange rates, upper tail dependence is a statement about the probability of two currency pairs experiencing strong concurrent depreciations against the dollar. Conversely, lower tail dependence is a statement about the probability of two currency pairs experiencing strong concurrent appreciations against the dollar. It is worth pointing out that tail dependence is a property of the copula and not of the marginal distributions. The following definition is due to Joe 1997.

Definition 5 *For a copula C, if $\lambda^U \equiv \lim_{t \nearrow 1} P\left[X > F_X^{-1}\left(t\right) | Y > F_Y^{-1}\left(t\right)\right] = \lim_{t \nearrow 1} \frac{1-2t+C(t,t)}{1-t}$ exists, then C has upper tail dependence if λ^U is in $(0,1]$ and no upper tail dependence if $\lambda^U = 0$. Likewise, if $\lambda^L \equiv \lim_{t \searrow 0} P\left[X \le F_X^{-1}\left(t\right) | Y \le F_Y^{-1}\left(t\right)\right] = \lim_{t \searrow 0} \frac{C(t,t)}{t}$ exists, C has lower tail dependence if λ^L is in $(0,1]$ and no lower tail dependence if $\lambda^L = 0$.*

When considering a mixture copula C^* as described in equation (9), the tail dependence is the weighted average of the tail dependence values for each component copula:

$$\lambda^{*,i} = \sum_{k=1}^{N} \pi_k \lambda_k^i \qquad i = \{L, U\} \tag{14}$$

The weights π_k in equation (14) are the same as the corresponding mixture weights in equation (9). This result can be verified by applying l'Hôpital's rule to the limits in definition (5), substituting in dC^*, and simplifying.

2.6 Estimation

Maximum likelihood is the most frequently used procedure for parametric estimation of copula models and is the method we employ here. A joint distribution is described by

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)) \tag{15}$$

Provided that F_X and F_Y are differentiable while F_{XY} and C are twice differentiable, the density function is found by taking the cross partial derivative of (15):

$$\begin{aligned} f_{XY}(x, y) &\equiv \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \\ &= \frac{\partial F_X(x)}{\partial x} \cdot \frac{\partial F_Y(y)}{\partial y} \cdot \frac{\partial^2 C(F_X(x), F_Y(y))}{\partial x \partial y} \\ &= f_X(x) \cdot f_Y(y) \cdot c(u, w) \end{aligned}$$

The log-likelihood value follows immediately.

$$\begin{aligned} \mathcal{L} &= \sum_{t=1}^{T} \log[f_X(x_t)] + \sum_{t=1}^{T} \log[f_Y(y_t)] + \sum_{t=1}^{T} \log[c(u_t, w_t)] \\ &= \mathcal{L}_x + \mathcal{L}_y + \mathcal{L}_c \end{aligned}$$

Full information maximum likelihood would require the simultaneous estimation of both the marginal distribution parameters along with the parameters of the copula. Although this approach would provide the most efficient estimates, the large dimensionality of the parameter space makes maximization of the likelihood function particularly difficult. A second, more feasible approach known as the 'inference on margins' (IFM) first estimates the marginal models and then separately estimates the copula model. This method leads to less efficient, but still consistent estimates.

3 Time-Varying Models

A brief survey of various time-varying copulas in the literature is provided by Manner and Reznikova 2012. All models are generalized to k different regimes.

3.1 Sequential Change Point Analysis

This first method we use to introduce time-varying dependence is the most simple. Dias and Embrechts 2004 and Dias and Embrechts 2009 outline a framework to detect structural breaks in the parameters of a copula over time. We employ that framework here. Consider the sequence of random vectors (U_t, W_t) created by the PIT from our marginal models where $U_t \equiv F_X\left(X_t|\Im_{t-1}\right)$ and $W_t \equiv F_Y\left(y_t|\Im_{t-1}\right)$. The dependence at each date can be modeled by a copula $C\left(U_t, V_t;\ \boldsymbol{\theta}_t\right)$ with parameter vector $\boldsymbol{\theta}_t = (\theta_{1,t}, ..., \theta_{N,t})$. The vector θ_t should contain the weights $\pi_{i,t}$ for any mixture copula described in section (2.3.3). The idea is to test the null hypothesis of an absence of any structural break in the copula parameters

$$H_0 : \boldsymbol{\theta}_t = \boldsymbol{\theta} \tag{16}$$

against the alternative hypothesis of a single structural break at unknown time $\tau \in [1, T]$.

$$H_A : \boldsymbol{\theta}_t = \begin{cases} \boldsymbol{\theta}_1 & \text{if} \quad 1 \leq t \leq \tau \\ \boldsymbol{\theta}_2 & \text{if} \quad \tau < t \leq T \end{cases} \tag{17}$$

Define $L_0\left(\theta_0\right)$, $L_1\left(\theta_1\right)$, and $L_2\left(\theta_2\right)$ as the log-likelihood functions of the copula model estimated on the entire sample, the first τ observations, and the last $T-\tau$ observations, respectively. The statistic Λ_τ for a structural break in the parameter of the copula model takes the form of a generalized likelihood ratio test:

$$-2\log\left(\Lambda_\tau\right) = 2\left[L_1\left(\hat{\theta}_1\right) + L_2\left(\hat{\theta}_2\right) - L_0\left(\hat{\theta}_0\right)\right] \tag{18}$$

Since the break point τ is not known, ex ante, a sequence of tests will be performed for each date in a subset of the full sample $(t_l, t_h) \subset [1, T]$. To avoid the decreasing power of the test as the break point nears the edge of the series, the subset of potential break dates (t_l, t_h) will exclude a number of dates at the beginning and end of the sample series. A five percent trim is used at both ends. With this setup, the null will be rejected for large values of:

$$Z_t = \max_{t_l \leq t \leq t_h} \left[-2 \log \left(\Lambda_t\right)\right] \tag{19}$$

Inference is based on the asymptotic distribution of $Z_t^{1/2}$, which has been derived by Csörgö and Horváth 1997. Due to the slow convergence of $Z_t^{1/2}$ to its asymptotic distribution, an approximation is used to improve the rejection regions of the test for small samples. Simulation results in Dias and Embrechts 2004 provided tentative evidence that this approximation provides accurate p-values.

$$\begin{aligned} P\left(Z_t^{1/2} \geq x\right) \quad &\approx \quad \frac{x^p \exp\left\{-x^2/2\right\}}{2^{p/2}\Gamma\left(p/2\right)} \times &(20)\\ &\left[\log \frac{\left(1-h\right)\left(1-l\right)}{hl} - \frac{p}{x^2} \log \frac{\left(1-h\right)\left(1-l\right)}{hl} + \frac{4}{x^2} + O\left(\frac{1}{x^4}\right)\right] &(21)\end{aligned}$$

Taking into consideration the possibility of more than one break, the approach described in this section can be iterated on smaller disjoint subsets produced by successful rejections of the null. Say we find an estimate $\hat{\tau}$ of a break date in our full series. We can test for additional breaks in the subsamples $[1, \hat{\tau}]$ and $[\hat{\tau}+1, T]$. This process can continue until the null hypothesis is no longer rejected for any subsample.

Once a collection of break dates $\hat{\tau}_1, \hat{\tau}_2, ..., \hat{\tau}_{k-1}$ has been estimated, Dias and Embrechts 2009 suggest a repartition method identical to Bai 1997. Since the alternative hypothesis of the test is that of a single break, the sequential process to find additional breaks could lead to biased estimates for the location of $\hat{\tau}_i : i > 1$. To alleviate these concerns, each break date is reestimated $\hat{\tau}_i^*$ by applying the test to the subsamples $[\hat{\tau}_{i-1} + 1 \leq t \leq \hat{\tau}_{i+1}]$ for $i = 1, ..., k-1$. Here $\hat{\tau}_0 = 0$ and $\hat{\tau}_k = T$.

The advantages of this approach is its relative simplicity. Many static copula have tractable gradients and hessians. The dimensionality of the parameter space is kept to a minimum as well. The only dynamics, so-to-speak, are the estimated break points. Other regime-varying models in the literature, such as the markov transition and smooth transition models discussed later, usually produce results where the dependence structure of the underlying data may change from day-to-day.

3.2 Markov Transition

In a second approach, this essay uses a latent state variable s_t to govern the dependence of our exchange rate series at time t. First brought into the econometrics literature by Hamilton 1989 and Hamilton 1994, changes in the state variable are modeled by a first-order markov transition matrix. In this setup, the probability of being in some particular state tomorrow s_{t+1} depends solely on what state the variable is in today s_t. A more concrete description of the process is given by the probability statement $P\left[s_{t+1}=j \mid s_t=i,\ s_{t-1}=m,\ldots\right]=P\left[s_{t+1}=j \mid s_t=i\right]=p_{ij}$, where we have $p_{i1}+p_{i2}+\cdots+p_{ik}=1$ for all $i=1,...,k$. The vectors of transition probabilities from one state to another are organized into a single $k\times k$ transition matrix $\mathbf{P}$.

$$\mathbf{P}=\begin{bmatrix} p_{11} & p_{21} & \cdots & p_{k1} \\ p_{12} & p_{22} & \cdots & p_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1k} & p_{2k} & \cdots & p_{kk} \end{bmatrix} \tag{22}$$

For each of the k possible states, a copula is used to model the dependence for that state. These copula densities are stacked into a $k \times 1$ vector and indexed by time:

$$\boldsymbol{\eta}_t=\begin{bmatrix} \eta_{t,1} \\ \vdots \\ \eta_{t,k} \end{bmatrix}=\begin{bmatrix} c_1\left(F_X\left(x_t|\Im_{t-1}\right), F_Y\left(y_t|\Im_{t-1}\right);\theta_1\right) \\ \vdots \\ c_k\left(F_X\left(x_t|\Im_{t-1}\right), F_Y\left(y_t|\Im_{t-1}\right);\theta_k\right) \end{bmatrix} \tag{23}$$

Since we can not know for sure which state the process is in at any given date, inference must be made using all the information available at that time. At each date, we assign a probability that the dependence governing the exchange rate returns comes from a state $P\left[s_t=j \mid x_t,x_{t-1},...,y_t,y_{t-1},...,\theta_1,...,\theta_k\right]$. These conditional probabilities are collected into another $k \times 1$ vector denoted $\hat{\boldsymbol{\xi}}_{t|t}$. Given a set of initial probabilities $\hat{\boldsymbol{\xi}}_{1|0}$, optimal inferences are found by iterating on the following equations:

$$\hat{\boldsymbol{\xi}}_{t|t} = \frac{\hat{\boldsymbol{\xi}}_{t|t-1} \odot \boldsymbol{\eta}_t}{\mathbf{1}^\top \left(\hat{\boldsymbol{\xi}}_{t|t-1} \odot \boldsymbol{\eta}_t \right)} \tag{24}$$

$$\hat{\boldsymbol{\xi}}_{t+1|t} = \mathbf{P} \hat{\boldsymbol{\xi}}_{t|t} \tag{25}$$

where the operator $\odot$ is the hadamard product denoting element-by-element multiplication and $\mathbf{1}$ is a column of ones of length k. Conveniently, the log-likelihood value can be calculated as a side product to this algorithm, with the value equaling

$$\mathcal{L}_c = \sum_{t=1}^{T} \log \left[\mathbf{1}^\top \left(\hat{\boldsymbol{\xi}}_{t|t-1} \odot \boldsymbol{\eta}_t \right) \right] \tag{26}$$

What's more, we can use these conditional probabilities to gain an even better gauge of the latent state variable. Starting with $\hat{\boldsymbol{\xi}}_{T|T}$, we can iterate backwards from T to 1 to calculate the smoothed probabilities $\hat{\boldsymbol{\xi}}_{t|T}$ by iterating on:

$$\hat{\boldsymbol{\xi}}_{t|T} = \hat{\boldsymbol{\xi}}_{t|t} \odot \left[\mathbf{P}^\top \left(\hat{\boldsymbol{\xi}}_{t+1|T} \, (\div) \, \hat{\boldsymbol{\xi}}_{t+1|t} \right) \right] \tag{27}$$

The operator $(\div)$ denotes element-by-element division. Changes in dependence through time will be captured by the different copula parameters governing each state.

3.3 Smooth Transition

The third, and final, approach to time-varying dependence this study implements is a version of the smooth-transition copula-GARCH (STCG) model. We employ a single copula but allow the parameters of the copula to evolve over time. More specifically, for any given copula $C(\cdot|\theta_t)$, we allow the possibility for each parameter θ_i to undergo multiple transitions.

$$\theta_{i,t} = \theta_{i,1} + \sum_{m=1}^{k-1} \left(\theta_{i,m+1} - \theta_{i,m} \right) G_m \left(s_t; c_{i,m}, \gamma_{i,m} \right) \tag{28}$$

In general, $G_m(\cdot)$ is a bounded function with respect to the continuous transition variable s_t. In this study, s_t will be constructed as a linear time trend $s_t = \frac{t}{T}$ and $G_m(\cdot)$ takes the form of the logistic function:

$$G(s_t; c, \gamma) = (1 + \exp\{-\gamma(s_t - c)\})^{-1} \qquad \gamma > 0 \tag{29}$$

Other specifications for $G(\cdot)$ can be found in Dijk and Franses 1999 and Öcal and Osborn 2000. The parameters $\boldsymbol{\gamma}_i = (\gamma_{i,1}, ..., \gamma_{i,k-1})$ control the pace of each transition while the members of vector $\mathbf{c}_i = (c_{i,1}, ..., c_{i,k-1})$ locate the inflection points of each transition. It is difficult to overstate how nonlinear a copula model becomes when subject to the full generality of specifications (28) and (29). To aid in identification, some restrictions are usually imposed. In our case, we restrict the location variables of each transition to be identical across copula parameters: $c_{i,m} = c_{j,m}$ for all i, j and each m. In addition, we also hold the weights π_i for mixture copulas constant.

4 Marginal Distributions

As stated earlier, the copula approach enables us to model the marginal distributions and dependence structure separately. In this section, we summarize how each univariate margin is modeled and validated. For any particular exchange rate, define the return as $x_t = \log(S_t) - \log(S_{t-1})$, where S_t is the spot price at time t. All exchange rates are denoted in foreign currency units per US dollar. Like many other financial assets, exchange rates are subject to volatility clustering, skewness, and leptokurtosis. To manage these properties, each return series $\{X_t\}_{t=1}^{T}$ is modeled using a ARMA-GARCH specification with a Student-t distribution or a skewed Student-t distribution. The overall dynamics of the ARMA-GARCH process can be summarized by the following two equations.

$$x_t = \mu_t + \varepsilon_t \tag{30}$$

$$\varepsilon_t = \eta_t \sqrt{h_t} \tag{31}$$

where η_t are independent and identically distributed with $E[\eta_t] = 0$ and $E[\eta_t^2] = 1$. and has a probability distribution g. The conditional mean for the log-returns can be formulated by the general ARMA process:

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{j=1}^{q} \theta_j \, \varepsilon_{t-j} \tag{32}$$

The value p is the order of of autoregressive component and q is the order of the moving average component. The conditional volatility process for the models under consideration can be generalized with the following formula. Functions A and B are generic stand-ins that will differ across the various GARCH specifications.

$$h_t = \omega + \sum_{j=1}^{m} \alpha_j A(\varepsilon_{t-j}^2) + \sum_{i=1}^{k} \beta_i B(h_{t-i}) \tag{33}$$

The conditional mean and variance are used to center and scale the innovation terms.

$$z_t = \frac{x_t - \mu_t}{\sqrt{h_t}} \tag{34}$$

By collecting all the parameters in the conditional mean and variance equations into one vector $\Delta = [\phi, \theta, \alpha, \beta]$, the distribution for the log-returns f can be written as a function of Δ, the chosen error distribution g, the conditional mean μ_t and conditional variance σ^2, and any necessary shape parameters λ of the distribution:

$$f(x_t|\mu_t, \sigma_t^2, \lambda, \Delta) = \frac{1}{\sqrt{h_t}} g(z_t|\lambda, \Delta) \tag{35}$$

In certain times the researcher may want the capability to accurately model return values that are not perfectly symmetric which is a characteristic of the t-distribution. To produce a skewed t-distribution, this essay follows Fernandez and Steel 1998. For a unimodal and symmetric density $f(x)$ that is decreasing in $|x|$, a skew parameter ξ can force asymmetry upon $f(x)$, generating a new distribution $p(x|\xi)$ where

$$p(x|\xi) = \frac{2}{\xi + \frac{1}{\xi}} \left[f(x\xi) \, \mathbf{I}_{\{x<0\}} + f\left(x\xi^{-1}\right) \mathbf{I}_{\{x\geq 0\}} \right] \tag{36}$$

4.1 Conditional Variance

For the conditional variance, two different specifications are employed in this essay. The first is the standard GARCH(b,k) model,which has the following form:

$$h_t = \omega + \sum_{i=1}^{b} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{k} \beta_j h_{t-j} \tag{37}$$

A separate approach proposed by Glosten, Jagannathan, and Runkle 1993 adds a bit of flexibility to the GARCH model specified in equation (37) by allowing asymmetrical responses to shocks[1] of opposing signs.

$$h_t = \omega + \sum_{i=1}^{b} \left(\alpha_i + \gamma_i \mathbf{I}_{\{\varepsilon<0\}}\right) \varepsilon_{t-i}^2 + \sum_{j=1}^{k} \beta_j h_{t-j} \tag{38}$$

Here, $\mathbf{I}_{\{.\}}$ is the indicator function with value one if the statement in the brackets is true and zero otherwise. This second model allows for asymmetric responses to past shocks ε_{t-i}. If the lagged shock is positive ($\varepsilon_{t-i} > 0$), the coefficient has the value α_i. If the lagged shock is negative ($\varepsilon_{t-i} < 0$), the coefficient has the value $\alpha_i + \gamma_i$.

4.2 Model Validation

The aim is to specify a parsimonious model that produces (standardized) residuals that are independent and identically draws from the selected distribution, g. To confirm that we have well specified models, a series of diagnostic checks are performed on the fitted ARMA-GARCH model for each currency pair and time-frame. The tests presented for validation are included in the same software that is used to model the log returns as an ARMA-GARCH process, Ghalanos 2020. There are five categories of tests. One set of tests checks for the correct ARMA specification while a second set of tests checks for the correct ARCH specification. A third test checks the empirical distribution of the standardized residuals against the theoretical distribution selected by the researcher, g. A fourth set of tests checks for asymmetric effects in the residuals. This is particular relevant when considering the gjr-GARCH model

[1] By shocks, we are referring to the error term in equation (30)

of equation (38). A fifth set of tests checks for parameter stability using Nyblom 1989.

To test for a correct ARMA specification the weighted Ljung-Box test of Fisher and Gallagher 2012 is used. The authors propose a statistic based on the determinant of a Toeplitz matrix whose elements are values of a series' sample autocorrelations. The statistic has a functional form that can be viewed as a weighted Ljung-Box statistic.

$$Q_W = n(n+2)\sum_{j=1}^{m}\frac{m-j+1}{m}\frac{[r_j(x)]^2}{n-j} \tag{39}$$

The value m is set to the largest lag and the weight term, $(m-j+1)/m$, decreases as the length of the lag increases. In the equation above, $\hat{r}_j(x)$ is used to denote the sample correlation value of the series x at the kth lag.

$$r_j(x) = \frac{\sum_{t=j+1}^{n}(x_t-\bar{x})(x_{t-j}-\bar{x})}{\sum_{t=1}^{n}(x_t-\bar{x})^2} \tag{40}$$

Two versions of this test are performed: one set checks for any remaining autocorrelation in the residuals of the model, $r_j(\hat{\varepsilon}_t)$, and a second version checks for any remaining autocorrelation in the squared residuals of the model, $r_j(\hat{\varepsilon}_t^2)$. Both versions of the test are conducted for m equal to one, three, and five. Under the null hypothesis of a correctly specified model, the statistic Q_W is asymptotically distributed as $\sum_{k=1}^{m}\lambda_k\chi_k^2$ where χ_k^2 are independent chi-squared random variables with one degree of freedom. The weight term, λ_k, is a function of the information matrix of ϕ and θ, the parameters of the ARMA model in equation (32), and the weight term in equation (39), $(m-j+1)/m$.

In Fisher and Gallagher 2012, the authors also provide a test for the correct ARCH specification. The statistic has the following functional form:

$$L_W(b,m) = n\sum_{j=b+1}^{m}\frac{m-j+(b+1)}{m}[\hat{r}_j\left(\hat{\varepsilon}_t^2/h_t\right)]^2 \tag{41}$$

Here, b is the ARCH order from equation (38). Under the null hypothesis of a correctly specified model, the statistic $L_W(b,m)$ is asymptotically distributed as

$\sum_{k=1}^{m} w_k \chi_k^2$ where χ_k^2 are independent chi-squared random variables with one degree of freedom and weight term equal to $(m - j + b + 1)/m$.

To verify the choice of the conditional distribution, g, a Pearson's Goodness-of-Fit Test is used to compare the histograms of the standard residuals and the theoretical distribution. The test is run multiple times with different number of bins.

The fourth diagnostic check reported here checks for the presence of any asymmetric effects remaining in the fitted residuals. The Joint Sign-Bias Test of Engle and Ng 1993 is run with the following regression:

$$\hat{z}_t^2 = a + b_1 \mathbf{I}_{\{\hat{\varepsilon}_t < 0\}} + b_2 \mathbf{I}_{\{\hat{\varepsilon}_t < 0\}} \hat{\varepsilon}_t + b_3 \mathbf{I}_{\{\hat{\varepsilon}_t \geq 0\}} \hat{\varepsilon}_t + e_t \tag{42}$$

The dependent variable, $\hat{z}_t^2$, is the square of the standardized residuals. The coefficient b_1 tests for any significant asymmetric effects remaining in the residuals while b_2 and b_3 capture the magnitude of the asymmetric effects. The t-ratios of b_1, b_2, and b_3 are the test statistics for the sign-bias, the negative sign-bias, and the positive sign-bias, respectively. The joint test evaluates the significance of the three coefficients together using a LM test whose statistic follows an asymptotic chi-squared distribution with three degrees of freedom. The null hypothesis is for a well specified model with no asymmetric effects present.

Finally, due to a time series' sequential nature, the fitted ARMA-GARCH models are checked for the constancy of their parameters over the duration of the sample. The test used in this essay was proposed by Nyblom 1989. Our approach is to start with a simple GARCH(1,1) model with no mean equation (e.g. no intercept, AR, or MA components). If our tests indicate that the model is well specified, we stop there. If not, we continue to add additional parameters or try different functional forms until an acceptable fit is obtained.

Table 1 summarizes the marginal models of the exchanges rates under review in this paper. This includes the exchange rates between the US Dollar and the following three major global currencies: the European Euro, the Japanese Yen, and the British Pound. A major takeaway from Table 1 is that the GARCH processes for all three exchange rates changes from one estimation period to the next. From 2000-2009, the Euro/USD and Pound/USD exchange rates have asymmetric volatility responses while the Yen/USD exchange rate is symmetric. These specifications flip during the 2010-2018 timeframe. Also, the shape parameters (the degrees of freedom parameter from the t-Distribution) for each exchange rate are collectively lower in 2010-2018 than in 2000-2009, indicating slightly fatter tails for the 2010-2018 timeframe.

Table 2 summarizing the diagnostic checks for the ARMA-GARCH models shows that, overall, the chosen marginal models are well specified and capture the conditional mean and conditional variance of the underlying processes. The top part of the table records the p-values for the ARCH-LM, ARMA-LM, Goodness-of-Fit, and Sign-Bias tests. The bottom portion of the table records the value of test statistic for Nyblom's test for parameter constancy. One problematic area is the constant parameter assumption for the GARCH intercept term, ω. It is consistently rejected at the 1% significance level which causes the rejection of the joint parameter test as well. The only other test that rejects the null of a well specified model at the 5% significance level concerns the 2010-2018 Pound/USD series; the ARMA-LM test on the fifth lag using a degree of 2. Finding a model specification that satisfies this particular test proved insurmountable. A wide range of GARCH models and distributional assumptions were tested but none were able to provide a specification that passed all ARMA-LM tests.

5 Results

Tables 3 through 7 summarize the various copulas estimated on exchange rate pairs after their marginal ARMA-GARCH models are estimated. Table 3 records the Bayesian Information Criterion (BIC) across the three different time-varying models of Section 3 and the three different exchange rate pairs. The bolded values highlight the minimum BIC value for each exchange rate pair and model. The dependency measures for the bolded models are contained in Tables 4 while Tables 5 through 7 each record the dates that indicate the transition from one regime to another for each exchange rate pair. The sequential break model returns the day of each regime explicitly. For the smooth transition model the date implied by parameter c from equation (28) is reported. For the markov switching model, the reported regime switching dates are the days the implied state probability of equation (27) crosses the 0.5 value in either direction.

It's clear that the t-Copula is the preferred distributional model for all three exchange rate pairs. It outperforms the mixture Clayton and mixture Gumbel copulas across all three time varying models under consideration in this essay. The sequential break point model is the only model of the three that produces the number of optimal regimes endogenously and its outcome provides a good cross-check to using the BIC for selecting the number of regimes to specify for the markov switching and smooth transition models. For all three exchange rate pairs it finds the optimal number

Table 1: Marginal ARMA-GARCH Regressions

	2000-2009			2010-2018		
	Euro	Pound	Yen	Euro	Pound	Yen
ϕ_0	-0.00024	-0.00006	0.00013	–	–	0.00012
	0.00011*	0.00010	0.00012	–	–	0.00010
ϕ_1	–	–	–	–	0.00166	–
	–	–	–	–	0.02098	–
ω	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	0.00000**	0.00000	0.00000	0.00000	0.00000	0.00000
α_1	0.00000	0.052	0.032	0.028	0.032	0.038
	0.00111	0.003**	0.007**	0.004**	0.019	0.005**
γ_1	-0.065	-0.023	–	–	–	0.009
	0.006**	0.011*	–	–	–	0.012
α_2	0.044	–	–	–	0.009	–
	0.004**	–	–	–	0.017	–
γ_2	0.063	–	–	–	–	–
	0.004**	–	–	–	–	–
β_1	0.952	0.953	0.957	0.970	0.878	0.949
	0.000**	0.006**	0.007**	0.003**	0.005**	0.005**
β_2	–	–	–	–	0.074	–
	–	–	–	–	0.004**	–
shape	9.387	9.953	6.879	8.388	8.301	4.907
shape	1.683**	1.551**	1.230**	0.989**	0.852**	0.456**
Distribution	T-Dist	T-Dist	T-Dist	T-Dist	T-Dist	T-Dist
GARCH Spec.	gjr	gjr	std	std	std	gjr
ARMA	0, 0	0, 0	0, 0	0, 0	1, 0	0, 0
GARCH	2, 1	1, 1	1, 1	1, 1	2, 2	1, 1

Signif. Codes: 0 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1

Table 2: Exchange Rate Model Diagnostics

		2000 - 2009			2010 - 2018		
Test	Spec	Euro	Pound	Yen	Euro	Pound	Yen
ARCH-LM	ARCH Lag[1]	0.417	0.930	0.109	0.905	0.063^{+}	0.761
ARCH-LM	ARCH Lag[3]	0.813	0.785	0.262	0.819	0.238	0.478
ARCH-LM	ARCH Lag[5]	0.900	0.837	0.460	0.807	0.428	0.651
ARMA-LM	Lag[1] degree=1	0.738	0.266	0.403	0.744	0.998	0.967
ARMA-LM	Lag[2] degree=1	0.830	0.375	0.605	0.847	0.297	0.663
ARMA-LM	Lag[5] degree=1	0.832	0.489	0.875	0.892	0.377	0.829
ARMA-LM	Lag[1] degree=2	0.782	0.319	0.921	0.682	0.351	0.146
ARMA-LM	Lag[2] degree=2	0.862	0.478	0.163	0.382	0.489	0.241
ARMA-LM	Lag[5] degree=2	0.879	0.756	0.078^{+}	0.517	0.027^{*}	0.385
GOF	20bins	0.122	0.186	0.291	0.264	0.656	0.887
GOF	30bins	0.310	0.252	0.290	0.295	0.312	0.441
GOF	40bins	0.327	0.425	0.260	0.077^{+}	0.540	0.340
GOF	50bins	0.676	0.563	0.072^{+}	0.291	0.665	0.734
SIGN-BIAS	Joint Effect	0.471	0.090^{+}	0.513	0.591	0.781	0.097^{+}
SIGN-BIAS	Negative Sign Bias	0.376	0.161	0.647	0.302	0.457	0.168
SIGN-BIAS	Positive Sign Bias	0.188	0.675	0.173	0.532	0.540	0.562
SIGN-BIAS	Sign Bias	0.234	0.338	0.510	0.487	0.700	0.319
NYBLOM	ϕ_0	0.161	0.276	0.118	–	–	0.258
NYBLOM	ϕ_1	–	–	–	–	0.251	–
NYBLOM	ω	228.7^{**}	292.1^{**}	186.3^{**}	293.9^{**}	232.3^{**}	134.8^{**}
NYBLOM	α_1	0.076	0.232	0.308	0.196	0.243	0.066
NYBLOM	γ_1	0.131	0.238	–	–	–	0.114
NYBLOM	α_2	0.076	–	–	–	0.272	–
NYBLOM	γ_2	0.138	–	–	–	–	–
NYBLOM	β_1	0.093	0.228	0.198	0.173	0.240	0.105
NYBLOM	β_2	–	–	–	–	0.239	–
NYBLOM	shape	0.132	0.172	0.105	0.459^{+}	0.971^{**}	0.032
NYBLOM	Joint Test	621.4^{**}	784.7^{**}	759.4^{**}	598.5^{**}	687.9^{**}	524.7^{**}

Signif. Codes: 0 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
Note: The values in the top part of the table are p-values for the ARCH-LM, ARMA-LM, GOF, and Sign-Bias tests.
Note: The bottom part of the table records the value of the test statistic for Nyblom's Test for Constancy of parameters over time. Significant code indicate if the test static exceeds threshold value for a particular significance level. Threshold levels differ between test for a single parameter or a joint test:
Individual statistic: 10% = 0.35, 5%=0.47, 1%=0.75
Joint statistic: 10% = 1.69, 5%=1.90, 1%=2.35.

of regimes over the 2000 to 2018 time period to be four. The dates of the regime changes, in 2001, 2004, and 2007, are close in time as well.

When it comes to the markov switching models, the optimal number of regimes for each exchange rate pair is two. For pairs containing the Euro there is one regime with high overall association—as judged by Kendall's τ—with heavy tail dependence. The alternate regime has a reduced level of association and reduced tail dependence. One of the key features of the markov switching models used in this essay is their flexible nature. They can switch back-and-forth between their regimes on almost a daily basis, something that is not feasible with the smooth transition model due to its particular functional form. This characteristic of the model is easily seen in the markov switching columns of Tables 5, 6, and 5. The number of times the model iterates between the two regimes far exceeds the number of total regimes modeled by the other time-varying methods.

6 Conclusion

This essay explored a set of tools that can be used study financial time series and model their bivariate distributions. These tools include the ARMA-GARCH framework for modeling the conditional mean and conditional volatility process of the exchange rates studied in this essay. The ARMA-GARCH models allow the research to control for asymmetry in response to positive versus negative returns and periods of volatility clustering. The use of Copulas allows for flexible distributional assumptions including excessive leptokurtosis and asymmetric tail dependencies.

The utility of using these three tools in conjunction with each other was demonstrated on three major global currencies and their exchange rate fluctuations with the US Dollar from 2000 to 2018.

Table 3: BIC Values for Estimated Copula Models

		Markov Switching			Smooth Trans.			Seq. Break[1]		
Distribution	Regime	tCop	Gumbel	Clayton	tCop	Gumbel	Clayton	tCop	Gumbel	Clayton
Euro-Pound	1	-2634.1	-2643.2	-2556.0	-2634.1	-2643.2	-2556.0	-2634.1	-2643.2	-2556.0
	2	**-2750.0**	-2727.7	-2626.5	-2681.3	-2672.5	-2590.1	–	–	–
	3	-2719.2	-2676.0	-2565.1	-2671.9	-2654.3	-2566.1	–	–	–
	4	-2643.2	-2593.0	-2507.5	**-2712.0**	-2699.8	-2601.0	**-2748.1**	-2710.6	-2618.8
Euro-Yen	1	-713.5	-682.8	-638.7	-713.5	-682.8	-638.7	-713.5	-682.8	-638.7
	2	**-939.8**	-892.6	-868.5	-714.0	-686.3	-639.3	–	–	–
	3	-880.6	-825.3	-801.1	-772.5	-733.3	-701.4	–	–	–
	4	-805.5	-762.8	-743.2	**-853.8**	-819.0	-781.9	**-840.6**	-780.8	-746.4
Pound-Yen	1	-365.1	-309.4	-273.5	-365.1	-309.4	-273.5	-365.1	-309.4	-273.4
	2	**-496.9**	-424.1	-407.7	-416.0	-341.7	-312.8	–	–	–
	3	-470.2	-366.0	-365.6	**-481.5**	-347.9	-325.5	–	–	–
	4	-406.8	-274.2	-298.5	-473.2	-397.1	80.8	**-511.1**	-418.0	-387.5

This table records the Bayesian information criterion (BIC) for each model. The values in bold represent the minimum BIC value by exchange rate distribution and regime switching model.

[1] The researcher does not need to provide the number of regimes to the sequential break point model.

Table 4: Estimated Dependence Measures from Fitted Models

		Markov Switching			Smooth Trans.			Seq. Break		
Currency	Regime	τ	LTD	UTD	τ	LTD	UTD	τ	LTD	UTD
Euro-Pound	1	0.558	0.255	0.255	0.155	0.004	0.004	0.381	0.111	0.111
	2	0.368	0.059	0.059	0.972	0.972	0.972	0.527	0.226	0.226
	3	–	–	–	0.529	0.154	0.154	0.621	0.524	0.524
	4	–	–	–	0.385	0.072	0.072	0.412	0.099	0.099
Euro-Yen	1	-0.017	0.021	0.021	-0.039	0.000	0.000	-0.004	0.000	0.000
	2	0.378	0.078	0.078	0.393	0.076	0.076	0.320	0.089	0.089
	3	–	–	–	0.073	0.075	0.075	0.453	0.002	0.002
	4	–	–	–	0.310	0.186	0.186	0.180	0.139	0.139
Pound-Yen	1	0.030	0.041	0.041	0.102	0.023	0.023	0.020	0.004	0.004
	2	0.376	0.015	0.015	-0.584	0.014	0.014	0.258	0.069	0.069
	3	–	–	–	0.183	0.084	0.084	0.439	0.002	0.002
	4	–	–	–	–	–	–	0.073	0.071	0.071

This table summarizes the changes of tail dependency and Kendall's τ as estimated from the Markov Switching, Smooth Transition, and Sequential Breakpoint models. The values of Kendall's τ and tail dependence come from the models that minimize the BIC value. This is recorded in Table 4.

References

Ang, Andrew and Joseph Chen (2002). "Asymmetric correlations of equity portfolios". In: *Journal of Financial Economics* 63.3, pp. 443–494.

Bai, Jushan (1997). "Estimating Multiple Breaks One at a Time". In: *Econometric Theory* 13.3, pp. 315–352.

Csörgö, Miklós and Lajos Horváth (1997). *Limit theorems in change-point analysis*. Wiley Series in Probability and Statistics. Wiley. ISBN: 978-0-471-95522-1.

Dias, Alexandra and Paul Embrechts (2004). "Change-point analysis for dependence structures in finance and insurance". English. In: *Risk Measures for the 21st Century*. Ed. by Giorgio Szegö. United States: Wiley-Blackwell, pp. 321–335. ISBN: 978-0-470-86154-7.

– (2009). "Testing for structural changes in exchange rates' dependence beyond linear correlation". In: *The European Journal of Finance* 15.7-8, pp. 619–637. DOI: `10.1080/13518470701705579`. eprint: `https://doi.org/10.1080/13518470701705579`.

Dijk, Dick van and Philip Hans Franses (1999). "Modeling Multiple Regimes in the Business Cycle". In: *Macroeconomic Dynamics* 3.3, pp. 311–340.

Embrechts, Paul, Filip Lindskog, and Alexander McNeil (2003). "Chapter 8 - Modelling Dependence with Copulas and Applications to Risk Management". In: *Handbook of Heavy Tailed Distributions in Finance*. Ed. by Svetlozar T. Rachev. Vol. 1. Handbooks in Finance. Amsterdam: North-Holland, pp. 329–384. DOI: `https://doi.org/10.1016/B978-044450896-6.50010-8`.

Embrechts, Paul, Alexander J. McNeil, and Daniel Straumann (2002). "Correlation and Dependence in Risk Management: Properties and Pitfalls". In: *Risk Management: Value at Risk and Beyond*. Ed. by M. A. H.Editor Dempster. Cambridge University Press, pp. 176–223.

Engle, Robert F. and Victor K. Ng (1993). "Measuring and Testing the Impact of News on Volatility". In: *The Journal of Finance* 48.5, pp. 1749–1778. ISSN: 00221082, 15406261.

Fama, Eugene F. (1965). "The Behavior of Stock-Market Prices". In: *The Journal of Business* 38.1, pp. 34–105. ISSN: 00219398, 15375374.

Fernandez, Carmen and Mark F. J. Steel (1998). "On Bayesian Modeling of Fat Tails and Skewness". In: *Journal of the American Statistical Association* 93.441, pp. 359–371. ISSN: 01621459.

Fisher, Thomas J. and Colin M. Gallagher (2012). "New Weighted Portmanteau Statistics for Time Series Goodness of Fit Testing". In: *Journal of the American*

Table 5: Dates of Regime Change: Euro-Pound

Markov Switching		Smooth Trans.		Seq. Break	
Date	Regime	Date	Regime	Date	Regime
2000-01-04	1	2000-01-04	1	2000-01-04	1
2000-02-03	2	–	1	–	1
–	2	2000-12-06	2	–	1
2001-08-16	1	–	2	–	1
–	1	–	2	2001-10-11	2
2003-01-30	2	–	2	–	2
2003-10-17	1	–	2	–	2
–	1	–	2	2004-12-29	3
–	1	–	2	2007-07-13	4
–	1	2007-04-17	3	–	4
2007-09-11	2	–	3	–	4
2008-06-12	1	–	3	–	4
2008-11-11	2	–	3	–	4
2009-03-25	1	–	3	–	4
2009-08-26	2	–	3	–	4
2011-06-09	1	–	3	–	4
2012-12-31	2	–	3	–	4
–	2	2013-01-08	4	–	4
2017-12-13	1	–	4	–	4

This table records the dates of regime switching as indicated by each model. The sequential break model returns the day of each regime explicitly. For the smooth transition model the date implied by parameter c from equation (28) is reported. For the Markov Switching model, the reported regime switching dates are the days the implied state probability of equation (27) crosses the 0.5 value in either direction.

Table 6: Dates of Regime Change: Euro-Yen

Markov Switching		Smooth Trans.		Seq. Break	
Date	Regime	Date	Regime	Date	Regime
2000-01-04	1	2000-01-04	1	2000-01-04	1
–	1	–	1	2001-03-02	2
2001-06-11	2	–	1	–	2
–	2	2001-07-03	2	–	2
2001-12-12	1	–	2	–	2
2002-02-20	2	–	2	–	2
2003-08-01	1	–	2	–	2
2003-09-19	2	–	2	–	2
–	2	–	2	2004-07-02	3
–	2	–	2	2007-01-03	4
2007-05-29	1	–	2	–	4
–	1	2007-02-23	3	–	4
2008-02-13	2	–	3	–	4
2008-08-13	1	–	3	–	4
2009-06-04	2	–	3	–	4
2009-06-10	1	–	3	–	4
2009-09-08	2	–	3	–	4
2009-10-12	1	–	3	–	4
2009-12-10	2	–	3	–	4
2010-01-15	1	–	3	–	4
2010-07-19	2	–	3	–	4
2010-08-27	1	–	3	–	4
2010-09-07	2	–	3	–	4
2011-03-09	1	–	3	–	4
2011-10-20	2	–	3	–	4
2012-01-27	1	–	3	–	4
–	1	2013-04-26	4	–	4
2013-04-30	2	–	4	–	4
2013-11-08	1	–	4	–	4
2014-03-18	2	–	4	–	4
2016-06-08	1	–	4	–	4
2016-07-18	2	–	4	–	4
2017-04-18	1	–	4	–	4
2017-05-05	2	–	4	–	4
2018-05-16	1	–	4	–	4

This table records the dates of regime switching as indicated by each model. The sequential break model returns the day of each regime explicitly. For the smooth transition model the date implied by parameter c from equation (28) is reported. For the Markov Switching model, the reported regime switching dates are the days the implied state probability of equation (27) crosses the 0.5 value in either direction.

Table 7: Dates of Regime Change: Pound-Yen

Markov Switching		Smooth Trans.		Seq. Break	
Date	Regime	Date	Regime	Date	Regime
2000-01-04	1	2000-01-04	1	2000-01-04	1
–	1	–	1	2001-06-21	2
2001-10-11	2	–	1	–	2
2001-12-04	1	–	1	–	2
2002-04-09	2	–	1	–	2
2003-02-12	1	–	1	–	2
2003-03-24	2	–	1	–	2
2003-04-25	1	–	1	–	2
2004-02-12	2	–	1	–	2
–	2	–	1	2004-04-06	3
2006-12-22	1	–	1	–	3
–	1	–	1	2007-01-03	4
–	1	2007-01-08	2	–	4
–	1	2007-01-09	3	–	4
2007-04-30	2	–	3	–	4
2007-05-23	1	–	3	–	4
2008-06-10	2	–	3	–	4
2008-07-23	1	–	3	–	4
2010-10-04	2	–	3	–	4
2011-01-20	1	–	3	–	4
2011-11-22	2	–	3	–	4
2012-01-13	1	–	3	–	4
2013-04-29	2	–	3	–	4
2013-09-02	1	–	3	–	4
2013-09-04	2	–	3	–	4
2013-11-08	1	–	3	–	4
2014-10-27	2	–	3	–	4
2015-01-19	1	–	3	–	4
2015-02-09	2	–	3	–	4
2015-06-30	1	–	3	–	4
2015-09-24	2	–	3	–	4
2015-12-14	1	–	3	–	4
2016-12-23	2	–	3	–	4
2017-03-30	1	–	3	–	4

This table records the dates of regime switching as indicated by each model. The sequential break model returns the day of each regime explicitly. For the smooth transition model the date implied by parameter c from equation (28) is reported. For the Markov Switching model, the reported regime switching dates are the days the implied state probability of equation (27) crosses the 0.5 value in either direction.

Statistical Association 107.498, pp. 777–787. DOI: 10.1080/01621459.2012.688465. eprint: https://doi.org/10.1080/01621459.2012.688465.
Genest, Christian and Anne-Catherine Favre (2007). "Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask". In: *Journal of Hydrologic Engineering* 12.4, pp. 347–368. DOI: 10.1061/(ASCE)1084-0699(2007)12:4(347). eprint: https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%291084-0699%282007%2912%3A4%28347%29.
Ghalanos, Alexios (2020). *rugarch: Univariate GARCH models.* R package version 1.4-4.
Glosten, Lawrence R, Ravi Jagannathan, and David E Runkle (1993). "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks". In: *Journal of Finance* 48.5, pp. 1779–1801.
Hamilton, James D. (1989). "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle". In: *Econometrica* 57.2, pp. 357–384.
– (1994). *Time Series Analysis*. Princeton University Press.
Hansen, Bruce (1994). "Autoregressive Conditional Density Estimation". In: *International Economic Review* 35.3, pp. 705–30.
Joe, Harry (1997). *Multivariate Models and Multivariate Dependence Concepts.* Monographs on Statistics and Applied Probability. Chapman and Hall. ISBN: 9780367803896.
Kroner, Kenneth F. and Victor K. Ng (1998). "Modeling Asymmetric Comovements of Asset Returns". In: *The Review of Financial Studies* 11.4, pp. 817–844. ISSN: 08939454, 14657368.
Longin, Francois M. (1996). "The Asymptotic Distribution of Extreme Stock Market Returns". In: *The Journal of Business* 69.3, pp. 383–408. ISSN: 00219398, 15375374.
Longin, Francois and Bruno Solnik (1995). "Is the correlation in international equity returns constant: 1960-1990?" In: *Journal of International Money and Finance* 14.1, pp. 3–26.
Mandelbrot, Benoit (1963). "The Variation of Certain Speculative Prices". In: *The Journal of Business* 36.4, pp. 394–419. ISSN: 00219398, 15375374.
Manner, Hans and Olga Reznikova (2012). "A Survey on Time-Varying Copulas: Specification, Simulations, and Application". In: *Econometric Reviews* 31.6, pp. 654–687.
Markowitz, Harry (1952). "Portfolio Selection". In: *The Journal of Finance* 7.1, pp. 77–91. ISSN: 00221082, 15406261.
– (1959). *Portfolio Selection: Efficient Diversification of Investments.* Yale University Press. ISBN: 9780300013726.

Nelsen, R.B. (2007). *An Introduction to Copulas.* Springer Series in Statistics. Springer New York. ISBN: 9780387286785.

Nyblom, Jukka (1989). "Testing for the Constancy of Parameters Over Time". In: *Journal of the American Statistical Association* 84.405, pp. 223–230. ISSN: 01621459.

Öcal, Nadir and Denise R. Osborn (2000). "Business Cycle Non-Linearities in UK Consumption and Production". In: *Journal of Applied Econometrics* 15.1, pp. 27–43.

Patton, Andrew J. (2006). "Modelling Asymmetric Exchange Rate Dependence". In: *International Economic Review* 47.2, pp. 527–556. DOI: `https://doi.org/10.1111/j.1468-2354.2006.00387.x`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2354.2006.00387.x`.

Ramchand, Latha and Raul Susmel (1998). "Volatility and cross correlation across major stock markets". In: *Journal of Empirical Finance* 5.4, pp. 397–416.

Rodríguez, Juan (2007). "Measuring financial contagion: A Copula approach". In: *Journal of Empirical Finance* 14.3, pp. 401–423.

Rosenblatt, Murray (1952). "Remarks on a Multivariate Transformation". In: *The Annals of Mathematical Statistics* 23.3, pp. 470–472. ISSN: 00034851.