SAPIENZA
UNIVERSITÀ DI ROMA

New Insights on Hidden Markov Models for Time Series Data Analysis

Facoltà di Economia
Dottorato di Ricerca in Scienze Economiche e Statistiche (XXXV cycle)

Beatrice Foroni
ID number 1529955

Advisor
Prof.ssa Lea Petrella

Academic Year 2022/2023

New Insights on Hidden Markov Models for Time Series Data Analysis
PhD thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Version: February 26, 2023

Author's email: beatrice.foroni@uniroma1.it

Abstract

The goal of this thesis is to develop novel methods for the analysis of financial data by using hidden Markov models based approaches. The analysis focuses on univariate and multivariate financial time series, modeling interrelationships between financial returns throughout different statistical methods, such as graphical models, quantile and expectile regressions. The dissertation is divided into three chapters, each of them examining different classes of assets returns for a comprehensive risk analysis. The methodologies we propose are illustrated using real-world data and simulation studies.

Contents

List of Figures

List of Tables

Introduction and Overview

Hidden Markov models (HMMs) are mathematical representations of systems embedding a hidden finite-state first-order Markov chain, describing the transition behavior between a set of states over time. An HMM's primary premise is that the observation process is distorted by some "noise", concealing the latent state of the system and other unobservable informations. The latent state flexibility of these models can provide an efficient methodological framework to capture the complex dynamics of the data-generating process that might not be reproduced by simpler models, such as unobserved time-dependent heterogeneity, serial correlation, clustered behaviors and transitions between clusters. The birth of hidden Markov models can be traced back to the 1960s when Baum & Petrie (1966) introduced them as new statistical methods. The use of HMMs in an economic framework was pioneered in Hamilton (1989) and Hamilton (1990), where for the first time the structure of an autoregressive model with parameters driven by a hidden two-state Markov chain was introduced. In the last decade financial literature has seen rise the number of applications of HMMs. The work of Mergner & Bulla (2008) is one of the first to employ a latent Markov dynamic to investigate the time-varying behavior of financial systemic risk. De Angelis & Paas (2013), Bae et al. (2014), Nystrup et al. (2015, 2017), Giudici & Abu Hashish (2020) have contributed to the discussion on the capabilities of hidden Markov-based models to infer abrupt changes of volatility dynamic among diverse financial markets, as stock, bond, commodity and cryptocurrency markets. Multivariate hidden Markov settings have also been studied in Bernardi et al. (2017) and Maruotti et al. (2019), where, in the last one, a hidden semi-Markov setting to explicitly model the sojourn distribution for a time series of stock market returns was introduced.

In a multivariate financial data framework recently a particular interest has grown respect to the risk propagation issue, since investors, fund managers and regulators aim for an early identification of systemic risk to proactively engage measures to control financial stability. For this reason, network science has emerged as a useful tool for describing systemic risk propagation and spillovers, where interactions among random variables in a system can be represented in the form of graphs, whose nodes

represent the variables and whose edges show their interactions. In this context, Gaussian Graphical Models (GGMs) have received an enormous attention because they provide a simple method to model the pair-wise conditional correlations of a collection of variables, see for instance Lauritzen (1996), Lauritzen & Wermuth (1989), Whittaker (2009).

As it is well known, in the context of quantitative risk management, beside the analysis on interconnectedness, one of the main goal is the study of the dynamic of extreme occurrences, being of utmost importance for market participants and regulators. Since the seminal work of Koenker & Bassett (1978), quantile regression has represented a valid approach for modeling the entire distribution of returns while accounting for the well-known stylized facts, i.e., high kurtosis, skewness and serial correlation, that typically characterize financial assets. Quantile models have been extensively applied in finance and economics for estimating Value at Risk (VaR) and quantile-based risk measures (Engle & Manganelli 2004, White et al. 2015, Taylor 2019, Merlo et al. 2021). Quantile regression methods have also been generalized to account for serial heterogeneity. For example, Liu (2016) consider a quantile autoregression in which the parameters are subject to regime shifts determined by the outcome of a latent, discrete-state Markov process, while Adam et al. (2019) propose a model-based clustering approach where groups are inferred from conditional quantiles. One of the most relevant extension related to quantile regression is provided by the expectile regression (Newey & Powell 1987), which is a "quantile-like" generalization of the mean regression using an asymmetric squared loss function. Similarly to the former, the latter allows to represent the entire conditional distribution of a response variable and it possesses several advantages theoretically and computationally (Tzavidis et al. 2016, Alfò et al. 2017, Nigri et al. 2022). For this reason, expectile models have been implemented in several fields, especially in the context of risk management (Taylor 2008, Kim & Lee 2016, Bellini & Di Bernardino 2017). However, to the best of our knowledge, hidden Markov expectile regression models and their multivariate generalizations for both quantiles and expectiles have not yet been proposed in the literature.

With this thesis we aim to fill this gap, introducing and developing new insights and applications of HMMs in a risk management context using different classes of assets returns. The dissertation is divided into three chapters, organized as follows. In Chapter 1 we investigate the interconnections among and within the energy, agricultural, and metal commodities, operating in a risk management framework with a twofold goal. First, we estimate the Value-at-Risk (VaR) employing GARCH and latent Markov GARCH models with different error term distributions. The use of such models allows us to take into account well-known stylized facts shown in the time series of commodities as well as possible regime changes in their conditional

variance dynamics. We rely on backtesting procedures to select the best model for each commodity. Second, we estimate the sparse Gaussian Graphical model of commodities exploiting the Graphical LASSO (GLASSO) methodology proposed by Friedman et al. (2008*a*) to detect the most relevant conditional dependence structure among and within the sectors. Unlike other studies on commodity connectedness, which are based on the single assumption on the vector of observations (Diebold et al. 2017, Balli et al. 2019, Zhang & Broadstock 2020), we simultaneously address the assessment of the market risk combining hidden Markov dynamics, backtesting procedures and network analysis, building up a novel perspective in the literature of risk management. We apply our approach to the sample of twenty-four series of commodity futures prices over the years 2005-2022. The content of this chapter is based upon a joint work with Prof. Petrella L. and Prof. Giacomo Morelli, which has been recently published in Foroni et al. (2022). In Chapter 2 we develop a linear expectile hidden Markov model for the analysis of cryptocurrency time series in a risk management framework. The methodology proposed allows to focus on extreme returns and describe their temporal evolution by introducing in the model time-dependent coefficients evolving according to a latent discrete homogeneous Markov chain. As it is often used in the expectile literature, estimation of the model parameters is based on the Asymmetric Normal distribution. Maximum likelihood estimates are obtained via an Expectation-Maximization (EM) algorithm using efficient M-step update formulas for all parameters. We evaluate the introduced method with both artificial data under several experimental settings and real data investigating the relationship between daily Bitcoin returns and major world market indices. The methods and findings of this analysis are the results of a collaborative work with Prof. Petrella and Dr. Luca Merlo and the relative paper has been submitted to an international journal. Finally, Chapter 3 generalizes the approach considered in Chapter 2 to a multivariate setting. Specifically, the proposed methodology introduces multivariate hidden Markov regression models for estimating quantiles and expectiles of cryptocurrency returns using regime-switching copulas. The proposed approach allows us to focus on extreme returns and describe their temporal evolution by introducing time-dependent coefficients evolving according to a latent Markov chain. To model the time-varying dependence structure of returns, we consider elliptical copula functions defined by state-specific parameters. Maximum likelihood estimates are obtained via an EM algorithm using efficient M-step update formulas for all parameters. The goodness of the method is evaluated by means of simulation studies, and the empirical analysis investigates the relationship between daily returns of five cryptocurrencies and major world market indices. This work is the result of a collaborative work with Prof. Petrella and Dr. Luca Merlo, and it has been submitted to an international journal for publication.

Chapter 1

The Network of Commodity Risk

1.1 Introduction

The financialization of commodities (Tang & Xiong 2012; Cheng & Xiong 2014; Basak & Pavlova 2016) has drawn the attention of risk managers and financial institutions on the propagation of the commodity risk that arises from the fluctuations of commodity future price values (Joëts 2015; Marvasti & Lamberte 2016). Commodities are actively traded in financial markets and have been largely used for hedging purposes (Kat & Oomen 2006). However, market volatility makes commodity prices vulnerable to highly correlated shocks (Diebold et al. 2017) creating significant business challenges that affect financial performances, exacerbate well-known spillovers effects among commodities, and tight credit availability.

Hidden severe consequences affect the economic system as well, especially in those countries where commodities are heavily employed as raw materials (Crude Oil, Gasoline, Natural Gas, Copper, Aluminium, and agricultural commodities) in the industrial sector. Indeed, UNCTAD (2019) reports that over the last two decades 67% of developing countries has been relying on commodities, a percentage that rises to 80% when considering only the least developed countries. Therefore, a big concern for risk managers and policy-makers becomes the monitoring of the propagation of the commodity risk in commodity markets which requires the development of new operational approaches (Aven 2016; Battiston et al. 2017; Giampietro et al. 2018). The understanding of the propagation of the commodity risk through the financialization of commodities requires to be detected designing a framework that accounts for the specific contribution of the single commodities to the market risk. This is the aim of our work. The relevance of the issue is particularly highlighted by the role of the commodity risk within the regulatory framework. The Basel

Accords establish a minimum capital standard to cover the risk of holding or taking positions in commodities and impose each bank subject to capital charges for market risk to monitor and report the level of commodity risk against which a capital requirement is to be applied (Basel Committee on Banking Supervision 2016). Over the past decades, rich literature has flourished to propose valuable instruments for measuring and quantifying such risk. The most employed market risk measure is the Value-at-Risk (VaR), defined as the worst expected loss of an asset or a portfolio given a certain confidence level and over a specific time period (Jorion 2006). It is a crucial component of risk management when designing and monitoring an appropriate modeling framework able to quantify commodity price risk exposure, avoiding unexpected large losses (Aven 2016). In this paper, we propose a framework with a twofold risk management goal: i) forecasting commodity risks and spillovers to identify and understand the factors that drive commodity markets and ii) capturing the impact of contagion of such risk on the stability of the financial system. We combine an econometric and statistical set-up where several models are compared and backtested to find the one that better estimates the VaR for each commodity futures returns. The network is built on the residuals of the models chosen according to a risk management approach. The network approach guides decision-makers in the field through the investigation of the extent to which uncertainty in commodity prices affects the practical transmissions of the commodity risks (Nguyen et al. 2020). More specifically, we carry out the first task estimating the VaR of the commodities through GARCH and Markov-switching GARCH (MS-GARCH) models with different distribution of the innovations. The choice of GARCH-type models accommodates the typical stylized facts of commodity time series such as volatility clustering, skewness, kurtosis (Das & Sundaram 1997; De Luca et al. 2006; Wilhelmsson 2006; Adcock et al. 2015), and regime changes in the conditional variance dynamics. We perform model selection relying on targeted tests procedures and evaluate the model that outperforms the others from a risk management point of view, i.e. from a VaR forecasting perspective (Laporta et al. 2018). To do this, we use three backtesting procedures: the Unconditional Coverage (UC) test of Kupiec (1995), the Conditional Coverage (CC) test of Christoffersen (1998), and the Dynamic Quantile (DQ) test of Engle & Manganelli (2004). We refer to Masala (2021) for a recent survey about the role of backtesting procedures in commodity portfolios. To contemplate the stylized facts of commodity returns, we consider various specifications in the model selection procedure such as Normal GARCH, Skewed Normal GARCH, Student's-t GARCH, Skewed Student's-t GARCH, Generalized Error GARCH, Skewed Generalized Error GARCH, Normal MS-GARCH, Skewed Normal MS-GARCH, Student's-t MS-GARCH, Skewed Student's-t MS-GARCH, Generalized Error MS-GARCH, Skewed Generalized Error MS-GARCH, thus empowering our framework with the

flexibility that the policy-maker needs. The second task deals with the estimation of the interconnections among and within the commodity sectors considered. To explore this issue we use a graphical model approach which is an intuitive way of representing and visualizing the relationships among many variables. In particular, to manage the interconnection structure of commodity markets, we exploit the Graphical LASSO (GLASSO) methodology proposed by Friedman et al. (2008*b*) to estimate a sparse Gaussian Graphical model and detect the strongest conditional dependencies among the commodities. Since GLASSO relies on the assumption of gaussianity, we build a Gaussian copula with marginals obtained from the residuals of the best model evaluated under the backtesting criteria. Moreover, to synthesize the information contained in the graphical model, we compute the eigenvector centrality measure that shows the most relevant commodities within the network structure in terms of the influence of each node in the graph. This allows us to guide decision makers (Nguyen et al. 2020) ranking the commodities in the network according to their importance in the propagation of the commodity risk. We collect the future prices of twenty-four commodities belonging to the commodity sectors Agriculture, Energy, and Metals over the sample period that spans from October 3, 2005 to March 25, 2022. Our findings show that the MS-GARCH outperforms GARCH models in 75% and 58% of the cases before and after the Covid-19 pandemic, respectively. From a graphical point of view, the network analysis reveals that commodities are overall densely connected and the Covid-19 shock hits such connectedness reducing the density of the graph by 18%. The degree of connectedness between the commodity sectors is also affected. Inter-sectorial linkages are severely weakened as opposed to those among commodities within the same sectors. In fact, before Covid-19 Coffee and Soybean Oil present the maximum rate of connections with the Energy sector coherently with Myint & El-Halwagi (2009), **?**, and Al-Maadid et al. (2017) where the spillover effect between Coffee and the energy sector results even increased after the global financial crisis in 2008. With the outbreak of Covid-19, these patterns have been disentangled. Concerning the specific roles of the commodities in the network, Soybean Oil occupies the most central node in the graph only after the pandemic while Natural Gas was the most connected until 2018 (in line with Ergen & Rizvanoghlu 2016). Remarkable changes are also encountered in the safe-haven commodities. Before Covid-19, they were Heating Oil, Soybean Meal, and Gold and after the pandemic, they become Natural Gas UK, Gold, and Heating Oil. In a similar spirit to our paper, other studies have investigated commodity connectedness (Diebold et al. 2017; Zhang & Broadstock 2018; Balli et al. 2019). However, their modeling frameworks are characterized by a common distributional assumption for the commodities included in the analysis and do not simultaneously address the assessment of the market risk. Indeed, the novelty of our approach lies in the

combination of backtesting procedures, Gaussian Copula, and GLASSO estimation approach for the network. Unlike traditional network estimation based on the single assumption on the vector of observations, we build the network of commodity risk basing on the residuals of GARCH-type models with underlying distribution selected for each commodity through the backtesting procedures thus accommodating for well-known stylized facts. The dependence structure of the residuals is captured by the Gaussian Copula. Centrality measures retrieved from the sparse estimated network are exploited to guide decision makers choices through commodity risk management. To the best of our knowledge, this is the first paper that addresses this issue. The rest of the paper is organized as follows. Section 1.2 discusses the major strands of the literature to which this paper contributes. Section 1.3 provides a brief outline of the employment of the GARCH-type models and the backtesting procedures for model selection. Section 1.4 presents the GLASSO model. Empirical results are reported in Section 1.5 and the conclusions are in Section 3.6.

1.2 Literature Review

The effects of commodity price fluctuations on the macro-economy have been pioneered in Hamilton (1983), and the connection with economic growth has been a fruitful thematic in the financial literature (Deaton 1999; Browne & Cronin 2010; Kilian & Vigfusson 2011; Umar & Spierdijk 2011; Cevik & Sedik 2014; Chen et al. 2014; Baumeister & Kilian 2016; Charfeddine et al. 2020, Zaremba et al. 2019; Zaremba, Umar, Mikutowski et al. 2021). Despite the key informational role of commodity futures in addressing the monetary policy (Awokuse & Yang 2003; Hess et al. 2008) the exposure of commodity price fluctuations to macro risk has been hard to price (Roache 2008) due to the strong inter-sectorial dependencies, first documented in Pindyck & Rotemberg (1990) and Ciner (2001). A rich literature has then focused on the spillover effects between oil price and financial markets (Kang et al. 2017), and oil price and other commodities (Baffes 2007; de Nicola et al. 2016) such as precious metals (Ewing & Malik 2013; Rehman et al. 2018), agricultural (Du et al. 2011; Nazlioglu 2011; Nazlioglu et al. 2013), energy (Reboredo 2015; Ferrer et al. 2018; Tiwari et al. 2019), and, more recently, the impact of climate related variables on the co-movements of commodity prices that affect the stability of the financial system (Flori et al. 2021). Besides, commodity price behavior shows small trends and big variability that affects market preferences also in the long-run (Cashin & McDermott 2002; Tiwari et al. 2021; Christodoulakis 2020). Spillovers effects have particularly intensified since 2004, the onset of the financialization of commodity markets (Tang & Xiong 2012; Cheng & Xiong 2014; Henderson et al. 2015; Basak & Pavlova 2016). From that moment onward, commodities have been considered among the likely

sources of financial distress due to the centrality of the role acquired. Gradually, they have revealed to be responsive to macro-economic shocks and to investors sentiment (Smales 2014; Gelos & Ustyugova 2017; Ramiah et al. 2019; Umar, Gubareva & Teplova 2021) and to be strongly connected to widely spread financial instruments (Malik & Umar 2019; Naeem et al. 2020; Zaremba, Umar & Mikutowski 2021). Major operational and management implications caused by the financialization of commodities spring up in the uncertainty of the decision-making processes for the related industries. Examples concern the role of supporting the management of refinery operations and productions of oil and gas, or the management generator operations and the supply chain due to the non-storable nature of the production of energy, as discussed in Andriosopoulos & Nomikos (2014), Gabrel et al. (2014), Joëts (2015), Aven (2016), and Nguyen et al. (2020). Such relevance of commodities from various perspectives of the financial system has encouraged the assessment of the market risk for commodity markets (see, among others, Pilipovic 2007). Giot & Laurent (2003) introduce the use of VaR to compute the contribution of the commodities to the market risk and find that for market participants trading short positions the risk arises from an increase in commodity prices, while for long positions it is given by a price drop. Marimoutou et al. (2009) apply VaR to the oil market and Aloui & Mabrouk (2010) study VaR estimations of energy commodities documenting financial stylized facts such as long-memory, asymmetry, and fat tails. Laporta et al. (2018) investigate the selection of VaR for energy commodities, whereas Algieri & Leccadito (2017) and Ji, Bouri, Roubaud & Shahzad (2018) propose a (Co)VaR based model to study risk spillovers between energy and non-energy commodity markets. Shen et al. (2018) integrate VaR estimation in a Vector Autoregression (VAR) to evaluate the risk transmission channel in energy markets. An alternative approach to address systemic risk in a VaR set-up is through copula (Mensi et al. 2017). VaR forecasting may help risk managers and regulators to evaluate the exposure to unexpected loss and consequently calibrate the overall riskiness of financial markets. Besides VaR computation, many studies have delved into the analysis of the connectedness among commodity markets (Diebold et al. 2017; Balli et al. 2019; Umar, Jareño & Escribano 2021; Umar et al. 2022) and between commodity and financial markets. Ji & Fan (2016) propose a graph analysis of the evolution of the world crude oil market whereas the works of Diebold et al. (2017), Zhang & Broadstock (2018), and Balli et al. (2019) derive the connectedness of commodities from the Diebold & Yılmaz (2014) forecast-error variance decomposition matrix of a vector autoregressive (VAR) model. In Diebold et al. (2017), the VAR is constructed from the range-based realized volatility of Garman & Klass (1980), whereas in Balli et al. (2019) it comes from the commodity uncertainty index proposed in Chuliá et al. (2017) that builds on the residuals of a generalized dynamic factor model. These studies

have tackled several relevant issues in the analysis of the mechanisms of commodity markets[1]. However, the major research questions concerning the financialization of commodities, commodity connectedness, and the assessment of the market risk remain separately addressed. In particular, in the field of the study of commodity connectedness, the modeling framework of the previous literature exploits common distributional assumptions for the commodities while commodity risk management requires models that account for the structural differences between the commodities to gauge the different risk exposures. Hence, for a robust detection of the major risk transmitters in commodity markets, methodologies must be integrated with new approaches that consider the stylized facts of the single commodities. For this reason, we propose a framework that merges commodity connectedness and modeling selection according to market risk criteria.

1.3 Model Specifications

It is well-known that the time series of commodities show most of the stylized facts detected in financial markets such as skewness, kurtosis, and volatility clustering. Moreover, recent studies have shown that the variance process often exhibits regime changes (Haas et al. 2004) and that ignoring this feature affects the precision of the volatility forecast (Danielsson 2011). Throughout the paper, we consider the GARCH(1,1) and MS-GARCH(1,1) models with different conditional distribution $\mathcal{D}_\Theta(\cdot)$ to account for stylized facts. Overall, our framework includes 12 model specifications, recovered as a combination of:

- the conditional variance specification: GARCH(1, 1) and MS-GARCH(1,1);
- the choice of the conditional distribution $\mathcal{D}_\Theta \in \{$norm, snorm, std, sstd, ged, sged$\}$.

We detail the description of the model specification in the Appendix. The choice of the conditional variance specifications are supported by the works of Bollerslev et al. (1992), Sadorsky (2006), Huang et al. (2009) and, especially, Hansen & Lunde (2003). Moreover, GARCH(1,1) models have proved good fitting performance on commodities (Laporta et al. 2018). We remark that the specifications considered are only intended to provide an example of design of policy for commodity markets that simultaneously addresses the commodity interdependencies and the exposure to the market risk. Among the conditional distributions, we include the standardized skewed version of each model implemented via the Fernández & Steel (1998) transformation.

[1]We refer to Tables 1.4, 1.5, 1.6, 1.7, 1.8 in the Appendix for a global summary of the works on commodity markets.

We denote the standardized Skewed Normal, the Skewed Student's-t, and the Skewed Generalized Error Distribution by "snorm", "sstd", and "sged", respectively. For the MS-GARCH specification, we focus on double-regime MS-GARCH models accounting for low and high volatility levels, thus the scale and asymmetry parameters vary with the regimes. In line with the first goal of the paper, we consider a wide range of models and use a risk management approach that selects the best models to accurately predict future risks, especially in the case of volatility models (Christoffersen & Diebold 2000). We select the model that provides the most reliable forecast of the VaR performing backtests (Christoffersen 2010).

1.4 Sparse Gaussian Graphical Model

The second goal of the paper is to study the interdependence among and within the commodity sectors. To achieve this purpose, we rely on the Gaussian Graphical Lasso (GLASSO) methodology proposed in Friedman et al. (2008*b*). GLASSO allows us to build an undirected Gaussian graphical model and perform a network representation of the connections of the *commodity risks* where only the most relevant intra- and inter- sectorial linkages are highlighted. This is conveniently accomplished estimating a sparse conditional dependence structure among the commodities. That is, we estimate the inverse Gaussian covariance matrix, $\Omega = \Sigma^{-1}$, where the zero off-diagonal elements correspond to a pair of commodities returns that are conditionally independent (Hastie et al. 2015). More specifically, GLASSO builds on Tibshirani (1996) where a penalized maximum likelihood problem shrinks to zero some coefficients through a L_1-norm penalty term as follows:

$$\Omega^* = \arg\max_{\Omega} \log(\det \Omega - tr(\Sigma\Omega) - \rho \|\Omega\|_1), \tag{1.1}$$

where $tr(\cdot)$ denotes the trace operator and $\|\Omega\|_1$ the L_1-norm that can be calculated as the sum of the absolute values of the elements of Ω. The parameter ρ controls for the size of the penalty and it determines the number of zeros in the sparse precision matrix Ω: a higher (lower) value is responsible for a more (less) sparse matrix. Like most of the shrinking methodologies, the right choice of the penalization parameter ρ is fundamental to obtain a reliable selection. To estimate the optimal value of ρ, we minimize the Extended Bayesian Information Criterion (EBIC, Chen & Chen 2008) which has been shown to work particularly well in retrieving the true network structure (Foygel & Drton 2010, Barber et al. 2015) and it is a computationally efficient alternative to cross-validation (Yuan & Lin 2007). The criterion is indexed by the hyperparameter $\gamma \in [0, 1]$. Typical values of γ are 0, 0.5 or 1 with values closer to 1 leading a stronger penalization. For this reason, in the empirical application of this paper we choose $\gamma = 1$. As stated above, GLASSO

relies on the assumption of Gaussianity introduced through the Gaussian copula. In particular, the copula approach provides the framework to model multivariate associations from the univariate distributions of the observed variables. In the case of the d–random vector, $X = (X_1, \ldots, X_i, \ldots, X_d)$, with marginal cumulative distributions $F_i(x_i) = P(X_i < x_i)$, we can define the joint cumulative distribution function (cdf) as $F(x) = P(\cap_{i=1}^{d} X_i < x_i)$. In many cases, the margins of the cdf are relatively easy to describe, but an explicit expression of the joint distribution may be difficult to obtain. When $X \sim N_d(\mu, \Sigma)$ is a Gaussian random vector then its copula is called Gaussian copula. Denoting $u_i \equiv F_i(x_i)$, the Gaussian copula is defined by the cdf $C^{Ga}(u_1, \ldots, u_i, \ldots, u_d) = P(\cap_{i=1}^{d} \Phi(X_i) \leq u_i) = \mathbf{\Phi}(\cap_{i=1}^{d} \Phi^{-1}(u_i))$ where $\Phi^{-1}(\cdot)$ is the univariate standard Gaussian quantile function and $\mathbf{\Phi}(\cdot)$ is the d-variate Gaussian cdf with mean 0 and covariance matrix Σ.

Network Metrics

Network metrics are used to synthesize the information contained in a graphical model. In this section, we briefly introduce the definitions and the metrics that we use to detect the position of a commodity within the network. We define an undirected graph as an ordered pair of two disjoint sets (V, E), where V is the set of vertices and E is the set of edges, consisting of pair of elements taken from V. We denote the number of vertices with $n = |V|$ and the number of edges with $m = |E|$. The density, D, of the graph is given by the ratio between the number of edges and the number of possible edges:

$$D = \frac{2m}{n(n-1)}. \tag{1.2}$$

Vertices i and j are adjacent if the undirected edge between i and j is in the set E, and a line connects them in the diagram of the graph. The matrix representation of such a graph is obtained via the adjacency matrix, A_G, of the inverse covariance matrix, Ω. The single element in $A_G = (a_{ij})$ is $a_{ij} = 1$ if the corresponding element of the inverse covariance matrix is positive, $a_{ij} = 0$ if the corresponding element is zero. Hence, the graph contains an edge that links two vertices i and j if and only if $a_{ij} = 1$. A simple yet fundamental metric is the degree of a node defined by $k_i = \sum_j a_{ij}$, which measures the number of neighbors of the node. Centrality measures are crucial metrics used in the network topology to highlight nodes that occupy critical positions in the graph. For instance, the eigenvector centrality, or Gould's index of accessibility (Gould 1967), indicates which are the most geographically central and important nodes, and it has been exploited in financial applications to capture the capacity of an agent to cause systemic risk (Billio et al. 2012). It builds on the eigenvector assigned to the leading eigenvalue of the adjacency matrix to assign a relative score to

the nodes, depending on how connected they are to the rest of the network. A metric that is strictly related to the eigenvector measure is the eigenvector community structure (Newman 2006), which allows us to create a subgraph starting from a group of vertices densely connected, linked with other groups of vertices through sparse connections. The eigenvector community structure depends on the spectrum of the modularity matrix B, with elements defined as:

$$b_{ij} = a_{ij} - \frac{k_i k_j}{2m}. \tag{1.3}$$

a_{ij} are the elements of the adjacency matrix and $\frac{k_i k_j}{2m}$ is the number of edges between vertices i and j if edges are placed at random, where k_i and k_j are the degrees of the vertices, and $m = \frac{1}{2}\sum_i k_i$ is the total number of edges in the network. The algorithm to compute the eigenvector community structure calculates the leading eigenvector of the *modularity matrix* and divides the vertices into two groups according to the signs of the elements in this vector. The values of the leading eigenvector assess the importance of each vertex in its community: a larger (smaller) value corresponds to a more (less) central member.

1.5 Empirical Results

1.5.1 Data Description

We collect data of twenty-four time series of commodity futures prices from Bloomberg over the period that spans from October 3, 2005 to March 25, 2022 for a total of 4300 observations. The commodities in the sample belong to the commodity sectors Agriculture, Energy, and Metals and are divided as follows:

- Agriculture: Coffee (KC1), Oats (O1), Soybeans (S1), Wheat (W1), Cocoa (CC1), Corn (C1), Rough Rice (RR1), Cotton (CT1), Sugar (SB1), Soybean Oil (BO1), Soybean Meal (SM1), and Orange Juice (JO1).
- Energy: Gasoline (XB1), Heating Oil (HO1), Low Sulfur Gasolio (QS1), Natural Gas (NG1), Ethanol (DL1), WTI Crude Oil (CL1), and Natural Gas UK (FN1).
- Metals: Gold (GC1), Silver (SI1), Palladium (PA1), Copper (HG1), and Zinc (LX1)

Daily returns with continuous compounding are calculated taking the logarithm of the difference between closing prices in consecutive trading days and then multiplied by 100[2]. In Table 1.3 we report the summary statistics of the log-returns of the

[2]Since daily returns are generally small in monetary units, they have been multiplied by 100 to avoid numerical errors in computer programs.

commodities as in Laporta et al. (2018) before and after the outbreak of Covid-19. The distribution of the returns for each commodity displays fat tails and serial correlation. The Jarque-Bera test significantly rejects the normality behavior of daily returns, the ARCH Lagrange Multiplier and the Augmented Dickey-Fuller tests suggest the presence of autoregressive conditional heteroskedasticity and the absence of unit roots, respectively. Ethanol and Natural Gas are the only commodities with negative average returns and show high standard deviation. The maximum, however, is reached by WTI Crude Oil with 20.7. The distributions of the commodity returns generally exhibit negative skewness with Natural Gas, Natural Gas UK, Wheat, and Coffee as the only exceptions. The returns of these four commodities are then featured by extreme positive values. The impact of Covid-19 is captured in the higher kurtosis and standard deviation which are amplified in response to the growing level of uncertainty in the market. Besides, the distribution of the returns of the commodities tends to be more negatively skewed indicating the stronger propensity to undergo high losses.

1.5.2 Backtesting Results

We perform backtests over 2699 observations using 1600 days as estimation window. The width of the rolling window allows us to obtain significant estimates of the parameters while providing a trustworthy picture of the fluctuations in the market. Also, an amplitude of six years is sufficient to capture the evolution and major regime changes in the returns. In particular, we include the entire period of the recession in the United States that began in December 2007 and ended in June 2009 according to the National Bureau of Economic Research[3]. For each time series of commodities, we fit GARCH and MS-GARCH with conditional distribution $\mathcal{D}_\Theta \in \{$norm, snorm, std, sstd, ged, sged$\}$. The computational analysis is conducted using the software R and the package "MSGARCH". We point out that the introduction of GARCH models is only intended to accommodate the different behaviors of commodity returns within a risk management perspective based on the VaR predictability of the model. We choose the best model out of the 12 models considered according to the backtesting procedures described in the Appendix. We backtest each GARCH-type estimated VaR using the Unconditional Coverage (UC) test of Kupiec (1995), the Conditional Coverage (CC) test of Christoffersen (1998), and the Dynamic Quantile (DQ) test of Engle & Manganelli (2004). We are concerned with the downside risk at two confidence levels: 95% and 99%. The best models are selected choosing among those that show a p-value higher than the 5% significance level in at least two out of three tests. To determine the final model, we pick the one with the highest

[3]The official arbiter of U.S. recessions.

p-value relative to the DQ test on the 95-th quantile level. This choice is motivated by a trade-off between the risk management perspective and the need to have an adequate number of observed violations of the estimated quantiles. Backtesting results are represented in Tables 1.9, 1.10, 1.11, 1.12 1.13, 1.14, 1.15 and 1.16. Overall, many of the model specifications succeed in forecasting the returns volatility of the commodities. On the one hand, we find no prevalence for the asymmetric distributions over the symmetric ones. On the other hand, we detect a considerable prevalence of the Markov-switching specification which feature 14 commodities out of 24 (Table 1.1). This result contributes, among others, to the findings in Bulla & Bulla (2006) and Ardia et al. (2018), that show that the Markov-switching specification better captures the breaks in the dynamics of the volatility of financial returns. The effect of Covid-19 is found in the reduced number of optimal MS-GARCH models. Thus, the shock of the pandemic has affected the propensity of regime changes of the volatility of the commodities to describe the evolution of VaR exceedances.

Commodity	Distribution	Regime
Gold	GED	MS-GARCH
Silver	Student's-t	MS-GARCH
Copper	Skewed GED	MS-GARCH
Palladium	Student's-t	MS-GARCH
Zinc	Gaussian	GARCH
WTI Crude Oil	Student's-t	MS-GARCH
Heating Oil	Skewed Gaussian	GARCH
Low Sulfur Gasolio	Skewed GED	MS-GARCH
Natural Gas	Skewed Student-t	GARCH
Gasoline	Gaussian	GARCH
Natural Gas UK	GED	GARCH
Ethanol	Skewed Gaussian	MS-GARCH
Corn	Student's-t	MS-GARCH
Oats	Student's-t	MS-GARCH
Rough Rice	Student's-t	MS-GARCH
Soybeans	Student's-t	GARCH
Wheat	Skewed Gaussian	MS-GARCH
Cocoa	Student's-t	GARCH
Cotton	GED	GARCH
Coffee	GED	GARCH
Sugar	Gaussian	MS-GARCH
Soybean Oil	Gaussian	MS-GARCH
Soybean Meal	Skewed Student's-t	MS-GARCH
Orange Juice	Student's-t	GARCH

Table 1.1. Best models for the commodities selected according to the backtesting criteria. Sample period: October 3, 2010 - March 25, 2022.

1.5.3 The Network of Commodity Risks

We turn the analysis to the estimation and discussion of the network structure among the commodities in the sample. To achieve the estimation of the graph, we fit the Gaussian Copula on the residuals obtained from the GARCH-type model selected according to the the backtesting results. Therefore, the marginals of the Gaussian Copula are given by the distribution of each series of residuals. Then, we estimate the tuning parameter for the GLASSO minimizing the EBIC. To quantify the degree of connectedness within the graph and among specific underlying clusters, we compute the network metrics described in Section 1.4. Figure 1.1 shows *the network of commodity risks*. The computational analysis is conducted using the software R and the package "GLASSO". The graph is estimated using the optimal tuning parameter $\rho = 0.0158$ and the size of the nodes in Figure 1.1 is proportional to the eigenvector centrality score. The density $D = 0.61$ highlights the strong interconnections in the network. This result is explained in part because of the well-known spillover effects (Du et al. 2011, Nazlioglu et al. 2013) originated from the dependence that links commodities to cycles of production and consumption, and in part because of the effects of the financialization of commodities. The graphical representation of the results has brought relevant information to the analysis of the network of commodity risk, making easy the interpretation of results, and strengthening previous literature in this field. For instance, the safe-haven role of Gold is immediately captured. In fact, it represents one of the least central nodes in the network. This implies that it has a poor dependence relation with the other commodities and therefore it can be a good investment in anticipation of high volatility periods. Copper is the first central metal commodity in the graph. It is a cheap and plentiful metal, among the most traded in the world markets, and in the years has seen its consumption rise, especially because of building construction and electronic products. Palladium has seen its consumption rise too. Their centrality, together with that of Coffee, provides useful information to investors in the financial markets and serves as a barometer for the stability of the financial markets of large parts of Asia and South America. In Table 1.17 we report the relative connections for each commodity defined as the ratio between the number of active connections that the commodity has with a certain sector and the maximum number of possible connections with that sector. For example, Gold has 3 connections with the Metal sector, dividing by the maximum number of possible connections of the metal sector, i.e. 4, we get the value of 0.75. Most of the commodities exhibit the highest connection with the corresponding sector. In particular, this is the case of Copper and Palladium and Silver for the Metal sector, WTI Crude Oil, Low Sulfur Gasoil, and Natural Gas for the Energy sector. It is interesting to notice that some commodity shows particular

rates of connection with other sectors. It is worth noting that some commodity show particular rates of connection with other sectors. For example, in the Metals sector, Palladium reports that connection rate of 0.75 with the Agriculture sector. Regarding the Energy sector, Low Sulphur Gasolio shows the maximum rate of connection with the Metals sector (0.80). Concerning the Agricultural sector, Oats, Cocoa, Coffee, and Soybean Oil have high connection rates with other sectors. The high rate of connections among Coffee and the Energy sector is coherent with Al-Maadid et al. (2017) where the authors find the spillover effect between Coffee and the Energy sector to be even increased during the post-crisis period. Soybean Oil presents the maximum connection rate with the Energy sector. This confirms a result found in Myint & El-Halwagi (2009) and Dunis et al. (2009). Still, Biodiesel production in the U.S is based predominantly on the use of Soybean Oil, precisely for the 82% (Ajanovic 2011).

In Table 1.18 the connection rate between the three considered sectors is shown. We divide the total number of connections that an entire sector has with another sector by the maximum number of possible connections between the two sectors considered. For example, there are 14 connections between metals and energy sectors, dividing by the maximum number of possible connections, i.e. 35, we get a ratio of 0.40. The highest connection rates lie on the main diagonal. Interestingly, the pandemic has strengthened the relationships between the commodities belonging to the same sector and weakened the inter-sectorial linkages. Before Covid-19, agriculture and energy commodities exhibited strong dependence also supported by other studies (Barbaglia et al. 2016; Tyner 2010; Rezitis 2015). Such a link can be explained as a cause of the introduction of biofuels, that have intensified the ties between the sectors, as well as the need for efficient transport for perishable commodities like the agricultural ones. The degree of connectedness between the energy and agricultural commodity market, and between agricultural and metals, highlights the great effect produced by the financialization of commodity markets. The outbreak of Covid-19 has remarkably diminished the strength of inter-sectorial relationships. One hypothesis for this lies on the fact that during normal times trades are more liberalized while during periods of market turmoils the level of trust of investors is impaired. Moreover, in the specific event of the Covid-19 pandemic, the demand and supply economic chain has majorly hit with the abrupt interruptions of most economic activity therefore concerning the classes of commodities that are employed in the related industrial processes. Furthermore, we sort the considered commodities in terms of eigenvector centrality identifying the group of the “most connected” commodities. Table 1.2 shows the eigenvector centrality scores. Soybean Oil is the most important in terms of centrality. This is explained by the fact that Soybean forms a large proportion (over 1/5-th) of the agricultural output of US

farmers (Dunis et al. 2009; Ajanovic 2011), mainly because it is the most used agricultural commodity for biodiesel production in the US (Myint & El-Halwagi 2009). Cotton, Soybean Oil, and Copper are used as raw material or inputs for industries (Balli et al. 2019) and their prices are subject to demand-side shocks that are highly correlated, which can explain their prominent position in the eigenvector centrality. On the contrary, it is worth noting that Gold and Natural Gas UK lies in the lowest positions in terms of centrality. Less central nodes suggest a more stable and isolated behavior, highlighting safe-haven assets. As Balli et al. (2019) point out in their study of commodity connectedness, Gold is known for his hedging abilities in crisis time and is therefore an alternative investment vehicle. We strengthen their findings ranking the position of commodities in the network with the use of eigenvector centrality and find that indeed that this commodity occupies one of the lowest positions in the Table 1.2. Finally, to investigate the existence of clusters in the *network of commodity risk* we split the estimated graph into different clusters. We estimate three sub-graphs as presented in Figure 1.2 and Table 1.19, which highlight the cluster for each commodity. The first cluster contains 6 nodes: it includes the entire Metal sector and one agricultural commodity (Cocoa). The second cluster contains 11 nodes: the entire agriculture sector, apart from Cocoa and Orange Juice, and Ethanol. The results confirm the strong interactions among Ethanol and agricultural commodities, which are generally higher than the interactions between oil and gas and agricultural markets (Chiou-Wei et al. 2019). The third cluster includes all the Energy commodities, apart from Ethanol, and Orange Juice. The results confirm the tendency of each node to attach preferably to a node of the same sector, with few exceptions.

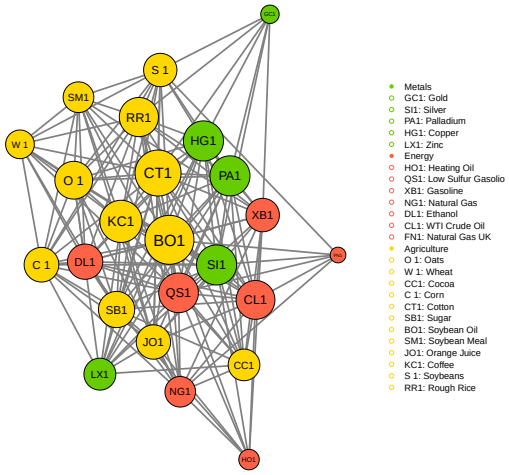

Figure 1.1. Sparse Gaussian graphical model built on the residuals following the distribution selected according to the backtesting criteria and then aggregated into a Gaussian copula. The optimal tuning parameter for the implementation of the GLASSO is 0.0158. The size of each vertex is proportional to the corresponding Eigenvector Centrality coefficient. Sample period: October 3, 2005 - March 25, 2022.

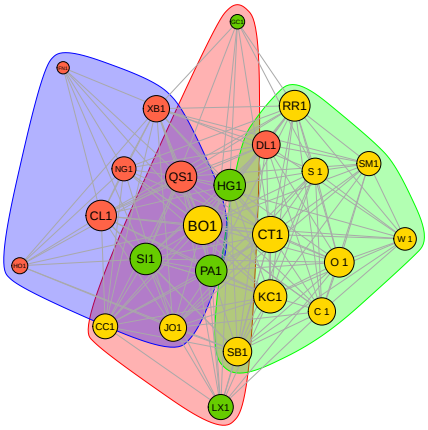

Figure 1.2. Community structure obtained via optimizing the modularity score. Size of the vertices change accordingly to the corresponding eigenvector centrality score. Sample period: October 3, 2005 - March 25, 2022.

	Eigenvector centrality
Soybean Oil	1.00
Cotton	0.94
Coffee	0.86
Copper	0.82
Silver	0.82
Palladium	0.82
Low Sulfur Gasolio	0.81
Rough Rice	0.80
WTI Crude Oil	0.79
Oats	0.77
Sugar	0.74
Ethanol	0.72
Corn	0.71
Orange Juice	0.70
Gasoline	0.68
Soybeans	0.68
Zinc	0.65
Cocoa	0.65
Soybean Meal	0.63
Natural Gas	0.62
Wheat	0.59
Heating Oil	0.42
Gold	0.38
Natural Gas UK	0.32

Table 1.2. Eigenvector centrality scores of the commodities in the sample in decreasing order. Sample period: October 3, 2005 - March 25, 2022.

1.6 Conclusions

In this paper, we investigate the connectedness within commodity markets relying on a risk management perspective. Building upon the Sparse Graphical Lasso (GLASSO) methodology, we build the network of commodity risk basing on the residuals of GARCH-type models with underlying distribution selected for each commodity through a risk management approach. The risk management approach enters the model selection criteria designed to identify the best model for the assessment of the contribution of the commodities to the market risk. Such criteria exploits traditional backtesting procedures to evaluate the Value-at-Risk predictability of the set of the models considered. We apply the methodology to the sample of twenty-four commodity futures prices over the period that spans from October 3, 2005 to March 25, 2022. Overall, we find that commodities show a moderate degree of connectedness within their network structure. The Covid-19 crisis has affected the interconnections increasing the heterogeneity within commodity markets as captured by the creation of three underlying clusters instead of the two that are estimated without including the pandemic in the sample period. However, the clusters identified mostly coincide with the three commodity sectors indicating that the additional heterogeneity is reflected in emphasized connections between commodities that belong to the same

commodity sector. From the quantitative viewpoint, this finding is also endorsed by the lower eigenvector centrality scores, weakened degree of connections between the commodities and the other sectors, and density of the graph which has decreased by 18.6%. The pandemic has also influenced the proportion of optimal MS-GARCH models in the sample which has gone from 75% to 58%. The response of *the network of the commodity risks* to the recent persistent market uncertainty proves that global increasing connectedness does not represent a crisis stylized fact. Our work is especially valuable to risk managers and policy-makers involved in the monitoring of the propagation of the commodity risk. Unlike previous studies on the assessment of the network of commodities, our methodology provides linkages that are determined conditionally on the ability of the models to forecast the VaR. In this way, the architecture of commodity markets depicts connections that incorporate information on the contribution to the market risk besides the underlying relationships that feature the commodities. This allows regulators and risk managers to infer more appropriate contribution of the commodity risks to the market risks since the risk exposure of each commodity has been conveniently modeled. Our findings reveal useful in this context providing information on the most relevant threats to the stability of commodity markets and, therefore, to design of tailored strategies for the mitigation of the commodity risk. An immediate extension of our framework would consider either the application of the Gaussian Graphical model set-up to different network specifications or the implementation of non-Gaussian copulas. Another departure from our model is the consideration of long memory among the typical stylized facts of the volatility process. A third possible insight regards the employment of a more flexible class of models for the MS-GARCH, as the semi-Markov model where the sojourn-distribution is explicitly modeled rather than implicitly assumed to be geometric. This is the object of ongoing research.

1.7 Appendix

1.7.1 A. Models Specifications

GARCH Models

Consider the time series of the log-return y_t whereby $t = 1, \ldots, T$. According to the GARCH(1,1) model, y_t can be decomposed into:

$$y_t = \mu + \epsilon_t, \tag{1.4}$$

where the series of the residuals ϵ_t presents conditional heteroskedasticity. More specifically, $\epsilon_t = \sqrt{h_t}\, z_t$ and the innovation term z_t follows the continuous standardized distribution $D_\Theta\,(0, 1,)$ with Θ as the set of parameters. The dynamics assumed for the conditional volatility h_t is:

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 h_{t-1}, \tag{1.5}$$

where α_0, α_1, and α_2 are simultaneously estimated through maximization of the log-likelihood. To ensure the positiveness of h_t and the covariance-stationarity, the parameters must satisfy respectively the conditions $\alpha_0 > 0$, $\alpha_1 \geq 0$, and $\alpha_2 \geq 0$ and $\alpha_1 + \alpha_2 < 1$.

MS-GARCH Models

Unlike GARCH models, Markov-switching GARCH (MS-GARCH) models allow the coefficients $(\alpha_0, \alpha_1, \alpha_2, \Theta)$ to change over the regimes considered. According to the MS-GARCH(1,1) specification, the dynamics log-returns and the conditional volatility are described as follows:

$$y_t = \mu + \epsilon_t, \qquad \epsilon_t = \sqrt{h_{k_t}}\, \eta_{k_t}, \qquad \eta_{k_t} | (s_t = k, \mathcal{F}_{t-1}) \sim \mathcal{D}_{\Theta_k}(0, 1), \tag{1.6}$$

$$h_{k_t} = \alpha_{k_0} + \alpha_{k_1} \epsilon_{t-1}^2 + \alpha_{k_2} h_{k_{t-1}}, \tag{1.7}$$

where the latent variable s_t, defined on the discrete space $\{1, ..., K\}$, evolves according to an unobserved first-order ergodic homogeneous Markov chain with transition probability matrix:

$$P = \begin{bmatrix} p_{1,1} & \cdots & p_{K,1} \\ \vdots & \ddots & \vdots \\ p_{1,K} & \cdots & p_{K,K} \end{bmatrix} \quad \text{with} \quad p_{i,j} = \mathbb{P}(s_t = j | s_{t-1} = i). \tag{1.8}$$

For each regime, the estimation of the parameters of the model is carried out with Maximum Likelihood. As for the GARCH model, the positiveness of h_{k_t} is

obtained requires $\alpha_{k_0} > 0$, $\alpha_{k_1} > 0$, and $\alpha_{k_2} \geq 0$, while covariance-stationarity relies on $\alpha_{k_1} + \alpha_{k_2} < 1$, $(k = 1, \ldots, K)$.

Description of Backtests

The Unconditional Coverage (UC) test of Kupiec (1995) assesses the statistical significance of the frequency of exceedances of the log-returns over the VaR. The test builds upon the following system of hypothesis:

$$\begin{cases} H_0 : \pi = \tau \\ H_1 : \pi \neq \tau \end{cases}, \tag{1.9}$$

with test statistics the likelihood ratio:

$$LR_{uc}(\tau) = -2 \times \log \left[\frac{L(\tau)}{L(\pi)} \right] = -2 \times \log \left[\frac{\tau^{n_1}(1-\tau)^{T-n_1}}{\pi^{n_1}(1-\pi)^{T-n_1}} \right], \tag{1.10}$$

where n_1 is the number of exceedances encountered. The statistic LRuc is asymptotically distributed as a chi-squared with one degree of freedom, χ^2_1.

The Conditional Coverage (CC) test of Christoffersen (1998) gauges the time dependence between VaR exceedances. A first-order Markov structure for the dependence of the hits is given by the transition probability matrix Π:

$$\Pi = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}, \tag{1.11}$$

with π_{11} (π_{10}) that indicates the probability of observing a VaR exception in t given that in $t-1$ no violation is registered:

$$\begin{aligned} \pi_{11} &= \mathbb{P}(I_t = 1 | I_{t-1} = 1) \\ \pi_{10} &= \mathbb{P}(I_t = 1 | I_{t-1} = 0) \end{aligned} \tag{1.12}$$

and $1 - \pi_{01}$ and $1 - \pi_{11}$ are the respective probability of a non-violation in t. The independence hypothesis that $\pi_{01} = \pi_{11}$ is tested with the likelihood ratio:

$$LR_{cc} = -2 \log \frac{(1-\pi)^{n_0}\pi^{n_1}}{(1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}\pi_{11}^{n_{11}}(1-\pi_{11})^{n_{10}}} \sim \chi^2_2, \tag{1.13}$$

where n_0 and n_{00} are the number of times a non-violation is encountered and the number of times it is followed by another non-violation, respectively. On the contrary, n_{11} is the number of times a violation is followed by another one, n_{10} (n_{01}) counts the times a violation is followed by a non-violation (viceversa). The Dynamic Quantile (DQ) test of Engle & Manganelli (2004) can be interpreted as an overall goodness-of-fit test for the estimated VaR. Engle & Manganelli (2004) consider that

the conditional expectation of a quantile violation given any information known at $t-1$ should be exactly τ. Hence, a linear regression model is set up where the sequence of violations represents the dependent variable while past violations or any other variables in the information set determine the set of explanatory variables. Denoting the set of parameters in the regression $\hat{\delta} = (\hat{\delta}_0, ..., \hat{\delta}_{q-1})'$, and $\mathbf{Z}$ the corresponding data matrix with columns given by the observations for the q explanatory variables, the DQ test statistic for the null hypothesis of correct unconditional and conditional coverage is:

$$DQ_{1-\tau} = \frac{\hat{\delta}'\mathbf{Z}'\mathbf{Z}\hat{\delta}}{(1-\tau)\tau}. \tag{1.14}$$

Under the null hypothesis, the test statistic is distributed as a χ^2 with q degrees of freedom where q is the number of lagged violations introduced in the aforementioned regression model.

1.7.2 B. Figures

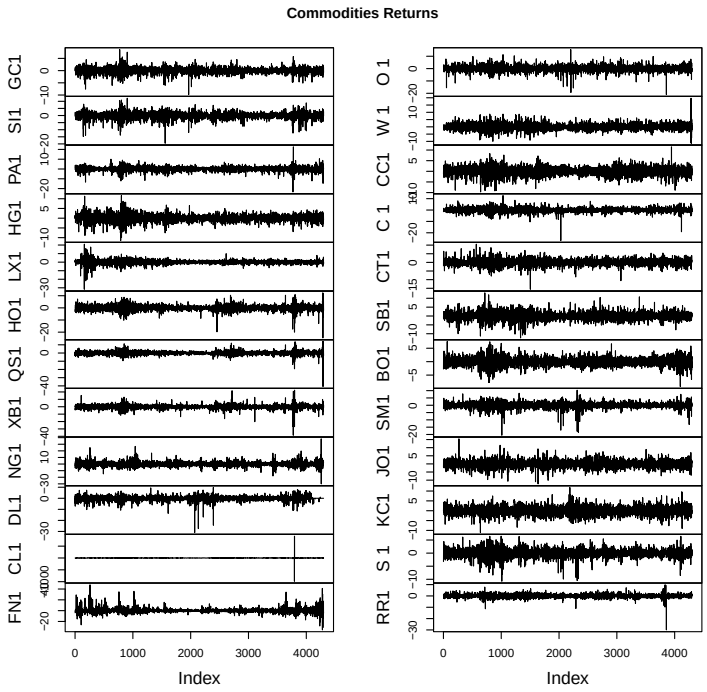

Figure 1.3. Log-returns of the commodities by sector. Agriculture: Coffee (KC1), Oats (O1), Soybeans (S1), Wheat (W1), Cocoa (CC1), Corn (C1), Rough Rice (RR1), Cotton (CT1), Sugar (SB1), Soybean Oil (BO1), Soybean Meal (SM1), Orange Juice (JO1). Energy: Gasoline (XB1), Heating Oil (HO1), Low Sulfur Gasolio (QS1), Natural Gas (NG1), Ethanol (DL1), WTI Crude Oil (CL1), Natural Gas UK (FN1). Precious Metals: Gold (GC1), Silver (SI1), Palladium (PA1). Industrial Metals: Copper (HG1), and Zinc (LX1). Sample period: October 3, 2005 - March 25, 2022.

	Mean	SD	Skewness	Kurtosis	J.B	L.B**	ARCH.LM	ADF
Gold	0.03	1.14	-0.35	5.75	0.00	0.00	0.00	0.01
Silver	0.03	2.08	-0.88	7.10	0.00	0.00	0.00	0.01
Palladium	0.06	2.12	-0.55	10.49	0.00	0.00	0.00	0.01
Copper	0.02	1.74	-0.14	4.07	0.00	0.00	0.00	0.01
Zinc	0.03	2.19	-0.94	21.27	0.00	0.00	0.00	0.01
WTI Crude Oil	0.01	20.70	-4.29	2084.02	0.00	0.00	0.00	0.01
Heating Oil	0.02	2.15	-0.80	11.35	0.00	0.00	0.00	0.01
Low Sulfur Gasolio	0.01	2.11	-1.64	38.33	0.00	0.00	0.00	0.01
Natural Gas	-0.02	3.31	0.62	9.64	0.00	0.00	0.00	0.01
Gasoline	0.01	2.65	-1.18	23.14	0.00	0.00	0.00	0.01
Natural Gas UK	0.04	3.99	1.58	21.08	0.00	0.00	0.00	0.01
Ethanol	-0.00	2.03	-2.44	29.37	0.00	0.00	0.00	0.01
Corn	0.03	1.87	-0.93	14.46	0.00	0.00	0.00	0.01
Oats	0.03	2.19	-0.59	7.93	0.00	0.00	0.00	0.01
Rough Rice	0.02	1.56	-1.78	35.34	0.00	0.00	0.00	0.01
Soybeans	0.03	1.50	-0.73	5.07	0.00	0.00	0.00	0.01
Wheat	0.03	2.07	0.24	3.88	0.00	0.00	0.00	0.01
Cocoa	0.01	1.79	-0.21	2.42	0.00	0.00	0.00	0.01
Cotton	0.02	1.75	-0.27	4.48	0.00	0.00	0.00	0.01
Coffee	0.02	1.96	0.14	1.96	0.00	0.00	0.00	0.01
Sugar	0.01	2.07	-0.04	3.15	0.00	0.00	0.00	0.01
Soybean Oil	0.03	1.47	-0.06	2.72	0.00	0.00	0.00	0.01
Soybean Meal	0.03	1.84	-1.45	13.99	0.00	0.00	0.00	0.01
Orange Juice	0.01	2.06	-0.01	2.98	0.00	0.00	0.00	0.01

Table 1.3. Summary statistics of the daily log-returns of the commodities in the sample. We report the mean (Mean), standard deviation (SD), skewness, kurtosis, test statistic of the Jarque-Bera Test (J.B), test statistic of the Ljung-Box Test on the squared log-returns with 20 lags (L.B), test statistic of the ARCH Lagrange Multiplier Test (ARCH.LM), and test statistic of the Augmented Dickey-Fuller unit root test (ADF). We denote with c the significance at the 1% level. Sample period: October 3, 2005 - March 25, 2022.

Study Reference	Study Period	Methods	Commodity class	Summary
Fong & See (2002)	1992-1997	GARCH, MS-GARCH	Crude oil	The regime switching model performs noticeably better than non-switching models regardless of evaluation criteria.
Cashin & McDermott (2002)	1862-1999	Null Hypothesis to test	Industrial commodities	Commodities have shifted from a boom phase to a slump phase if prices have declined from their previous peak
Giot & Laurent (2003)	1987-2002	RiskMetrics, Student APARCH and ARCH model	Metal, energy and agriculture commodities	Assess the performance of the RiskMetrics, skewed Student APARCH and skewed student ARCH models.
Fong & See (2003)	1992-1997	MSGARCH	Crude oil	Incorporating regime shifts improves the accuracy of short-term volatility forecasts
Baffes (2007)	1960-2005	OLS regression	Primary Commodities	Effect of crude oil prices on the prices of 35 internationally traded primary commodities
Marimoutou et al. (2009)	1983-2007	EVT models, GARCH, Historical Simulation and Filtered Historical	Crude oil	Extreme Value Theory and Filtered Historical Simulation procedures offer a major improvement over the traditional methods
Aloui & Mabrouk (2010)	1987-2007	GARCH-type models	Energy commodities	Considering for long-range memory, fat-tails and asymmetry performs better in predicting a one- day-ahead VaR for both short and long trading positions
Tyner (2010)	2006-2008	Price analysis	Energy and agriculture commodities	Exploration of the drivers in these markets as well as other major issues facing the corn ethanol industry in the United States such as the blend wall
Du et al. (2011)	1998-2009	Stochastic volatility models	Oil and agriculture commodities	Linkage between crude oil volatility and agricultural commodity markets
Nazlioglu & Soytas (2011)	1994-2010	Toda-Yamamoto; Nonparametric causality	Oil, agriculture and exchange rates	Nonlinear feedback relationship between the oil and the agricultural prices

Table 1.4. Summary of the works of reference on the literature on commodity markets. We report study reference, the study period, core method employed in the analysis (Methods), the commodity sectors on which the analysis is focused (Commodity class), and a brief description of the major findings (Summary).

Study Reference	Study Period	Methods	Commodity class	Summary
Ewing & Malik (2013)	1993-2010	GARCH	Gold and oil	Significant volatility transmission between gold and oil
Nazlioglu et al. (2013)	1986-2011	Variance causality test	Oil and agriculture commodities	A shock to oil price volatility is transmitted to agricultural markets only in the post-crisis period.
Spierdijk & Umar (2013)	1970-2011	VAR	Agriculture, energy, industrial metals, live cattle, and precious metals commodities	Significant hedging ability for commodity futures indices
Smales (2014)	2003-2012	Sentiment Analysis	Gold	Constraints imposed on traders have a significant impact on the net positions of both speculators and hedgers; this influences the way in which prices in the gold futures market react to news sentiment
Andriosopoulos & Nomikos (2014)	2007-2010	Genetic and Differential Evolution algorithms	Energy and equity	The proposed methodology suggests an effective, and at the same time, least-expensive way to operate such a fund, giving the full flexibility of any investment style, long or short, that equities can provide
Reboredo (2015)	2005-2013	Copulas	Oil and energy commodities	Oil and renewable energy displayed time-varying average and symmetric tail dependence
Rezitis (2015)	1983-2013	VAR; Granger Causality	Agricultural commodity prices, crude oil prices and US dollar exchange rates	Bidirectional panel causality effects between crude oil prices and international agricultural prices as well as between US exchange rates and international agricultural prices
Joëts (2015)	2005-2010	HAM	Energy commodities	The recent surge in energy prices is viewed as the consequence of irrational exuberance.
de Nicola et al. (2016)	1970-2013	VAR	Energy, agricultural and food commodities	Price returns of energy and agricultural commodities are highly correlated
Ji & Fan (2016)	2000-2011	Graph theory	Crude oil	The integration of the world crude oil market is verified. Furthermore, the world crude oil market is characterised as a geographical and organisational structure

Table 1.5. Summary of the works of reference on the literature on commodity markets (continued).

Study Reference	Study Period	Methods	Commodity class	Summary
Barbaglia et al. (2016)	2013-2015	Multi-class and VAR model; network analysis	Global, energy, metal and agricultural commodities	More common commodity price effects among portfolios than among markets
Kang et al. (2017)	2002-2016	DECO-GARCH	Oil and agricultural commodities and precious metals	Strong spillover during crisis; Gold and silver are transmitters to other commodities
Algieri & Leccadito (2017)	2005-2013	delta CoVaR	Energy, food and metals commodities	Commodity markets generate contagion risks which are mainly triggered by financial factors for energy and metal markets and by financial and economic fundamentals for food markets
Mensi et al. (2017)	1998-2016	VMD method and static and time-varying symmetric and asymmetric copula functions	Commodities and equity	The dependence structure varies across market conditions and under investment horizons; risk spillovers are higher in the long than the short run investment horizon.
Diebold et al. (2017)	2011-2016	VAR; FEVD; Network Analysis	Energy, livestock and agricultural commodities, precious and industrial metals	Clustering of commodities into groups; high overall connectedness and energy sector sends shocks to other commodities
Rehman et al. (2018)	1989-2016	SVAR	Crude oil, precious and industrial metals	Structural oil shocks impact precious metal returns tails except gold
Ferrer et al. (2018)	2003-2017	VAR; FEVD	Crude oil, US renewable energy stocks, high technology stocks, conventional energy stocks, US 10-year Treasury bond yields	Most of return and volatility connectedness is found in the short-term; Crude oil prices are not the key driver of renewable energy companies' performance
Laporta et al. (2018)	2002-2017	GARCH, GAS and CAViaR models; Dynamic Quantile Regression	Energy commodities	CAViaR and DQR models provide more accurate VaR estimates at high confidence levels
Ji, Bouri, Roubaud & Shahzad (2018)	2000-2017	CoVaR, delta CoVaR, dependence-switching copula	Energy and agricultural commodities	Lower tail dependence is stronger in bearish regime than in bullish one; agricultural commodities are more sensitive to shock from oil than from gas
Shen et al. (2018)	2000-2014	VAR	Energy commodities	Asymmetric patterns in response of gains and losses transmission between energy markets; Extreme market risk is more easily transmitted across markets than moderate risk.

Table 1.6. Summary of the works of reference on the literature on commodity markets (continued).

Study Reference	Study Period	Methods	Commodity class	Summary
Zhang & Broadstock (2018)	1982-2017	VAR; FEVD; Granger Causality; Network Analysis	Beverage, Fertilizers, Food, Metal, Precious metal, Raw materials, Oil	Significant rising of connectedness has been found after the global financial crisis.
Tiwari et al. (2019)	2007-2013	Wavelet analysis	Oil and energy commodities	During the shale gas revolution period of 2007-2013, oil and natural gas prices were procyclical and oil prices were leading natural gas prices.
Balli et al. (2019)	2007-2016	GDFM; SV models	Energy commodities, precious and industrial metals, and agricultural commodities	Spillovers increase during the GFC and 2014-16 oil price collapse; Network analysis shows more spillover within a specific commodity class.
Ramiah et al. (2019)	1990-2017	Non-parametric ranking test and kernel regression	Metal, chemical, precious, energy and agriculture commodities	There is a delayed reaction of investor in commodity markets compared to the equity market
Malik & Umar (2019)	1996-2019	VAR	Exchange rates of major oil-exporting and oil-importing countries	Results show that demand shocks have a major impact while supply shocks have no impact.
Zaremba et al. (2019)	1265-2017	Wavelet analysis	Agriculture, energy, industrial commodities	Robust inflation hedging properties of agricultural, energy, and industrial commodities for the 4- to 8-year horizon through almost the entire seven centuries
Nguyen et al. (2020)	1992-2017	local Gaussian correlation measure	Commodity and U.S. financial markets	Financialization hypothesis is confirmed; Special role of Gold is highlighted.
Christodoulakis (2020)	2001-2013	GMM	Energy, agriculture, livestock, industrial metals and precious metals	Joint preference asymmetries for longer maturities, joint preference symmetries for short maturities.
Naeem et al. (2020)	2007-2018	EGARCH-Copula approach	Energy and commodity ETFs, oil	Positive correlations of energy and commodity ETFs with oil prices is found
Flori et al. (2021)	1980-2017	Graph-theoretical approach; Granger causal connectivity analysis	Commodities and climate related variables	Climate conditions affect financial stability by impacting commodity comovements

Table 1.7. Summary of the works of reference on the literature on commodity markets (continued).

Study Reference	Study Period	Methods	Commodity class	Summary
Umar, Gubareva & Teplova (2021)	2020-2020	Wavelet coherence and phase-difference methods	Energy commodities, precious and agricultural commodities	Non-precious metals offer the best diversification during the recovery from crisis
Zaremba, Umar & Mikutowski (2021)	1850-2019	Pearson's product-moment pairwise correlation coefficients; Gerber statistic; R^2 analysis.	Precious metals, energy, industrials, and agriculturals	Findings cast doubt on the link between commodity dependence and financialisation
Balcilar et al. (2021)	1990-2019	TVP-VAR	Crude Oil and agricultural commodities	Crude Oil not only affects commodity assets but is also equally responsive to their innovations
Tiwari et al. (2021)	1986-2018	time-varying generalised Hurst exponent	Energy commodities	After the subprime crisis, the persistence of energy spot market products has increased
Umar, Jareño & Escribano (2021)	2000-2020	TVP-VAR	Oil and agriculture commodities	Higher directional return and volatility connectedness to oil risk and demand than supply shocks
Umar et al. (2022)	2020-2021	TVP-VAR	Agricultural and livestock commodity, Coronavirus Media Coverage Index	Dynamic total return and volatility connectedness fluctuate over time, reaching a peak during both the first and the third waves of the global pandemic crisis

Table 1.8. Summary of the works of reference on the literature on commodity markets (continued).

Confidence level	95%				99%			
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
Gold								
GARCHnorm	0.376169**	0.673278**	0.272021**	0.926612	1.8e-05	7.00E-06	0	1.927354
MSGARCHnorm	0.538857**	0.768622**	0.318919**	0.948851	0.848702**	0.078492**	0.025813	0.963677
GARCHstd	0.377801**	0.646796**	0.514667**	1.07487	0.996913**	0.09181**	0.075842**	1.000741
MSGARCHstd	0.992953**	0.952407**	0.353444**	1.000741	0.191705**	0.013278	0.000506	1.260193
GARCHged	0.481144**	0.780227**	0.547514**	0.941438	0.844493**	0.010177	4.4e-05	1.037806
MSGARCHged	0.866417**	0.959651**	0.775744**	0.985915	0.447366**	0.001146	0	1.148999
GARCHsnorm	0.247085**	0.500426**	0.212008**	0.904374	0.000457	7.3e-05	0	1.742031
MSGARCHsnorm	0.046827	0.116861**	0.051335**	0.837658	0.848702**	0.078492**	0.016064	0.963677
GARCHsstd	0.784965**	0.87801**	0.525519**	1.02298	0.69809**	0.064088**	0.035234	0.926612
MSGARCHsstd	0.729269**	0.792302**	0.250049**	0.97109	0.844493**	0.010177	4.7e-05	1.037806
GARCHsged	0.329276**	0.616101**	0.407365**	0.919199	0.69809**	0.064088**	0.046315	0.926612
MSGARCHsged	0.599692**	0.869062**	0.142255**	0.956264	0.844493**	0.010177	3.3e-05	1.037806
Silver								
GARCHnorm	0.426842**	0.728406**	0.01555	0.934025	0	0	0	2,149,741
MSGARCHnorm	0.93306**	0.644346**	0.090229**	0.99296	0.066456**	0.049959	0	1.371386
GARCHstd	0.163877**	0.370471**	0.082125**	1.118933	0.191705**	0.013278	1e-04**	1.260193
MSGARCHstd	0.536826**	0.815249**	0.096226**	1.052242	0.004327	0.00181	0	1.593773
GARCHged	0.481144**	0.712248**	0.067553**	0.941438	0.191705**	0.013278	8.6e-05**	1.260193
MSGARCHged	0.797141**	0.944272**	0.042973	0.978503	0.011746	0.003498	0	1.519644
GARCHsnorm	0.180153**	0.390089**	0.017609	0.889548	4e-06**	0	0	2.001483
MSGARCHsnorm	0.284285**	0.46621**	0.037643	0.911449	0.699371**	0.11023**	0	1.074870
GARCHsstd	0.428915**	0.642174**	0.088059**	1.067062	0.699371**	0.0118	1.2e-05**	1.074870
MSGARCHsstd	0.793748**	0.574531**	0.038624	0.97814	0.565949**	0.113531**	9e-06**	1.111935
GARCHsged	0.376169**	0.593415**	0.02916	0.926612	0.447366**	0.013942	2.9e-05**	1.148999
MSGARCHsged	0.535971**	0.759965**	0.045742	0.948499	0.844493**	0.102753**	0.000173	1.037806
Copper								
GARCHnorm	0.247085**	0.417994**	0.479086**	0.904374	0.000826	0.000109	0	1.704967
MSGARCHnorm	0.599692**	0.869062**	0.837004**	0.956264	0.018733	0.000678	0	1.482580
GARCHstd	0.866417**	0.836154**	0.547015**	0.985915	0.260377**	0.001327	0	1.223128
MSGARCHstd	0.481144**	0.780227**	0.855243**	0.941438	0.096832**	0.001232	0	1.334322
GARCHged	0.286243**	0.47466**	0.539922**	0.911787	0.447366**	0.001146	0	1.148999
MSGARCHged	0.329276**	0.616101**	0.058863**	0.919199	0.191705**	0.001345	0	1.260193
GARCHsnorm	0.211755**	0.364273**	0.417435**	0.896961	0.004327	0.00031	0	1.593773
MSGARCHsnorm	0.286243**	0.558099**	0.68568**	0.911787	0.066456**	0.001117	0	1.371386
GARCHsstd	0.6633**	0.711657**	0.523842**	0.963677	0.345368**	0.001259	0	1.186064
MSGARCHsstd	0.329276**	0.533458**	0.617516**	0.919199	0.260377**	0.001327	0	1.223128
GARCHsged	0.286243**	0.47466**	0.526265**	0.911787	0.565949**	0.000999	0	1.111935
MSGARCHsged	0.180153**	0.390089**	0.406984**	0.889548	0.260377**	0.001327	0	1.223128

Table 1.9. Backtesting results for the commodities Gold, Silver, Copper (Metal commodity sector). We report the p-values of the Unconditional Coverage (uc.LRp), Conditional Coverage (cc.LRp), Dynamic Quantile (DQp) tests and the Actual over Expected Exceedances Ratio (AE) computed at the 95% and 99% confidence levels. The models included in the assessment are GARCH(1,1) with innovations following a Normal (GARCHnorm), Skewed Normal (GARCHsnorm), Student's-t (GARCHstd), Skewed Student's-t (GARCHsstd), Generalized Error Distribution (GARCHged), and Skewed Generalized Error Distribution (GARCHsged) and MS-GARCH(1,1) following a Normal (MSGARCHnorm), Skewed Normal (MSGARCHsnorm), Student's-t (MSGARCHstd), Skewed Student's-t (MSGARCHsstd), Generalized Error Distribution (MSGARCHged), and Skewed Generalized Error Distribution (MSGARCHsged).
* denotes the p-values higher than the 5% significance level
** denotes the p-values higher than the 1% significance level.
Sample period: October 3, 2005 - March 25, 2022.

Confidence level		95%				99%		
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
Palladium								
GARCHnorm	0.922692**	0.038134	0.055372**	1.008154	0	0	0	2.149741
MSGARCHnorm	0.220265**	0.001177	0.000205	1.104522	0.004327	0.00181	1.8e-05**	1.593773
GARCHstd	0.220265**	0.036004	0.012632	1.104522	0.011746	0.003498	0.000101	1.519644
MSGARCHstd	0.53389**	0.012154	0.007221	1.052632	0.011746	0.003498	0.000163	1.519644
GARCHged	0.922692**	0.038134	0.053203**	1.008154	0.007206	0.002552	7.5e-05**	1.556709
MSGARCHged	0.718505**	0.022332	0.009323	1.030393	0.000457	0.000348	0	1.742031
GARCHsnorm	0.329276**	0.004679	0.013865	0.919199	9.00E-06	4.00E-06	0	1.964418
MSGARCHsnorm	0.599692**	0.033631	0.016613	0.956264	0.004327	0.00181	4.3e-05	1.593773
GARCHsstd	0.797141**	0.022501	0.028856	0.978503	0.137853**	0.012129	0.000299	1.297257
MSGARCHsstd	0.6633**	0.087649**	0.045702	0.963677	0.260377**	0.014018	0.000167	1.223128
GARCHsged	0.329276**	0.004679	0.01127	0.919199	0.029222	0.00601	0.00022	1.445515
MSGARCHsged	0.426842**	0.046269	0.062604**	0.934025	0.096832**	0.010698	0.000293	1.334322
Zinc								
GARCHnorm	0.936567**	0.440549**	0.128704**	0.993328	0.260377**	0.384782**	0.35774**	1.223128
MSGARCHnorm	0.922692**	0.489923**	0.001407	1.008154	0.844493**	0.569028**	0.1717**	1.037806
GARCHstd	0.85319**	0.509112**	0.010503	1.015567	0.556782**	0.384266**	0.397618**	0.889548
MSGARCHstd	0.85319**	0.723387**	0.02322	1.015567	0.429471**	0.311619**	0.290963**	0.852483
GARCHged	0.797141**	0.380745**	0.020441	0.978503	0.429471**	0.311619**	0.343904**	0.852483
MSGARCHged	0.797141**	0.588467**	0.033023	0.978503	0.556782**	0.384266**	0.175018**	0.889548
GARCHsnorm	0.797141**	0.380745**	0.083603**	0.978503	0.447366**	0.501277**	0.5536**	1.148999
MSGARCHsnorm	0.718505**	0.889211**	0.037931	1.030393	0.699371**	0.566416**	0.240576**	1.074870
GARCHsstd	0.936567**	0.440549**	0.020503	0.993328	0.429471**	0.60052**	0.670183**	0.852483
MSGARCHsstd	0.784965**	0.523944**	0.068473**	1.022980	0.319515**	0.240722**	0.58721**	0.815419
GARCHsged	0.729269**	0.348669**	0.039292	0.97109	0.228641**	0.410883**	0.419957**	0.778354
MSGARCHsged	0.797141**	0.800231**	0.05046**	0.978503	0.556782**	0.384266**	0.294271**	0.889548

Table 1.10. Backtesting results for the commodities Palladium, Zinc (Metal commodity sector).

Confidence level	95%				99%			
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
WTI								
GARCHnorm	0.286243**	0.215639**	0.324009**	0.911787	0.000457	0.002108	0.998099**	1.742031
MSGARCHnorm	0	0	0.999998**	9.870322	0	0	1**	49.351612
GARCHstd	0.162516**	0.324307**	0.993754**	1.119348	0.018733	0.05598**	0.999999**	1.48258
MSGARCHstd	0.138392**	0.063562**	0.99884**	1.126761	0.011746	0.037856	1**	1.519644
GARCHged	0.6633**	0.505691**	0.998082**	0.963677	0.096832**	0.201986**	1**	1.334322
MSGARCHged	0.426311**	0.50965**	0.987309**	1.067457	0.007206	0.024923	1**	1.556709
GARCHsnorm	0.07161**	0.079492**	0.769626**	0.852483	0.007206	0.024923	0.999884**	1.556709
MSGARCHsnorm	0	0	0.999998**	9.870322	0	0	1**	49.351612
GARCHsstd	0.922692**	0.707753**	0.989617**	1.008154	0.096832**	0.201986**	0.999997**	1.334322
MSGARCHsstd	0.85319**	0.179494**	0.994373**	1.015567	0.096832**	0.201986**	0.999999**	1.334322
GARCHsged	0.426842**	0.499202**	0.981284**	0.934025	0.345368**	0.446807**	0.999988**	1.186064
MSGARCHsged	0	0	0.988484**	1.452928	0	0	0.999999**	3.854707
Heating Oil								
GARCHnorm	0.53389**	0.693124**	0.120038**	1.052632	0.001464	0.006095	0.00027	1.667902
MSGARCHnorm	0.0467	0.137776**	0.007206	1.171238	0.029222	0.080541**	0.008966	1.445515
GARCHstd	0.068662**	0.178165**	0.020561	1.156412	0.191705**	0.3205**	0.010064	1.260193
MSGARCHstd	0.098622**	0.24531**	0.038368	1.141586	0.191705**	0.3205**	0.007833	1.260193
GARCHged	0.592603**	0.745351**	0.099405**	1.045219	0.447366**	0.501277**	0.029084	1.148999
MSGARCHged	0.098622**	0.233768**	0.022659	1.141586	0.007206	0.024923	0.001097	1.556709
GARCHsnorm	0.784965**	0.963175**	0.215632**	1.02298	0.007206	0.024923	0.00079	1.556709
MSGARCHsnorm	0.098622**	0.24531**	0.029798	1.141586	0.096832**	0.201986**	0.008677	1.334322
GARCHsstd	0.162516**	0.371858**	0.031443	1.119348	0.699371**	0.566416**	0.033559	1.07487
MSGARCHsstd	0.068662**	0.190635**	0.023278	1.156412	0.137853**	0.25846**	0.004244	1.297257
GARCHsged	0.654254**	0.794345**	0.091228**	1.037806	0.565949**	0.542742**	0.033717	1.111935
MSGARCHsged	0.332944**	0.624808**	0.056914**	1.082283	0.096832**	0.201986**	0.038381	1.334322
Low Sulfur Gasoline								
GARCHnorm	0.016083	0.008514	0.013422	1.208302	0	3.00e-06	0	2.112676
MSGARCHnorm	0.003693	0.003696	0.012349	1.25278	1.8e-05	6.6e-05	9.00e-06	1.927354
GARCHstd	0.000201	0.000119	9.8e-05	1.326909	0.002544	0.009995	0.021922	1.630838
MSGARCHstd	0.000924	0.000303	0.000321	1.289844	0.011746	0.01534	0.007292	1.519644
GARCHged	0.020139	0.009492	0.014018	1.20089	0.137853**	0.25846**	0.611293**	1.297257
MSGARCHged	0.02507	0.005551	0.004621	1.193477	0.011746	0.037856	0.248321**	1.519644
GARCHsnorm	0.02507	0.010489	0.014526	1.193477	2.00e-06	1.2e-05	1.00e-06	2.038547
MSGARCHsnorm	0.08253**	0.030287	0.023212	1.148999	0.000131	0.000391	3.9e-05	1.81616
GARCHsstd	0.001637	0.001343	0.001067	1.275019	0.044571	0.112663**	0.084314**	1.408451
MSGARCHsstd	0.020139	0.002545	0.002498	1.20089	0.066456**	0.049959	0.011816	1.371386
GARCHsged	0.038177	0.012473	0.012606	1.178651	0.066456**	0.153107**	0.097539**	1.371386
MSGARCHsged	0.138392**	0.107628**	0.026985	1.126761	0.260377**	0.384782**	0.689165**	1.223128

Table 1.11. Backtesting results for the commodities WTI, Heating Oil, Low Sulfur Gasoline (Energy commodity sector).

Confidence level		95%				99%		
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
Natural Gas UK								
GARCHnorm	0.481144**	0.009157	0.000199	0.941438	0.018733	0.021533	0	1.48258
MSGARCHnorm	0.038177	0.006409	2.00e-06	1.178651	0.565949**	0.113531**	0	1.111935
GARCHstd	0.031028	0.001312	1.8e-05	1.186064	0.447366**	0.501277**	0	1.148999
MSGARCHstd	0.162516**	0.007884	5.00e-06	1.119348	0.066456**	0.153107**	0	1.371386
GARCHged	0.538857**	0.011202	0.000345	0.948851	0.996913**	0.549507**	0	1.000741
MSGARCHged	0.162516**	0.007884	7.8e-05	1.119348	0.096832**	0.062199**	0	1.334322
GARCHsnorm	0.332944**	0.006588	0.000186	1.082283	0.000131	3.00e-05	0	1.816160
MSGARCHsnorm	0.003693	0.000623	0	1.252780	0.018733	0.004655	0	1.482580
GARCHsstd	0.000689	3.1e-05	0	1.297257	0.137853**	0.25846**	0	1.297257
MSGARCHsstd	0.01008	0.000949	0	1.223128	0.044571	0.038892	0	1.408451
GARCHsged	0.592603**	0.026594	0.00089	1.045219	0.191705**	0.3205**	0	1.260193
MSGARCHsged	0	0	0	10.455724	0	0	0	52.278622
Ethanol								
GARCHnorm	0.001242	8.7e-05	0.003903	0.741015	0.018838	0.004379	7.4e-05	1.482030
MSGARCHnorm	0	0	0	1.170804	0	0	0	5.854020
GARCHstd	4.00e-06	0	0	1.407929	0	0	0	3.853279
MSGARCHstd	0	0	0	1.682104	0	0	0	3.001112
GARCHged	0.105096**	0	0	0.866988	0.007252	0	0	1.556132
MSGARCHged	0	0	0	1.170804	0	0	0	5.854020
GARCHsnorm	5.9e-05	4.4e-05	0.002908	0.681734	0.138392**	0.254925**	0.585762**	1.296777
MSGARCHsnorm	0.139598**	0	0.051353**	1.126343	4.00e-06	0	0.75966**	2.000741
GARCHsstd	7.7e-05	0	0	1.348648	0	0	0	3.705076
MSGARCHsstd	1.00e-06	0	0	1.444980	0	0	0.054491**	2.260096
GARCHsged	0.327124**	0	0	0.918859	0.000133	0	0	1.815487
MSGARCHsged	0.481164**	0	0.013414	1.059652	0	0	0.113669**	2.111893
Gasoline								
GARCHnorm	0.797141**	0.949723**	0.999712**	0.978503	1.00e-06	5.00e-06	1.00e-06	2.075612
MSGARCHnorm	0.098622**	0.24531**	0.429401**	1.141586	0.096832**	0.201986**	0.067329**	1.334322
GARCHstd	0.332944**	0.589478**	0.655888**	1.082283	0.044571	0.112663**	0.040811	1.408451
MSGARCHstd	0.377801**	0.646796**	0.465694**	1.074870	0.011746	0.037856	0.035793	1.519644
GARCHged	0.797141**	0.944272**	0.835113**	0.978503	0.565949**	0.605186**	0.228264**	1.111935
MSGARCHged	0.291763**	0.532529**	0.730573**	1.089696	0.044571	0.112663**	0.052285**	1.408451
GARCHsnorm	0.426842**	0.67518**	0.977582**	0.934025	1.8e-05	6.6e-05	8.8e-05	1.927354
MSGARCHsnorm	0.25423**	0.495833**	0.614729**	1.097109	0.066456**	0.153107**	0.076079**	1.371386
GARCHsstd	0.53389**	0.809986**	0.677071**	1.052632	0.565949**	0.605186**	0.260247**	1.111935
MSGARCHsstd	0.377801**	0.646796**	0.121419**	1.074870	0.191705**	0.3205**	0.064282**	1.260193
GARCHsged	0.866417**	0.959651**	0.712483**	0.985915	0.260377**	0.352689**	0.18532**	1.223128
MSGARCHsged	0.729269**	0.758381**	0.914729**	0.97109	0.137853**	0.25846**	0.06866**	1.297257
Natural Gas								
GARCHnorm	0.106023**	0.248188**	0.000613	0.867309	0.260377**	0.352689**	0	1.223128
MSGARCHnorm	0.481144**	0.373426**	0.000525	0.941438	0.137853**	0.209911**	0	1.297257
GARCHstd	0.286243**	0.47466**	0.000333	0.911787	0.556782**	0.678314**	0	0.889548
MSGARCHstd	0.6633**	0.711657**	0.00055	0.963677	0.345368**	0.436289**	0	1.186064
GARCHged	0.046827	0.116861**	0.000287	0.837658	0.228641**	0.410883**	0	0.778354
MSGARCHged	0.211755**	0.254273**	0.000536	0.896961	0.844493**	0.73126**	0	1.037806
GARCHsnorm	0.599692**	0.821237**	0.001347	0.956264	0.011746	0.022208	0	1.519644
MSGARCHsnorm	0.797141**	0.800231**	0.000541	0.978503	0.011746	0.022208	0	1.519644
GARCHsstd	0.936567**	0.86523**	0.001241	0.993328	0.996913**	0.761058**	0	1.000741
MSGARCHsstd	0.592603**	0.841768**	0.00054	1.045219	0.018733	0.03455	0	1.482580
GARCHsged	0.211755**	0.444122**	0.00056	0.896961	0.996913**	0.761058**	0	1.000741
MSGARCHsged	0.599692**	0.661202**	0.001528	0.956264	0.066456**	0.110924**	0	1.371386

Table 1.12. Backtesting results for the commodities Natural Gas UK, Ethanol, Gasoline, Natural Gas (Energy commodity sector).

Confidence level		95%				99%		
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
Corn								
GARCHnorm	1e-06**	0	9.3e-05**	0.615271	0.011746	6.1e-05**	0	1.519644
MSGARCHnorm	0.087482**	0.011279	0.036897	0.859896	0.191705**	0.013278	8.7e-05**	1.260193
GARCHstd	0.023255	0.000595	0.010703	0.815419	0.447366**	0.112431**	0.089284**	1.148999
MSGARCHstd	0.087482**	0.02633	0.094703**	0.859896	0.137853**	0.001312	0	1.297257
GARCHged	0	0	1.8e-05**	0.570793	0.102975**	0.231256**	0.29938**	0.704225
MSGARCHged	0.002433	0.000638	0.019307	0.756116	0.447366**	0.013942	0.000126	1.148999
GARCHsnorm	1e-06**	0	0.000183	0.607858	0.066456**	0.001117	1e-06**	1.371386
MSGARCHsnorm	0.037398	0.003423	0.032947	0.830245	0.565949**	0.013098	0.000222	1.111935
GARCHsstd	0.01068	0.000187	0.004173	0.79318	0.699371**	0.11023**	0.059708**	1.07487
MSGARCHsstd	0.013966	0.002355	0.014881	0.800593	0.345368**	0.001259	0	1.186064
GARCHsged	2e-06**	0	6e-05**	0.630096	0.996913**	0.09181**	0.034724	1.000741
MSGARCHsged	0.003333	0.00035	0.001891	0.763529	0.066456**	0.001117	0	1.371386
Oats								
GARCHnorm	0.010551	0.025755	0.028003	0.792886	0.191705**	0.3205**	0.45965**	1.260193
MSGARCHnorm	0.793748**	0.201993**	0.105016**	0.97814	0.319515**	0.240722**	0.702016**	0.815419
GARCHstd	0.856665**	0.483646**	0.083857**	1.015191	0.848702**	0.509321**	0.836011**	0.963677
MSGARCHstd	0.660119**	0.161073**	0.256858**	0.96332	0.844493**	0.569028**	0.633835**	1.037806
GARCHged	0.046355	0.078011**	0.031655	0.837347	0.228641**	0.176721**	0.594603**	0.778354
MSGARCHged	0.478429**	0.106085**	0.050126**	0.941089	0.996913**	0.549507**	0.877766**	1.000741
GARCHsnorm	0.007993	0.019337	0.026297	0.785476	0.191705**	0.3205**	0.459643**	1.260193
MSGARCHsnorm	0.793748**	0.201993**	0.097417**	0.97814	0.102975**	0.231256**	0.790905**	0.704225
GARCHsstd	0.856665**	0.483646**	0.110756**	1.015191	0.996913**	0.549507**	0.844961**	1.000741
MSGARCHsstd	0.862953**	0.115043**	0.012937	0.98555	0.565949**	0.113531**	9.8e-05**	1.111935
GARCHsged	0.070937**	0.121375**	0.046818	0.852167	0.429471**	0.311619**	0.763183**	0.852483
MSGARCHsged	0.327124**	0.129665**	0.037439	0.918859	0.848702**	0.509321**	0.662643**	0.963677
Rough Rice								
GARCHnorm	0.037398	0.000112	0.001064	0.830245	0.260377**	0.098552**	0.127786**	1.223128
MSGARCHnorm	0.106023**	0.000653	0.004699	0.867309	0.556782**	0.384266**	0.700468**	0.889548
GARCHstd	0.729269**	0.000836	0.000241	0.97109	0.429471**	0.036898	0.015663	0.852483
MSGARCHstd	0.426842**	6.7e-05	2.3e-05	0.934025	0.69809**	0.452346**	0.83786**	0.926612
GARCHged	0.013966	1.00e-06	7.00e-06	0.800593	0.429471**	0.036898	0.010444	0.852483
MSGARCHged	0	0	0	10,778,354	0	0	0	5.3891772
GARCHsnorm	0.058145**	0.000235	0.002011	0.84507	0.191705**	0.087449**	0.134392**	1.260193
MSGARCHsnorm	0.936567**	0.001529	0.000183	0.993328	0.447366**	0.501277**	0.666702**	1.148999
GARCHsstd	0.592603**	0.00154	0.001052	1.045219	0.996913**	0.008388	3.5e-05	1,000,741
MSGARCHsstd	0.992953**	0.000591	3.7e-05	1.000741	0.137853**	0.012129	0.000517	1.297257
GARCHsged	0.247085**	0.000271	0.001173	0.904374	0.69809**	0.064088**	0.032217	0.926612
MSGARCHsged	0.481144**	0.001061	0.001119	0.941438	0.044571	0.000978	0	1.408451

Table 1.13. Backtesting results for the commodities Corn, Oats, Rough Rice (Agriculture commodity sector).

Confidence level		95%				99%		
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
Cocoa								
GARCHnorm	0.481144**	0.373426**	0.404176**	0.941438	0.007206	0.000415	0	1.556709
MSGARCHnorm	0.729269**	0.197239**	0.133631**	0.97109	0.029222	0.02937	0.000264	1.445515
GARCHstd	0.922692**	0.489923**	0.531656**	1.008154	0.565949**	0.113531**	0.027145	1.111935
MSGARCHstd	0.936567**	0.262422**	0.16268**	0.993328	0.447366**	0.112431**	0.006144	1.148999
GARCHged	0.180153**	0.127769**	0.29398**	0.889548	0.848702**	0.509321**	0.455834**	0.963677
MSGARCHged	0.329276**	0.03151	0.009584	0.919199	0.699371**	0.566416**	0.040742	1.07487
GARCHsnorm	0.329276**	0.251301**	0.332859**	0.919199	0.018733	0.000678	0	1.48258
MSGARCHsnorm	0.538857**	0.251877**	0.083059**	0.948851	0.096832**	0.201986**	0.051822**	1.334322
GARCHsstd	0.729269**	0.348669**	0.355458**	0.97109	0.844493**	0.102753**	0.021506	1.037806
MSGARCHsstd	0.866417**	0.24103**	0.109761**	0.985915	0.699371**	0.566416**	0.423868**	1.07487
GARCHsged	0.247085**	0.183078**	0.396397**	0.904374	0.848702**	0.509321**	0.601433**	0.963677
MSGARCHsged	0.426842**	0.099668**	0.072971**	0.934025	0.69809**	0.452346**	0.554777**	0.926612
Cotton								
GARCHnorm	0.426842**	0.099668**	0.111656**	0.934025	0.000247	0.000216	1.4e-05	1.779096
MSGARCHnorm	0.784965**	0.098907**	0.170881**	1.02298	0.191705**	0.3205**	0.475666**	1.260193
GARCHstd	0.654254**	0.112376**	0.199126**	1.037806	0.345368**	0.446807**	0.610373**	1.186064
MSGARCHstd	0.936567**	0.067909**	0.120201**	0.993328	0.345368**	0.446807**	0.67507**	1.186064
GARCHged	0.329276**	0.141126**	0.200114**	0.919199	0.996913**	0.09181**	0.062845**	1.000741
MSGARCHged	0.211755**	0.038545	0.080121**	0.896961	0.345368**	0.107205**	0.007552	1.186064
GARCHsnorm	0.180153**	0.030838	0.050062**	0.889548	0.000826	0.000547	7.2e-05	1,704,967
MSGARCHsnorm	0.866417**	0.126714**	0.055129**	0.985915	0.565949**	0.542742**	0.535499**	1.111935
GARCHsstd	0.718505**	0.105951**	0.137861**	1.030393	0.137853**	0.074986**	0.082556**	1.297257
MSGARCHsstd	0.797141**	0.05291**	0.049236	0.978503	0.191705**	0.087449**	0.06821**	1.260193
GARCHsged	0.329276**	0.141126**	0.187072**	0.919199	0.996913**	0.09181**	0.06281**	1.000741
MSGARCHsged	0.599692**	0.033631	0.058223**	0.956264	0.345368**	0.107205**	0.062089**	1.186064
Coffee								
GARCHnorm	0.599692**	0.292951**	0.182489**	0.956264	0.096832**	0.201986**	0.517523**	1.334322
MSGARCHnorm	0.599692**	0.292951**	0.082859**	0.956264	0.69809**	0.734075**	0.989647**	0.926612
GARCHstd	0.332944**	0.318135**	0.327701**	1.082283	0.102975**	0.231256**	0.864046**	0.704225
MSGARCHstd	0.592603**	0.140428**	0.077869**	1.045219	0.69809**	0.734075**	0.950101**	0.926612
GARCHged	0.538857**	0.293491**	0.336994**	0.948851	0.038264	0.104931**	0.666176**	0.630096
MSGARCHged	0.797141**	0.535237**	0.134944**	0.978503	0.69809**	0.452346**	0.734606**	0.926612
GARCHsnorm	0.797141**	0.099346**	0.215052**	0.978503	0.066456**	0.153107**	0.399228**	1.371386
MSGARCHsnorm	0.866417**	0.264959**	0.091541**	0.985915	0.996913**	0.761058**	0.915913**	1.000741
GARCHsstd	0.189745**	0.178477**	0.209194**	1.111935	0.102975**	0.231256**	0.875625**	0.704225
MSGARCHsstd	0.426311**	0.224706**	0.06562**	1.067457	0.848702**	0.762395**	0.176027**	0.963677
GARCHsged	0.538857**	0.113066**	0.218791**	0.948851	0.064408**	0.160255**	0.758104**	0.667161
MSGARCHsged	0.85319**	0.4361**	0.172491**	1015567	0.319515**	0.508443**	0.853383**	0.815419

Table 1.14. Backtesting results for the commodities Cocoa, Cotton, Coffee (Agriculture commodity sector).

Confidence level	95%				99%			
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
Soybeans								
GARCHnorm	0.070937**	0.074143**	0.339663**	0.852167	0.011746	0.022208	0.068071**	2
MSGARCHnorm	0.373825**	0.425806**	0.84635**	0.926269	0.996913**	0.761058**	0.644437**	1.000741
GARCHstd	0.793748**	0.777613**	0.962079**	0.97814	0.848702**	0.762395**	0.58281**	0.963677
MSGARCHstd	0.478429**	0.205038**	0.569573**	0.941089	0.556782**	0.678314**	0.636652**	0.889548
GARCHged	0.150896**	0.169852**	0.637644**	0.881808	0.429471**	0.60052**	0.876634**	0.852483
MSGARCHged	0.245318**	0.282074**	0.577385**	0.904039	0.848702**	0.762395**	0.859678**	0.963677
GARCHsnorm	0.010551	0.025755	0.251902**	0.792886	0.137853**	0.209911**	0.318147**	1.297257
MSGARCHsnorm	0.210175**	0.43667**	0.724079**	0.896628	0.228641**	0.410883**	0.898333**	0.778354
GARCHsstd	0.210175**	0.351503**	0.770771**	0.896628	0.429471**	0.60052**	0.463687**	0.852483
MSGARCHsstd	0.037009	0.061497**	0.309101**	0.829937	0.102975**	0.231256**	0.807443**	0.704225
GARCHsged	0.029299	0.072187**	0.42747**	0.822527	0.228641**	0.410883**	0.766017**	0.778354
MSGARCHsged	0.046355	0.045807	0.219669**	0.837347	0.319515**	0.508443**	0.808987**	0.815419
Wheat								
GARCHnorm	0.018101	0.046651	0.272935**	0.808006	0.228641**	0.410883**	0.29739**	0.778354
MSGARCHnorm	0.127491**	0.17592**	0.731896**	0.874722	0.064408**	0.160255**	0.781315**	0.667161
GARCHstd	0.211755**	0.366135**	0.43829**	0.896961	0.102975**	0.231256**	6.3e-05	0.704225
MSGARCHstd	0.152128**	0.300643**	0.533684**	0.882135	0.005665	0.020242	0.069853**	0.518903
GARCHged	0.023255	0.056412**	0.205177**	0.815419	0.038264	0.104931**	0.05844**	0.630096
MSGARCHged	0.152128**	0.193574**	0.630419**	0.882135	0.011397	0.037439	0.117407**	0.555967
GARCHsnorm	0.058145**	0.153375**	0.474211**	0.84507	0.844493**	0.73126**	0.633005**	1.037806
MSGARCHsnorm	0.329276**	0.456365**	0.880336**	0.919199	0.102975**	0.231256**	0.414457**	0.704225
GARCHsstd	0.538857**	0.740983**	0.577715**	0.948851	0.228641**	0.410883**	0.000269	0.778354
MSGARCHsstd	0.797141**	0.535237**	0.677215**	0.978503	0.102975**	0.231256**	0.402083**	0.704225
GARCHsged	0.106023**	0.236974**	0.457106**	0.867309	0.102975**	0.231256**	0.133795**	0.704225
MSGARCHsged	0.286243**	0.428291**	0.69054**	0.911787	0.038264	0.104931**	0.24896**	0.630096
Orange Juice								
GARCHnorm	0.784965**	0.046657	0.006078	1.02298	0.000131	0.000664	2.00e-06	1.81616
MSGARCHnorm	0.53389**	0.221842**	0.003341	1.052632	0.848702**	0.762395**	0.197053**	0.963677
GARCHstd	0.08253**	0.030287	0.001254	1.148999	0.429471**	0.60052**	0.29819**	0.852483
MSGARCHstd	0.25423**	0.208511**	0.001421	1.097109	0.699371**	0.677201**	0.129704**	1.07487
GARCHged	0.654254**	0.024493	0.003877	1.037806	0.102975**	0.231256**	0.238843**	0.704225
MSGARCHged	0.922692**	0.038134	0.008726	1008154	0.996913**	0.761058**	0.010889	1.000741
GARCHsnorm	0.784965**	0.046657	0.005183	1.02298	3.5e-05	0.000191	0	1.890289
MSGARCHsnorm	0.718505**	0.050828**	0.000302	1.030393	0.844493**	0.73126**	0.107508**	1.037806
GARCHsstd	0.068662**	0.028494	0.001259	1.156412	0.319515**	0.508443**	0.224408**	0.815419
MSGARCHsstd	0.53389**	0.061884**	0.000893	1.052632	0.844493**	0.73126**	0.015246	1.037806
GARCHsged	0.784965**	0.020155	0.001836	1.02298	0.156918**	0.316237**	0.273965**	0.74129
MSGARCHsged	0.654254**	0.024493	0.001054	1.037806	0.565949**	0.605186**	0.007519	1.111935

Table 1.15. Backtesting results for the commodities Soybeans, Wheat, Orange Juice (Agriculture commodity sector).

Confidence level		95%				99%		
	uc.LRp	cc.LRp	DQp	AE	uc.LRp	cc.LRp	DQp	AE
Sugar								
GARCHnorm	0.247085**	0.497879**	0.732411**	0.904374	0.096832**	0.201986**	0.033076	1.334322
MSGARCHnorm	0.127491**	0.268296**	0.767797**	0.874722	0.002628	0.010177	0.301284**	0.481838
GARCHstd	0.85319**	0.910519**	0.689048**	1.015567	0.319515**	0.240722**	0.096096**	0.815419
MSGARCHstd	0.729269**	0.542411**	0.456444**	0.97109	0.064408**	0.160255**	0.58733**	0.667161
GARCHged	0.211755**	0.449912**	0.560649**	0.896961	0.064408**	0.050112**	0.128026**	0.667161
MSGARCHged	0.127491**	0.268296**	0.090227**	0.874722	0.228641**	0.176721**	0.002815	0.778354
GARCHsnorm	0.376169**	0.635087**	0.664289**	0.926612	0.011746	0.037856	0.015267	1.519644
MSGARCHsnorm	0.936567**	0.960533**	0.02223	0.993328	0.102975**	0.080877**	0.028972	0.704225
GARCHsstd	0.592603**	0.745351**	0.606945**	1.045219	0.996913**	0.549507**	0.012688	1.000741
MSGARCHsstd	0.53389**	0.807748**	0.203627**	1.052632	0.848702**	0.509321**	0.23263**	0.963677
GARCHsged	0.376169**	0.673278**	0.767099**	0.926612	0.228641**	0.176721**	0.084522**	0.778354
MSGARCHsged	0.866417**	0.959651**	0.423016**	0.985915	0.429471**	0.311619**	0.137356**	0.852483
Soybean Oil								
GARCHnorm	0.247085**	0.01933	0.068365**	0.904374	0.260377**	0.098552**	0	1.223128
MSGARCHnorm	0.538857**	0.061719**	0.129585**	0.948851	0.447366**	0.112431**	2.00e-06	1.148999
GARCHstd	0.376169**	0.036285	0.069837**	0.926612	0.848702**	0.078492**	0	0.963677
MSGARCHstd	0.6633**	0.037262	0.0381	0.963677	0.844493**	0.102753**	0	1.037806
GARCHged	0.211755**	0.015324	0.054546**	0.896961	0.69809**	0.064088**	0	0.926612
MSGARCHged	0.180153**	0.012011	0.046329	0.889548	0.996913**	0.09181**	0	1.000741
GARCHsnorm	0.481144**	0.022126	0.060051**	0.941438	0.066456**	0.049959	0	1.371386
MSGARCHsnorm	0.922692**	0.035904	0.08193**	1.008154	0.029222	0.02937	1.00e-06	1.445515
GARCHsstd	0.992953**	0.031922	0.043938	1.000741	0.699371**	0.11023**	0	1.07487
MSGARCHsstd	0.592603**	0.055455**	0.043967	1.045219	0.447366**	0.112431**	0	1.148999
GARCHsged	0.286243**	0.024111	0.087524**	0.911787	0.996913**	0.09181**	0	1.000741
MSGARCHsged	0.992953**	0.07185**	0.073024**	1.000741	0.137853**	0.012129	0	1.297257
Soybean Meal								
GARCHnorm	6.00e-05	0.00032	0.011323	0.681987	0.260377**	0.352689**	0.001867	1.223128
MSGARCHnorm	0.0024	0.009907	0.066647**	0.755835	0.996913**	0.761058**	0.020671	1.000741
GARCHstd	0.003333	0.011939	0.094688**	0.763529	0.156918**	0.316237**	0.007041	0.74129
MSGARCHstd	0.004522	0.015539	0.123757**	0.770941	0.319515**	0.508443**	0.010816	0.815419
GARCHged	4.00e-06	2.3e-05	0.002788	0.637509	0.038264	0.104931**	0.00335	0.630096
MSGARCHged	0.000204	0.000995	0.0244	0.704225	0.556782**	0.678314**	0.031888	0.889548
GARCHsnorm	4.00e-06	2.3e-05	0.002749	0.637509	0.699371**	0.677201**	0.004273	1.07487
MSGARCHsnorm	0.003288	0.011234	0.0703**	0.763246	0.69809**	0.734075**	0.004694	0.926612
GARCHsstd	0.003333	0.011939	0.096203**	0.763529	0.319515**	0.508443**	0.005971	0.815419
MSGARCHsstd	0.023255	0.074293**	0.256971**	0.815419	0.69809**	0.734075**	0.007531	0.926612
GARCHsged	3.9e-05**	0.000214	0.011576	0.674574	0.156918**	0.316237**	0.002567	0.74129
MSGARCHsged	0.003333	0.013443	0.11186**	0.763529	0.556782**	0.678314**	0.007598	0.889548

Table 1.16. Backtesting results for the commodities Sugar, Soybean Oil, Soybean Meal (Agriculture commodity sector).

	Metals	Energy	Agriculture
Gold	0.75	0.14	0.25
Silver	1.00	0.71	0.67
Palladium	1.00	0.43	0.75
Copper	1.00	0.57	0.67
Zinc	0.75	0.14	0.67
Heating Oil	0.20	0.67	0.25
Low Sulfur Gasolio	0.80	1.00	0.58
Gasoline	0.60	0.83	0.50
Natural Gas	0.00	1.00	0.58
Ethanol	0.40	0.67	0.67
WTI Crude Oil	0.60	1.00	0.58
Natural Gas UK	0.20	0.50	0.17
Oats	0.80	0.29	0.82
Wheat	0.00	0.29	0.82
Cocoa	0.80	0.57	0.36
Corn	0.40	0.43	0.82
Cotton	1.00	0.43	1.00
Sugar	0.80	0.43	0.73
Soybean Oil	0.80	1.00	0.91
Soybean Meal	0.00	0.43	0.82
Orange Juice	0.60	0.57	0.64
Coffee	0.80	0.43	0.91
Soybeans	0.60	0.14	0.91
Rough Rice	0.60	0.71	0.73

Table 1.17. Degree of connectedness between each commodity and the three commodity sectors. Sample period: October 3, 2005 - March 25, 2022.

	Metals	Energy	Agriculture
Metals	0.90	0.40	0.60
Energy	0.40	0.81	0.48
Agriculture	0.60	0.48	0.79

Table 1.18. Degree of connectedness between the commodity sectors. Sample period: October 3, 2005 - March 25, 2022.

1.7.3 C. Tables

	Cluster analysis
Gold	1
Silver	1
Palladium	1
Copper	1
Zinc	1
Heating Oil	3
Low Sulfur Gasolio	3
Gasoline	3
Natural Gas	3
Ethanol	2
WTI Crude Oil	3
Natural Gas UK	3
Oats	2
Wheat	2
Cocoa	1
Corn	2
Cotton	2
Sugar	2
Soybean Oil	2
Soybean Meal	2
Orange Juice	3
Coffee	2
Soybeans	2
Rough Rice	2

Table 1.19. Cluster identification of the commodities in the sample in Figure 1.2. Number 1 denotes the red cluster, number 2 the green cluster, and number 3 the purple cluster. Sample period: October 3, 2005 - March 25, 2022.

Chapter 2

Expectile Hidden Markov Regression Models for Analyzing Cryptocurrency Returns

2.1 Introduction

In the last ten years, investors have been increasingly attracted by the exploit of the cryptocurrency market, mostly because of its peculiar characteristics. Born merely as a peer-to-peer electronic cash system (Nakamoto 2008), the 70 billion increase in market capitalization (in particular Bitcoin during 2016-2017), enormous price jumps and levels of high volatility that were never seen before have made cryptos a new category of investment assets. Their unusual behavior makes them prone to some speculative bubbles that may in turn threaten the stability of financial markets (Cheah & Fry 2015, Yarovaya et al. 2016). Being crucial to address the level of integration between cryptocurrencies and traditional financial assets, many contributions have analyzed the relationship with equities mainly relying on well-known econometric techniques such as GARCH models (Katsiampa et al. 2019, Guesmi et al. 2019), variance decomposition (Ji, Bouri, Gupta & Roubaud 2018, Corbet et al. 2018, Yi et al. 2018) and Granger causality test (Bouri et al. 2020*b*). Part of the related literature has focused on extreme returns by using models capturing the tail behavior, rather than inferring such occurrences from models based on conditional central tendency. For instance, Kristjanpoller et al. (2020) and Naeem et al. (2021) employ a multifractal asymmetric analysis, indicating the presence of heterogeneity in the cross-relationship between most cryptocurrencies and equity ETFs and showing different behaviors between upward and downward trends. Shahzad et al. (2022) investigate tail-based connectedness among major cryptocurrencies in extreme downward and upward market conditions using LASSO penalized quantile

regressions, while Zhang et al. (2021) apply a risk spillover approach based on generalized quantiles, showing the existence of a downside risk spillover between Bitcoin and traditional assets. In quantitative risk management, indeed, investigating the dynamic of extreme occurrences is of utmost importance for market participants and regulators. Among the different methods considered throughout the literature, quantile regression, introduced by Koenker & Bassett (1978), has represented a valid approach for modeling the entire distribution of returns while accounting for the well-known stylized facts, i.e., high kurtosis, skewness and serial correlation, that typically characterize financial assets. In the financial literature, the quantile regression framework has been positively applied to estimate and forecast Value at Risk (VaR) and quantile-based risk measures (Engle & Manganelli 2004, White et al. 2015, Taylor 2019, Merlo et al. 2021).

Several generalizations of the concept of quantiles have also been introduced over the years. One important extension is provided by the expectile regression (Newey & Powell 1987), which can be thought of as a generalization of the classical mean regression based on an asymmetric squared loss function. Similar to quantile regression, expectile regression allows to characterize the entire conditional distribution of a response variable, but possesses several advantages over the former. First, expectiles are more informative than quantiles since they rely on tail expectations whereas quantiles use only the information on whether an observation is below or above the predictor. Second, the squared loss is continuously differentiable which makes the estimators and their covariance matrix easier to compute using fast and efficient algorithms. For these reasons, expectile models have been implemented in several fields, such as longitudinal data (Tzavidis et al. 2016, Alfò et al. 2017, Barry et al. 2021), spatial analysis (Sobotka & Kneib 2012, Spiegel et al. 2020), life expectancy (Nigri et al. 2022), economics and finance (Taylor 2008, Kim & Lee 2016, Bellini & Di Bernardino 2017, Bottone et al. 2021). Especially in the context of risk management, expectiles have gained an important role as potential competitors to the VaR and the Expected Shortfall measures. Indeed, they possess several interesting properties in terms of risk measures (see for instance Bellini 2012, Bellini et al. 2014 and Ziegel 2016), and are the only risk measure that is both coherent (Artzner et al. 1999) and elicitable (Lambert et al. 2008). Moreover, when modeling financial time series, returns often exhibit a clustering behavior over time which cannot be captured by traditional homogeneous regression models. Risk managers and regulators are increasingly interested in determining whether, and how, their temporal evolution can be influenced by hidden variables, e.g., the state of the market, during tranquil and crisis periods. In this context, Hidden Markov Models (HMMs, see MacDonald & Zucchini 1997, Zucchini et al. 2016) have been successfully employed in the analysis of time series data, with applications to asset allocation

and stock returns as discussed in Mergner & Bulla (2008), De Angelis & Paas (2013), Nystrup et al. (2017) and Maruotti et al. (2019). Quantile regression methods have also been generalized to account for serial heterogeneity. For example, Liu (2016) consider a quantile autoregression in which the parameters are subject to regime shifts determined by the outcome of a latent, discrete-state Markov process, while Adam et al. (2019) propose a model-based clustering approach where groups are inferred from conditional quantiles; see also Ye et al. (2016), Maruotti et al. (2021) and Merlo et al. (2022) for other applications of regime-switching models to financial and environmental time series. In longitudinal data, Farcomeni (2012) and Marino et al. (2018) introduce linear quantile regression models where time-dependent unobserved heterogeneity is described through dynamic coefficients that evolve according to a homogeneous hidden Markov chain. Within a Bayesian framework, a quantile nonhomogeneous HMM for longitudinal data has been recently proposed by Liu et al. (2021). To the best of our knowledge, however, a HMM for estimating conditional expectiles has not yet been proposed in the literature.

Motivated by the advantages of expectiles and the versatility of HMMs, we develop a linear expectile hidden Markov regression model to analyze the tail relation between cryptocurrencies and traditional asset classes. The method introduced allows to examine the entire conditional distribution of returns given the hidden state and potential covariates, where the dynamics of returns over time is described by state-specific regression coefficients which follow a latent discrete homogeneous Markov chain. Inference about model parameters is carried out in a Maximum Likelihood (ML) approach using an Expectation-Maximization (EM) algorithm based on the asymmetric normal distribution of Waldmann et al. (2017) as working likelihood. From a risk management standpoint, the proposed methodology contributes to identify and control for potential inherent risks related to the participation in crypto exchanges to develop appropriate policies and risk assessment procedures.

The study period considered starts from September 2014 until October 2022, comprising numerous events that heavily impacted financial stability, as the Chinese stock market crash of 2015, the crypto currency bubble crisis in 2017-2018, the COVID-19 outbreak in 2020 and the Russian invasion of Ukraine at the beginning of 2022. Following Corbet et al. (2018), we model Bitcoin daily returns as a function of major stock and global market indices, including Crude Oil, Standard & Poor's 500 (S&P500), Gold COMEX daily closing prices and the Volatility Index (VIX). Our results show that Bitcoin returns exhibit a clear temporal clustering behavior in calm and turbulent periods, and they are strongly associated with traditional assets at low and high expectile levels.

In concluding, we also evaluate the performance of our approach in a simulation study, generating observations from a two-state HMM under two different sample

sizes and two different distributions for the error terms.
The rest of the paper is organized as follows. Section 2.2 briefly reviews the expectile regression. In Section 3.3 we specify the proposed model with the EM algorithm for estimating the model parameters and the computational aspects. In Section 3.4, we evaluate the performance of our proposal in a simulation study. Section 3.5 shows the empirical analysis and discusses the results obtained while Section 2.6 concludes.

2.2 Expectile regression

Expectile regression has been proposed by Newey & Powell (1987) as a "quantile-like" generalization of standard mean regression based on asymmetric least-squares estimation. Similarly to quantile regression of Koenker & Bassett (1978), this is an alternative approach for characterizing the entire conditional distribution of a response variable where the quantile loss function is substituted with an asymmetric squared loss function. Formally, the expectile of order $\tau \in (0,1)$ of a continuous response Y given the P-dimensional vector of covariates $\boldsymbol{X} = \boldsymbol{x}$, is defined as the minimizer, $\mu_{\boldsymbol{x}}(\tau)$, of the following problem:

$$\mu_{\boldsymbol{x}}(\tau) = \arg\min_{\mu \in \mathbb{R}} \mathbb{E}[\omega_\tau(Y - \mu_{\boldsymbol{x}}(\tau))], \tag{2.1}$$

where $\omega_\tau(u) = u^2|\tau - \mathbb{I}(u < 0)|$ is the asymmetric square loss and $\mathbb{I}(\cdot)$ denotes the indicator function.
In a regression framework, for a given τ, a linear expectile model is defined as $\mu_{\boldsymbol{x}}(\tau) = \boldsymbol{x}'\boldsymbol{\beta}(\tau)$, where $\boldsymbol{\beta}(\tau) \in \mathbb{R}^P$ is the regression parameter vector. If $\tau = \frac{1}{2}$, expectile regression reduces to the standard mean regression while for $\tau \neq \frac{1}{2}$ it allows to target the entire conditional distribution of the response given the covariates similarly to quantile regression. When we turn from quantiles to expectiles, the latter possess several advantages over the former. Particularly, we gain uniqueness of the ML solutions which is, indeed, not granted in the quantile context. From a computational standpoint, since the squared loss function $\omega_\tau(\cdot)$ is differentiable, the regression parameters $\boldsymbol{\beta}(\tau)$ can be estimated by efficient Iterative Reweighted Least Squares (IRLS), in contrast to algorithms used for fitting quantile regression models. Proofs of consistency, asymptotic normality and a robust estimator of the variance-covariance matrix of the regression coefficients for inference have been established in Newey & Powell (1987). These properties make the expectile regression versatile and computationally appealing from a statistical point of view.
In a likelihood approach, Gerlach & Chen (2015) and Waldmann et al. (2017) originally introduced the idea of expectile regression by employing a likelihood function that is based on the Asymmetric Normal (AN) distribution. The AN

distribution can be thought of as a generalization of the normal distribution to allow for non-zero skewness, having the following density:

$$f_Y(y) = \frac{2\sqrt{\tau(1-\tau)}}{\sqrt{\pi\sigma^2}(\sqrt{\tau}+\sqrt{1-\tau})} \exp\left[-\omega_\tau\left(\frac{y-\mu}{\sigma}\right)\right], \tag{2.2}$$

where $\mu \in \mathbb{R}$ is a location parameter corresponding to the τ-th expectile of Y, $\sigma > 0$ is a scale parameter and $\tau \in (0,1)$ determines the asymmetry of the distribution. Particularly, when $\tau = \frac{1}{2}$ the density in (2.2) reduces to the well-known normal distribution, and μ and σ coincide with its mean and standard deviation, respectively. As discussed by Waldmann et al. (2017), the minimization of the asymmetric squared loss function in (2.1) is equivalent, in terms of parameter estimates, to the maximization of the likelihood associated with the AN density.
In the following section, we extend the expectile regression to the HMM setting by using the AN distribution as working likelihood.

2.3 Methodology

In this section we describe the expectile hidden Markov regression model in order to take into account the temporal evolution of the time series under analysis. We then show how inference about model parameters can be carried out in a ML approach using the AN distribution introduced in the previous section. Formally, let $\{S_t\}_{t=1}^T$ be a latent, homogeneous, first-order Markov chain defined on the discrete state space $\{1,\dots,K\}$. Let $\pi_k = Pr(S_1 = k)$ be the initial probability of state k, $k = 1,\dots,K$, and $\pi_{k|j} = Pr(S_{t+1} = k|S_t = j)$, with $\sum_{k=1}^K \pi_{k|j} = 1$ and $\pi_{k|j} \geq 0$, denote the transition probability between states j and k, that is, the probability to visit state k at time $t+1$ from state j at time t, $j,k = 1,\dots,K$ and $t = 1,\dots,T$. More concisely, we collect the initial and transition probabilities in the K-dimensional vector $\boldsymbol{\pi}$ and in the $K \times K$ matrix $\boldsymbol{\Pi}$, respectively.

To build the proposed model, let Y_t denote a continuous observable response variable and $\boldsymbol{X}_t = (1, X_{t2},\dots,X_{tP})'$ be a vector of P exogenous covariates, with the first element being the intercept, at time $t = 1,\dots,T$. For a given expectile level $\tau \in (0,1)$, the proposed linear Expectile Hidden Markov Model (EHMM) is defined as follows:

$$Y_t = \boldsymbol{X}_t'\boldsymbol{\beta}_k(\tau) + \epsilon_{tk}(\tau), \tag{2.3}$$

with $\boldsymbol{\beta}_k(\tau) = (\beta_{1k}(\tau),\dots,\beta_{Pk}(\tau))' \in \mathbb{R}^P$ being a state-specific coefficient vector that assumes one of the values $\{\boldsymbol{\beta}_1(\tau),\dots,\boldsymbol{\beta}_K(\tau)\}$ depending on the outcome of the unobservable Markov chain S_t and where $\epsilon_{tk}(\tau)$ is the error term whose conditional τ-th expectile is assumed to be zero.
Extending the approach of Waldmann et al. (2017) to the HMM setting, we use

the AN distribution to describe the conditional distribution of the response given covariates and the state occupied by the latent process at time t, whose probability density function is now given by

$$f_Y(y_t|\boldsymbol{X}_t = \boldsymbol{x}_t, S_t = k) = \frac{2\sqrt{\tau(1-\tau)}}{\sqrt{\pi\sigma_k^2}(\sqrt{\tau}+\sqrt{1-\tau})} \exp\left[-\omega_\tau\left(\frac{y_t - \mu_{tk}}{\sigma_k}\right)\right], \tag{2.4}$$

where the location parameter μ_{tk} is defined by the linear model $\mu_{tk} = \boldsymbol{x}_t'\boldsymbol{\beta}_k(\tau)$.
In the following section we use the AN distribution as a working likelihood for estimating the model parameters in a regression framework.

2.3.1 Likelihood inference

In this section we consider a ML approach to make inference on model parameters. As is common for HMMs, and for latent variable models in general, we develop an EM algorithm (Baum et al. 1970) to estimate the parameters of the method proposed based on the observed data. To ease the notation, unless specified otherwise, hereinafter we omit the expectile level τ, yet all model parameters are allowed to depend on it.
For a given number of hidden states K, the EM algorithm runs on the complete log-likelihood function of the model introduced, which is defined as

$$\begin{aligned}\ell_c(\boldsymbol{\theta}_\tau) = \sum_{k=1}^{K}\gamma_1(k)\log\pi_k + \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{j=1}^{K}\xi_t(j,k)\log\pi_{k|j} \\ + \sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_t(k)\log f_Y(y_t|\boldsymbol{x}_t, S_t = k),\end{aligned} \tag{2.5}$$

where $\boldsymbol{\theta}_\tau = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K, \sigma_1, \ldots, \sigma_K, \boldsymbol{\pi}, \boldsymbol{\Pi})$ represents the vector of all model parameters, $\gamma_t(k)$ denotes a dummy variable equal to 1 if the latent process is in state k at occasion t and 0 otherwise, and $\xi_t(j,k)$ is a dummy variable equal to 1 if the process is in state j in $t-1$ and in state k at time t and 0 otherwise.

To estimate $\boldsymbol{\theta}_\tau$, the algorithm iterates between two steps, the E- and M-steps, until convergence, as outlined below.

E-step:

In the E-step, at the generic $(h+1)$-th iteration, the unobservable indicator variables $\gamma_t(k)$ and $\xi_t(j,k)$ in (3.12) are replaced by their conditional expectations given the observed data and the current parameter estimates $\boldsymbol{\theta}_\tau^{(h)}$. To compute such quantities we require the calculation of the probability of being in state k at time t given the observed sequence

$$\gamma_t^{(h)}(k) = P_{\boldsymbol{\theta}_\tau^{(h)}}(S_t = k|y_1, \ldots, y_T) \tag{2.6}$$

and the probability that at time $t-1$ the process is in state j and then in state k at time t, given the observed sequence

$$\xi_t^{(h)}(j,k) = P_{\boldsymbol{\theta}_\tau^{(h)}}(S_{t-1}=j, S_t=k|y_1,\dots,y_T). \tag{2.7}$$

The quantities in (3.13) and (3.14) can be obtained using the Forward-Backward algorithm of Welch (2003). Then, we use these to calculate the conditional expectation of the complete log-likelihood function in (3.12) given the observed data and the current estimates:

$$\begin{aligned} Q(\boldsymbol{\theta}_\tau|\boldsymbol{\theta}_\tau^{(h)}) = \sum_{k=1}^{K}\gamma_1^{(h)}(k)\log\pi_k + \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{j=1}^{K}\xi_t^{(h)}(j,k)\log\pi_{k|j} \\ + \sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_t^{(h)}(k)\log f_Y(y_t|\boldsymbol{x}_t, S_t=k). \end{aligned} \tag{2.8}$$

M-step:

In the M-step we maximize $Q(\boldsymbol{\theta}_\tau|\boldsymbol{\theta}_\tau^{(h)})$ in (3.15) with respect to $\boldsymbol{\theta}_\tau$ to obtain the update parameter estimates $\boldsymbol{\theta}_\tau^{(h+1)}$. The maximization of $Q(\boldsymbol{\theta}_\tau|\boldsymbol{\theta}_\tau^{(h)})$ can be partitioned into orthogonal subproblems, where the updating formulas for the hidden Markov chain and state-dependent regression parameters are obtained independently maximizing each of these terms. Formally, the initial probabilities π_k and transition probabilities $\pi_{k|j}$ are updated using:

$$\pi_k^{(h+1)} = \gamma_1^{(h)}(k), \quad k=1,\dots,K \tag{2.9}$$

and

$$\pi_{k|j}^{(h+1)} = \frac{\sum_{t=1}^{T}\xi_t^{(h)}(j,k)}{\sum_{t=1}^{T}\sum_{k=1}^{K}\xi_t^{(h)}(j,k)}, \quad j,k=1,\dots,K. \tag{2.10}$$

To update the regression coefficients, the first-order condition of (3.15) with respect to $\boldsymbol{\beta}_k$, $k=1,\dots,K$, yields

$$\frac{\partial Q(\boldsymbol{\theta}_\tau|\boldsymbol{\theta}_\tau^{(h)})}{\partial\boldsymbol{\beta}_k} \propto \sum_{t=1}^{T}\gamma_t^{(h)}(k)|\tau - \mathbb{I}(y_t < \boldsymbol{x}_t'\boldsymbol{\beta}_k)|\boldsymbol{x}_t(y_t - \boldsymbol{x}_t'\boldsymbol{\beta}_k) = \mathbf{0}_P, \tag{2.11}$$

so the M-step update expression for $\boldsymbol{\beta}_k$ is

$$\boldsymbol{\beta}_k^{(h+1)} = \Big(\sum_{t=1}^{T}\gamma_t^{(h)}(k)|\tau - \mathbb{I}(y_t < \boldsymbol{x}_t'\boldsymbol{\beta}_k)|\boldsymbol{x}_t\boldsymbol{x}_t'\Big)^{-1}\Big(\sum_{t=1}^{T}\gamma_t^{(h)}(k)|\tau - \mathbb{I}(y_t < \boldsymbol{x}_t'\boldsymbol{\beta}_k)|\boldsymbol{x}_t y_t\Big), \tag{2.12}$$

which can be computed using IRLS for cross-sectional data with appropriate weights. Similarly, from the first-order condition of (3.15) with respect to the scale parameters we obtain the following M-step update formula for σ_k^2:

$$\sigma_k^{2(h+1)} = \frac{2}{\sum_{t=1}^{T}\gamma_t^{(h)}(k)}\sum_{t=1}^{T}\gamma_t^{(h)}(k)|\tau - \mathbb{I}(y_t < \boldsymbol{x}_t'\boldsymbol{\beta}_k^{(h+1)})|(y_t - \boldsymbol{x}_t'\boldsymbol{\beta}_k^{(h+1)})^2. \tag{2.13}$$

The E- and M- steps are alternated until convergence, that is when the observed likelihood between two consecutive iterations is smaller than a predetermined threshold. In this paper, we set this threshold criterion equal to 10^{-4}.
Following Maruotti et al. (2021) and Merlo et al. (2022), for fixed τ and K we initialize the EM algorithm by providing the initial states partition, $\{S_t^{(0)}\}_{t=1}^{T}$, according to a Multinomial distribution with probabilities $1/K$. From the generated partition, the elements of $\mathbf{\Pi}^{(0)}$ are computed as proportions of transition, while we obtain $\boldsymbol{\beta}_k^{(0)}$ and $\sigma_k^{(0)}$ by fitting mean regressions on the observations within state k. To deal with the possibility of multiple roots of the likelihood equation and better explore the parameter space, we fit the proposed EHMM using a multiple random starts strategy with different starting partitions and retain the solution corresponding to the maximum likelihood value.
Once we computed the ML estimate of the model parameters, to estimate the standard errors we employ the parametric bootstrap scheme of Visser et al. (2000). In practice, we refit the model to R bootstrap samples and approximate the standard error of each model parameter with the corresponding standard deviation of the bootstrap estimates.

2.4 Simulation study

We conduct a simulation study to validate the performance of our method under different scenarios in terms of recovering the true values of the parameters and the clustering performance. Similar to the work of Maruotti et al. (2021), we analyze two different sample sizes ($T = 500,\ T = 1000$) and two different distributions for the error term. For each scenario we conduct 500 Monte Carlo simulations. We draw observations from a two state HMM ($K = 2$) using the following data generating process:

$$Y_t = \begin{cases} -1 + 2X_t + \epsilon_{t1}, & S_t = 1 \\ 1 - 2X_t + \epsilon_{t2}, & S_t = 2, \end{cases} \tag{2.14}$$

with $X_t \sim \mathcal{N}(0, 1)$. We consider two distributions for the error terms in (2.14). In the first scenario, ϵ_{tk} is generated from a normal distribution with standard deviation 1, for $k = 1, 2$. In the second one, ϵ_{tk} is generated from a skew-t distribution with 5 degrees of freedom and asymmetry parameter 2, for $k = 1, 2$. Finally, the matrix of transition probabilities is set equal to $\mathbf{\Pi} = \left(\begin{smallmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{smallmatrix}\right)$.

In order to assess the validity of the model we fit the proposed EHMM at five expectile levels, i.e., $\tau = \{0.10, 0.25, 0.50, 0.75, 0.90\}$, and compute the bias and standard errors associated to the state-specific coefficients, averaged over the Monte

Carlo replications, for each combination of sample size and error distribution. Tables 2.1 and 2.2 report the simulation outputs for the normal and skew-t distributions, respectively. As can be observed, the precision of the estimates is higher at the center of the distribution rather than on the tails, mainly due to the reduced number of observations at extreme expectile levels, but the bias always remains under control. Evidently, in Table 2.2 a higher standard deviation shows up for the skew-t distribution due to the asymmetry and heavier tails than the normal density, but both the bias and the standard deviation tend to decrease as the sample size increases. Concerning the hidden process, given the true values of the transition probabilities in $\mathbf{\Pi}$, we see that the coefficients corresponding to the first state are estimated with lower precision because fewer transitions occur from one state to the other, as expected.

To evaluate the ability in recovering the true states partition we consider the Adjusted Rand Index (ARI) of Hubert & Arabie (1985). The state partition provided by the fitted models is obtained by taking the maximum, $\max_k \gamma_t(k)$, posteriori probability for every $t = 1, \ldots, T$, and report the box-plots of ARI for the posterior probabilities in Figure 3.1 for the four settings considered. Firstly, we observe that the distribution error plays a fundamental role in estimating the true states partition as, for all five expectile levels, we obtain a better clustering performance for the model with Gaussian errors with respect to the skew-t case. Secondly, the goodness of the clustering obtained partially depends on the specific expectile level, being the values slightly higher at the mean ($\tau = 0.50$) than at the tails. Finally, when increasing the sample size to $T = 1000$, results slightly improve reporting a lower variability for both error distributions. Overall, the proposed EHMM is able to recover the true values of the parameters and the true state partition highly satisfactory in all the cases examined.

2.5 Empirical application

In this section we apply the methodology proposed to analyze the Bitcoin daily returns as a function of global leading financial indices. Over the last decade, cryptocurrencies and in particular the Bitcoin market played a leading role, attracting attentions of researchers and investors. Their peculiar characteristics, such their extreme price volatility, driven by market speculation and technology applications, often lead to price bubbles, euphoria and market instability. In order to address these periods of upheaval, it is crucial to understand the association between Bitcoin and globally relevant market indices in such circumstances of financial turmoil. Consistently with Corbet et al. (2018), here we consider the Bitcoin, Crude Oil, Standard & Poor's 500 (S&P500), Gold COMEX daily closing prices and the

τ	0.10		0.25		0.50		0.75		0.90	
	Bias	Std.Err	Bias	Std.Err	Bias	Std.Err	Bias	Std.Err	Bias	Std.Err
Panel A: T=500										
State 1										
$\beta_{1,1} = -1$	0.020	0.093	0.010	0.076	-0.002	0.073	-0.018	0.080	-0.048	0.100
$\beta_{2,1} = 2$	0.001*	0.109	0.001	0.093	0.004	0.087	0.010	0.089	0.023	0.101
State 2										
$\beta_{1,2} = 1$	0.040	0.055	0.013	0.041	-0.002	0.037	-0.013	0.039	-0.027	0.047
$\beta_{2,2} = -2$	-0.008	0.067	0.001*	0.058	0.001	0.055	-0.003	0.058	-0.012	0.068
Panel B: T = 1000										
State 1										
$\beta_{1,1} = -1$	0.021	0.068	0.010	0.055	-0.001	0.051	-0.016	0.056	-0.042	0.070
$\beta_{2,1} = 2$	0.006	0.071	0.003	0.060	0.004	0.057	0.008	0.060	0.017	0.069
State 2										
$\beta_{1,2} = 1$	0.039	0.038	0.014	0.029	0.001*	0.026	-0.010	0.028	-0.023	0.034
$\beta_{2,2} = -2$	-0.012	0.053	-0.004	0.045	-0.002	0.043	-0.005	0.044	-0.014	0.050

Table 2.1. Bias and standard error values of the state-regression parameter estimates with Gaussian distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B). * represents values smaller (in absolute value) than 0.001.

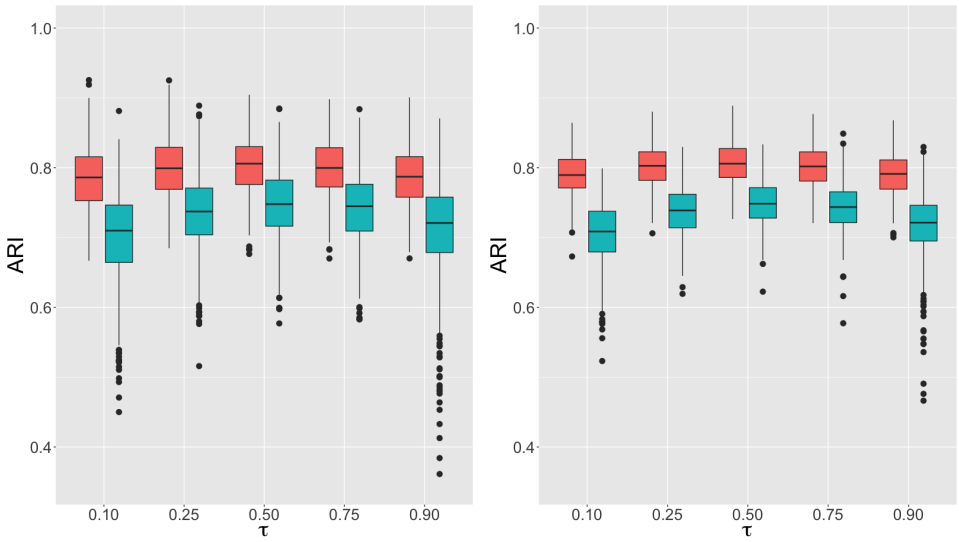

Figure 2.1. From left to right, box-plots of ARI for the posterior probabilities for Gaussian (red) and skew-t (blue) distributed errors with $T = 500$ and $T = 1000$.

Volatility Index (VIX) from September 2014 to October 2022. All series are expressed in US dollars and have been downloaded from the Yahoo finance database. Daily returns with continuous compounding are calculated taking the logarithm of the difference between closing prices in consecutive trading days and then multiplied by 100, i.e., $r_t = 100 \cdot \log(P_t/P_{t-1})$ where P_t is the closing price on day t, for a total of $T = 2025$ observations.

τ	0.10		0.25		0.50		0.75		0.90	
	Bias	Std.Err	Bias	Std.Err	Bias	Std.Err	Bias	Std.Err	Bias	Std.Err
Panel A: T=500										
State 1										
$\beta_{1,1}$ = -1	-0.064	0.228	-0.009	0.119	-0.004	0.099	-0.018	0.166	0.083	0.489
$\beta_{2,1}$ = 2	-0.153	0.319	-0.049	0.183	0.001	0.125	0.027	0.161	-0.045	0.471
State 2										
$\beta_{1,2}$ = 1	0.181	0.144	0.066	0.073	0.013	0.051	-0.018	0.060	-0.052	0.131
$\beta_{2,2}$ = -2	-0.060	0.100	-0.029	0.080	-0.015	0.074	-0.015	0.079	-0.038	0.103
Panel B: T = 1000										
State 1										
$\beta_{1,1}$ = -1	-0.053	0.162	0.001	0.077	-0.004	0.069	-0.022	0.123	0.025	0.369
$\beta_{2,1}$ = 2	-0.133	0.236	-0.024	0.111	0.013	0.082	0.039	0.116	0.027	0.319
State 2										
$\beta_{1,2}$ = 1	0.168	0.101	0.057	0.048	0.010	0.035	-0.019	0.041	-0.059	0.083
$\beta_{2,2}$ = -2	-0.067	0.067	-0.033	0.052	-0.016	0.048	-0.015	0.054	-0.031	0.072

Table 2.2. Bias and standard error values of the state-regression parameter estimates with skew-t distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B).

The considered timespan is marked by numerous crises that may have impacted cross-market integration patterns, including the Chinese stock market crash of 2015, the cryptocurrency bubble crisis in 2017-2018, the COVID-19 pandemic and the Russian invasion of Ukraine at the beginning of 2022, which have caused unprecedented levels of uncertainty and risk. In Table 2.3 we report the list of examined variables and the summary statistics for the whole sample. First thing to notice is that Bitcoin is generally much more volatile than the other assets, having the highest standard deviation. The Bitcoin returns also show very high negative skewness and very high kurtosis, as well as S&P500. The highest level of kurtosis is reported by Crude Oil, probably determined by the prices' fluctuations after the COVID-19 outbreak. On the other side, the large positive skewness of VIX indicates longer and fatter tails on the right side of the distribution, highlighting the well-known inverse relationship with the S&P500. In concluding, the Augmented Dickey-Fuller (ADF) test Dickey & Fuller (1979) shows that all daily returns are stationary at the 1% level of significance. Following these considerations, and motivated by the reforms considered by markets authorities to protect investors and preserve stability in response to cryptocurrencies' downturns, the proposed EHMM can provide insights into the temporal evolution of Bitcoin returns and describe how this is affected by rapid changes in markets volatility.

To this end, we consider the following linear EHMM:

$$\mu_{tk}^{Bitcoin} = \beta_{1k}(\tau) + \beta_{2k}(\tau) r_t^{Crude\ Oil} + \beta_{3k}(\tau) r_t^{S\&P500} + \beta_{4k}(\tau) r_t^{Gold} + \beta_{5k}(\tau) r_t^{VIX}, \tag{2.15}$$

	Minimum	Mean	Maximum	Std.Err.	Skewness	Kurtosis	Jarque-Bera test	ADF test
Bitcoin	-46.4730	0.1859	22.5119	4.6165	-0.6817	8.7172	**6568.469**	**-11.117**
Crude Oil	-28.2206	0.0280	31.9634	3.1077	0.0942	21.1219	**37645.730**	**-10.324**
S&P500	-12.7652	0.0316	8.9683	1.1716	-0.9033	16.3473	**22823.380**	**-12.390**
Gold	-5.1069	0.0152	5.7775	0.9358	-0.0698	4.1741	**1471.712**	**-12.292**
VIX	-29.9831	0.0379	76.8245	8.3414	1.2683	6.6648	**4290.737**	**-14.209**

Table 2.3. Descriptive statistics for the whole sample. The Jarque-Bera test and the ADF test statistics are displayed in boldface when the null hypothesis is rejected at the 1% significance level.

with $\mu_{tk}^{Bitcoin}$ corresponding to the τ-th conditional expectile of Bitcoin return at time t in state k, while $r_t^{Crude\ Oil}$ denotes the return of the same date for Crude Oil, and similarly for the other indices.

As a first step of the empirical analysis, in order to select the number of latent states we fit the proposed EHMM for different values of K varying from 2 to 5 at three expectile levels $\tau = \{0.10, 0.50, 0.90\}$, which allow us to focus on both downside and upside risks. To compare models with differing number of states, Table 2.4 reports three widely employed penalized likelihood selection criteria for K, namely the AIC (Akaike 1998), the BIC (Schwarz et al. 1978) and the ICL (Biernacki et al. 2000). As one can see, the AIC selects 5, or more, states, while BIC chooses $K = 4$ for all three expectile levels. This should not be surprising since the AIC tends to overestimate the true number of hidden states. On the contrary, ICL favors a more parsimonious choice as $K = 2$ is always considered to be the optimal number of states at $\tau = 0.10$, $\tau = 0.50$ and $\tau = 0.90$. For these reasons, and in order to clearly identify high and low volatility market conditions we thus consider the proposed EHMM with $K = 2$ states for all three τ levels.

For the selected models, we report the clustering results in Figure 2.2 at the investigated expectile levels. Each plot shows the time series of Bitcoin daily returns colored according to the estimated posterior probability of class membership, $\max_k \gamma_t(k)$, with the vertical dashed lines representing globally relevant events such as the Chinese stock market crash in 2015, the cryptocurrencies crash at the beginning of 2018, the COVID-19 market crash in March 2020, Biden's election at the USA presidency in November 2020 and the Russian invasion of Ukraine at the beginning of 2022. Here we clearly see that the latent components can be associated to specific market regimes characterized by low and high volatility periods. Specifically, light-blue points (State 1) tend to identify low returns, while dark-blue ones (State 2) correspond to periods of extreme positive and negative returns.

Moving on to the state-specific model parameters, Table 2.5 shows the parameter estimates along with the standard errors (in brackets) computed by using the parametric bootstrap technique illustrated in Section 3.3.1 over $R = 1000$ resamples.

First, consistently with the quantile regression literature, the intercepts are increasing with τ, with State 1 having lower values than State 2 for all τs. Second, it is interesting to observe that in the not-at-risk state (State 1) the S&P500, Gold and the VIX index positively influence extreme left-tail ($\tau = 0.10$) movements of Bitcoin returns, while only S&P500 and Gold significantly influence the right-tail ($\tau = 0.90$) expectiles of the cryptocurrency, exposing a connection during high volatility periods between traditional financial markets and Bitcoin both for negative and positive returns. At $\tau = 0.50$, instead, Bitcoin can be considered as a weak hedge during high volatility periods since it is not statistically associated with all the assets considered (Bouri, Jalkh, Molnár & Roubaud 2017, Bouri, Molnár, Azzi, Roubaud & Hagfors 2017). In the at-risk state (State 2) we observe a positive influence of the S&P500 and Gold across the conditional distribution of returns. Also, one can see that Crude Oil is negatively associated with the crypto returns at the 10-th expectile. This finding is in line with Bouri et al. (2020*a*) but it is contrary to the works of Dyhrberg (2016) and Corbet et al. (2018), which may be due to the events and crises occurred in the last years.

Finally, the estimated state-dependent scale parameter σ_1 reflects more stable periods for the first state, meanwhile σ_2 contemplates rapid (positive and negative) peak and burst returns for the second state, confirming the graphical analysis conducted in Figure 2.2.

	$\tau = 0.10$	$\tau = 0.50$	$\tau = 0.90$
AIC			
$K = 2$	11347.4122	11231.4286	11389.7954
$K = 3$	11210.6074	11126.1327	11204.4727
$K = 4$	11109.3930	11051.9556	11115.4655
$K = 5$	**11055.7980**	**11014.8696**	**11079.7912**
BIC			
$K = 2$	11431.6121	11315.6285	11473.9953
$K = 3$	11356.5539	11272.0791	11350.4192
$K = 4$	**11328.3127**	**11270.8753**	**11334.3852**
$K = 5$	11358.9175	11317.9892	11382.9108
ICL			
$K = 2$	**12784.4362**	**12787.4471**	**12926.5182**
$K = 3$	13490.5599	13406.8221	13616.8609
$K = 4$	13880.2346	13487.8461	13759.7332
$K = 5$	14111.8066	13511.4108	13908.7362

Table 2.4. AIC, BIC and ICL values with varying number of states for the investigated expectile levels. Bold font highlights the best values for the considered criteria (lower-is-better).

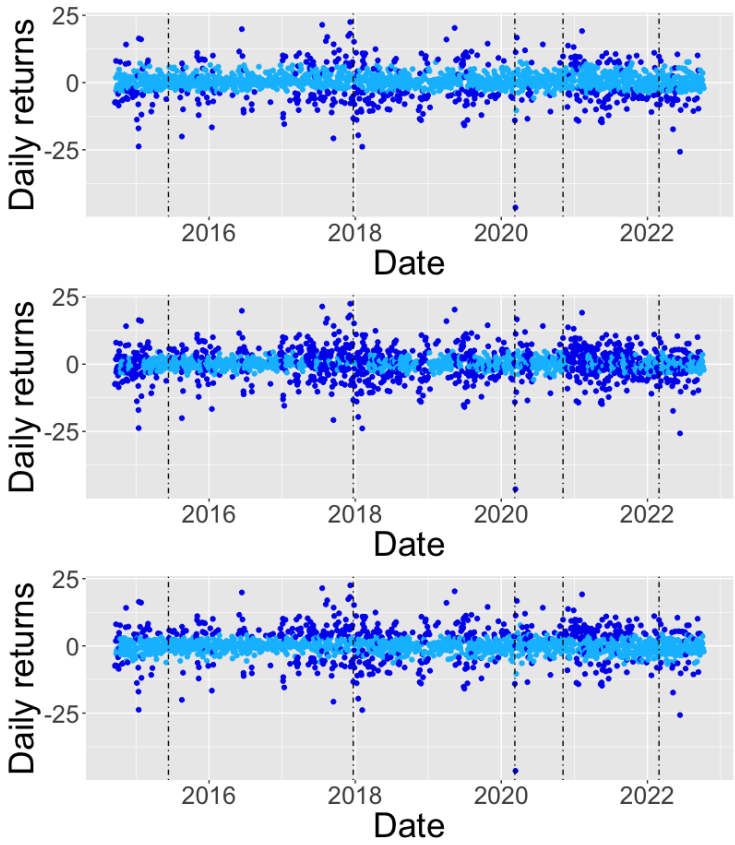

Figure 2.2. From top to bottom, Bitcoin returns series classified according to the estimated posterior probability of class membership at $\tau = 0.10$, $\tau = 0.50$ and $\tau = 0.90$. Vertical dashed lines indicate globally relevant events in the financial markets that occurred in 2015,06; 2017,12; 2020,03; 2020,11; and 2022,02.

2.6 Conclusions

The increasing popularity and importance of Bitcoin in the financial landscape have made scholars and practitioners interrogated about its properties and its relation to other assets. In this regard, we contribute to the existing literature in two ways. From a theoretical standpoint, we develop a linear expectile hidden Markov model for the analysis of time series where temporal behaviors of the data are captured via time-dependent coefficients that follow an unobservable discrete homogeneous Markov chain. The proposed method enables us to model the entire conditional distribution of asset returns and, at the same time, to grasp unobserved serial heterogeneity and rapid volatility jumps that would otherwise go undetected. From a practical point of view, we analyze the association between Bitcoin and a collection of global market indices, not only at the average, but also during times of market stress.

	Intercept	Crude Oil	S&P500	Gold	VIX	σ_k
State 1						
$\tau = 0.10$	**-1.036 (0.280)**	0.024 (0.021)	**0.595 (0.096)**	**0.189 (0.072)**	**0.029 (0.012)**	1.433 (0.040)
$\tau = 0.50$	0.122 (0.158)	0.031 (0.072)	0.409 (0.383)	0.263 (0.249)	0.009 (0.036)	1.695 (0.062)
$\tau = 0.90$	**1.297 (0.061)**	-0.009 (0.020)	**0.589 (0.088)**	**0.134 (0.065)**	0.014 (0.011)	1.335 (0.041)
State 2						
$\tau = 0.10$	**-6.52 (0.060)**	**-0.256 (0.096)**	**2.072 (0.476)**	**1.032 (0.320)**	-0.055 (0.058)	4.964 (0.157)
$\tau = 0.50$	0.242 (0.092)	-0.056 (0.055)	**1.087 (0.357)**	**0.613 (0.214)**	-0.025 (0.026)	6.164 (0.169)
$\tau = 0.90$	**6.244 (0.229)**	0.017 (0.079)	**0.948 (0.291)**	**0.835 (0.249)**	-0.002 (0.041)	4.692 (0.128)

Table 2.5. State-specific parameter estimates for three expectile levels, with bootstrapped standard errors (in brackets) obtained over 1000 replications. Point estimates are displayed in boldface when significant at the standard 5% level.

Empirically, we find evidence of strong and positive interrelations among Bitcoin returns and S&P500 and Gold, and, at the same time, we observe the capacity of the Bitcoin of working as a weak hedge during high volatility periods, contributing to the existing strands of literature on the subject (Baur et al. 2018, Corbet et al. 2018, 2019). Its partial capacity of being a weak hedge but not a safe haven it is consistent with the excess volatility of Bitcoin and indications that assets with no history as a safe haven are unlikely to be considered "safe" in an economic or financial crisis (Baur et al. 2018).

As a possible next step, our methodology could be extended to the hidden semi-Markov model setting where the sojourn-distributions, that is, the distributions of the number of consecutive time points that the chain spends in each state, can be modeled by the researcher using either parametric or non-parametric approaches instead of assuming geometric sojourn densities as in HMMs.

Chapter 3

Expectile Copula-Based Hidden Markov Regression Models for the Analysis of the Cryptocurrency Market

3.1 Introduction

Since the creation of Bitcoin more than a decade ago (Nakamoto 2008), exchange volumes have expanded tremendously, facilitated by the speed of transaction and the lack of any central authority or financial intermediary. This exploitation has piqued the interest of policymakers, risk managers and academics in the peculiar characteristics of cryptocurrencies, which can serve both as an efficient payment method and an investment asset. The rapid development of many alternative crypto assets attempting to replicate the path of the Bitcoin has resulted in massive speculative manoeuvres, which have increased the willingness of investors and practitioners to comprehend this challenging financial market (Borri 2019). Speculative periodic crypto-bubbles have repeatedly jeopardized the stability of financial markets and have been extensively studied, especially after the 2017 boom and the subsequent catastrophic crash in 2018 (Cheah & Fry 2015, Cheung et al. 2015, Corbet et al. 2018, Agosto & Cafferata 2020, Xiong et al. 2020). At the beginning of 2020, the global COVID-19 pandemic and the prospect of a global recession exacerbated the volatility of cryptoassets, prompting practitioners to speculate about the impact of the pandemic on digital currency returns and volatility, as well as their effect on international stock markets. Different aspects of cryptocurrencies, such as long-memory and efficiency (Duan et al. 2021, López-Martín et al. 2021, Assaf et al.

2022), hedging properties (Demir et al. 2020, Das et al. 2020), and relationships with other asset classes, have been investigated in the literature. In particular, empirical studies focused on whether cryptocurrencies could be used as optimal instruments to diversify investors' portfolios. Conlon & McGee (2020) analyzed if Bitcoin could be used as a safe-haven for the Standard & Poor's 500 (S&P500). Mariana et al. (2021) similarly tested Bitcoin and Ethereum safe-haven properties, while Corbet et al. (2020) studied the potential increases in volatility or correlation between Bitcoin and traditional markets and commodities, such as gold and oil prices. In the last years, the financial literature on the empirical characteristics of digital currencies have documented the existence of stylized facts, such as volatility clustering, asymmetry and leptokurticity, and the presence of spillover effects within the crypto market and toward other global financial markets. Consequently, it is indeed of utmost importance to be able to develop adequate statistical tools to take into account all these features. In the financial literature, Hidden Markov Models (HMMs, see MacDonald & Zucchini 1997, Zucchini et al. 2016) have been successfully employed in understanding whether, and how, time series temporal evolution can be influenced by hidden variables during tranquil and crisis periods. Numerous HMMs applications can be found in asset allocation and stock returns, as discussed for example in Mergner & Bulla (2008), De Angelis & Paas (2013), Nystrup et al. (2017), Maruotti et al. (2019). In the context of cryptocurrencies time series, Giudici & Abu Hashish (2020), Caferra & Vidal-Tomás (2021), Pennoni et al. (2022) and Cremaschini et al. (2022) consider latent Markov processes to analyze volatility clustering and serial heterogeneity of crypto returns. Besides the clustering behavior of returns, the analysis of the dynamics of extreme returns is of the utmost importance for regulators and policymakers for modeling the entire distribution of returns while accounting for the well-known stylized facts, i.e., high kurtosis, skewness and serial correlation. In order to address these features, since the seminal work of Koenker & Bassett (1978), quantile regression represented a valid approach and a widely used technique in many empirical applications; see Koenker (2005) and Koenker et al. (2017). It allows to model the conditional quantiles of a response variable with respect to a set of covariates, providing a much more complete picture of the conditional distribution compared with traditional mean regression. Quantile models have been extensively applied in finance and economics for estimating Value at Risk (VaR) and quantile-based risk measures (Engle & Manganelli 2004, White et al. 2015, Taylor 2019, Merlo et al. 2021). One of the most relevant extension related to quantile regression is provided by the expectile regression (Newey & Powell 1987), which is a "quantile-like" generalization of the mean regression using an asymmetric squared loss function. Even though many quantile regression methods are now well consolidated in the literature, few studies have been conducted on tail

events and their link to traditional assets within the context of digital currencies. Borri (2019) employs the Conditional VaR to estimate the conditional tail risk in the cryptocurrency markets, indicating that cryptocurrencies are highly exposed to tail risk within the cryptocurrency markets but are disconnected from other global assets. Ciner et al. (2022) analyze cryptocurrency returns during highly volatile period of the COVID-19 pandemic using penalized quantile regressions, finding an impact of Gold and S&P500 at the median of the crypto return distributions while, in the case of expectiles, Foroni et al. (2023) propose an expectile HMM for the analysis of Bitcoin daily returns. These proposals, however, are confined to the modeling of univariate financial time series only. When multiple cryptocurrencies returns are analyzed jointly, their dependence structure must be incorporated in the modeling framework in order to provide adequate risk control measures and produce effective asset allocation and diversification strategies. In these cases, taking into account the degree of association among different digital currencies, that cannot be detected by univariate methods, could be extremely important also for regulatory interventions. To address all these features simultaneously, we develop multivariate hidden Markov models for estimating conditional quantiles and expectiles of multiple cryptocurrencies returns using regime-switching copulas in a regression framework. On the one side, to model the temporal evolution of returns we introduce in the regression model time-dependent coefficients evolving according to a discrete, homogeneous latent Markov chain. On the other side, to take into account for the time-varying dependence structure of returns, we consider elliptical copulas defined by state-specific parameters. With this work, we unify in a common approach quantile and expectile HMMs with copulas, which allow us to pursue a two-fold goal. First, we incorporate within state-dependencies among the cryptocurrencies, which is crucial for investors whose investment portfolios contain a portion of crypto-assets as well as for policymakers whose role is to maintain the stability of financial markets. Second, we investigate the relationship among crypto and traditional financial assets at different volatility states, and in lower and upper tails of cryptocurrency return distributions. The estimation is carried out in a Maximum Likelihood (ML) framework using, respectively, the Asymmetric Laplace (Yu & Moyeed 2001) and Normal (Waldmann et al. 2017) distributions for quantiles and expectiles as working likelihoods through suitable Expectation-Maximization (EM) algorithms. The good performances of our methods are illustrated through a simulation study generating observations from a bivariate two-state HMM under different sample sizes, error distributions and copula functions. The real data analysis considers daily returns from July 2017 until December 2022, which comprise numerous events that heavily impacted financial stability, as the crypto currency bubble crisis in 2017-2018, the COVID-19 outbreak in 2020, Biden's election at the USA presidency in November 2020 and the Russian

invasion of Ukraine at the beginning of 2022. As for the choice of the dependent variables, we select five cryptocurrencies, namely Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), and Bitcoin Cash (BCH), following the criteria adopted in Pennoni et al. (2022). Digital currencies are modeled as functions of major stock and global market indices, including S&P500, S&P US Treasury Bond, US dollar index, WTI Crude Oil and Gold COMEX daily closing prices. Our results show that cryptocurrency returns exhibit a clear temporal clustering behavior in calm and turbulent periods, and the association with traditional financial assets is strong at extreme tails of returns distribution, especially with S&P500, S&P US Treasury Bond and Gold.

The rest of the paper is organized as follows. Section 3.2 briefly reviews univariate quantile and expectile regressions. In Section 3.3 we describe the proposed models, the EM algorithms for estimating the model parameters and the computational aspects of the estimation procedure. In Section 3.4, we evaluate the performance of our proposal in a simulation study. Section 3.5 is devoted to the empirical analysis and discusses the results obtained while Section 3.6 concludes.

3.2 Preliminaries on quantile and expectile regressions

In order to better explain the proposed models, in this section we briefly revise the univariate quantile and expectile regressions.

For a continuous response variable, quantile and expectile regressions provide a much more flexible approach and complete picture of the conditional distribution of the response than classical regression models targeting the mean. The former, introduced by Koenker & Bassett (1978), can be considered as a generalisation of median regression, while the latter, proposed by Newey & Powell (1987), can be thought of as a generalization of mean regression based on asymmetric least-squares estimation. More generally, both quantiles and expectiles can be embedded in a common framework within the wider class of generalized quantiles defined as the minimizers of an asymmetric l-power loss function.

Formally, let $\tau \in (0,1)$ and consider the following asymmetric loss function

$$\omega_{l,\tau}(u) = |u|^l \cdot |\tau - \mathbb{I}(u < 0)|, \quad u \in \mathbb{R} \tag{3.1}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function.

Given a set of covariates $\boldsymbol{X} = \boldsymbol{x}$, it is easy to see that when $l = 1$, the conditional quantile of order τ of a continuous response Y is defined as

$$q_{\boldsymbol{x}}(\tau) = \arg\min_{m \in \mathbb{R}} \mathbb{E}[\omega_{1,\tau}(Y - m_{\boldsymbol{x}}(\tau))] \tag{3.2}$$

meanwhile, for $l = 2$, the τ-th conditional expectile of Y is defined as

$$e_{\boldsymbol{x}}(\tau) = \arg\min_{m \in \mathbb{R}} \mathbb{E}[\omega_{2,\tau}(Y - m_{\boldsymbol{x}}(\tau))]. \tag{3.3}$$

In particular, $q_{\boldsymbol{x}}(\tau)$ and $e_{\boldsymbol{x}}(\tau)$ with $\tau = \frac{1}{2}$ correspond respectively to the conditional median and the mean of Y given covariates $\boldsymbol{x}$. When $\tau \neq \frac{1}{2}$, both methods allow to target the entire conditional distribution of the response. In practice, quantiles have a more intuitive interpretation than expectiles even if they target essentially the same part of the distribution of interest. However, despite the popularity and the easy interpretability of the former, the latter offer some advantage: (a) we gain uniqueness of the ML solutions which is not granted in the quantile context; (b) from a computational standpoint, since the squared loss function $\omega_{2,\tau}(\cdot)$ is differentiable, $e_{\boldsymbol{x}}(\tau)$ can be estimated by efficient Iterative Reweighted Least Squares (IRLS), in contrast to algorithms used for fitting quantile regression models.

From a likelihood perspective, both methods have been implemented in a ML approach by exploiting the relationship between the minimization of the loss function in (3.1) and the maximization of a likelihood function formed by combining independently distributed densities with kernel function $\omega_{l,\tau}(\cdot)$. That is,

$$f(y; \mu, \sigma, \tau) = B_{l,\tau}(\sigma) \exp\left[-\omega_{l,\tau}\left(\frac{y-\mu}{\sigma}\right)\right] \tag{3.4}$$

where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter and $B_{l,\tau}(\sigma)$ is a normalizing constant that ensures the density integrates to one. In the case of quantiles for $l = 1$, the density in (3.4) reduces to the Asymmetric Laplace (AL) distribution introduced by Yu & Moyeed (2001), $f_{AL}(y; \mu, \delta, \tau)$, where μ coincides with the τ-th quantile of Y with $B_{1,\tau}(\sigma) = \frac{\tau(1-\tau)}{\sigma}$. As regards to expectiles when $l = 2$, (3.4) corresponds to the Asymmetric Normal (AN) distribution proposed by Gerlach & Chen (2015) and Waldmann et al. (2017), $f_{AN}(y; \mu, \sigma, \tau)$, and μ is the τ-th expectile of Y with $B_{2,\tau}(\sigma) = \frac{2\sqrt{\tau(1-\tau)}}{\sqrt{\pi\sigma^2}(\sqrt{\tau}+\sqrt{1-\tau})}$. It is easy to verify that in both cases, as discussed in Yu & Moyeed (2001) and Waldmann et al. (2013) respectively, the minimization of the respective expected loss functions $\omega_{1,\tau}(\cdot)$ and $\omega_{2,\tau}(\cdot)$ is equivalent, in terms of parameter estimates, to the maximization of the likelihood functions associated with the AL and AN densities.

In the following section, we extend quantile and expectile regressions to the HMM setting for the analysis of multivariate time series by considering elliptical copulas.

3.3 Methodology

In this section we formally introduce the quantile and expectile copula-based hidden Markov regression models. We then build suitable EM algorithms for ML estimation

using the AL and AN distributions as working likelihoods for the proposed models.

Formally, let $\{S_t\}_{t=1}^T$ be a latent, homogeneous, first-order Markov chain defined on the discrete state space $\{1,\dots,K\}$. Let $\pi_k = Pr(S_1 = k)$ be the initial probability of state k, $k = 1,\dots,K$, and $\pi_{k|j} = Pr(S_{t+1} = k|S_t = j)$, with $\sum_{k=1}^K \pi_{k|j} = 1$ and $\pi_{k|j} \geq 0$, denote the transition probability between states j and k, that is, the probability to visit state k at time $t+1$ from state j at time t, $j,k = 1,\dots,K$ and $t = 1,\dots,T$. More concisely, we collect the initial and transition probabilities in the K-dimensional vector $\boldsymbol{\pi}$ and in the $K \times K$ matrix $\boldsymbol{\Pi}$, respectively.

Let $\boldsymbol{Y}_t = (Y_{t,1},\dots,Y_{t,d})$ denote a continuous d-dimensional dependent variable and $\boldsymbol{X}_t = (1, X_{t,2},\dots,X_{t,p})$ be a vector of p exogenous covariates, with the first element being the intercept, at time $t = 1,\dots,T$. Finally, let $\boldsymbol{\tau} = (\tau_1,\dots,\tau_d)$ denote a d-dimensional vector of fixed scalars with $\tau_j \in (0,1)$, $j = 1,\dots,d$.

As mentioned in the Introduction, our goal is to jointly model the univariate component-wise quantiles (expectiles) of the conditional distribution of the vector $\boldsymbol{Y}_t$ given $\boldsymbol{X}_t = \boldsymbol{x}_t$ and $S_t = k$, capturing the possible dependence structure between cryptocurrencies returns'. To this end, we construct a multivariate state-dependent distribution allowing for within-state correlation among the elements in $\boldsymbol{Y}_t$ by using a copula-based approach. As introduced by the seminal paper of Sklar (1959), the idea of a copula is to split a multivariate distribution into its univariate margins and the dependence structure, where the latter depends on the copula considered. Formally, a d-dimensional copula C is a d-dimensional distribution function on $[0,1]^d$ with standard uniform marginal distributions. Denoting with $F_{Y_{t,j}}(y_{t,j}|\boldsymbol{X}_t = \boldsymbol{x}_t, S_t = k)$, $j = 1,\dots,d$, the distribution functions of the marginals, the state-dependent multivariate distribution of $\boldsymbol{Y}_t$ given covariates and $S_t = k$, is defined by

$$F_{\boldsymbol{Y}_t}(\boldsymbol{y}_t|\boldsymbol{x}_t, S_t = k) = C(F_{Y_{t,1}}(y_{t,1}|\boldsymbol{x}_t, S_t = k),\dots,F_{Y_{t,d}}(y_{t,d}|\boldsymbol{x}_t, S_t = k);\boldsymbol{\eta}_k), \quad (3.5)$$

where $C(\cdot;\boldsymbol{\eta}_k)$ is a d-variate copula with time-varying parameter vector $\boldsymbol{\eta}_k$ that evolves over time according to the hidden process S_t and takes one of the values in the set $\{\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_K\}$. As a consequence of Sklar's theorem (Sklar 1959), when the marginal distribution functions are continuous and strictly increasing, the joint density $f_{\boldsymbol{Y}_t}(\boldsymbol{y}_t|\boldsymbol{x}_t, S_t = k)$ can be written as

$$f_{\boldsymbol{Y}_t}(\boldsymbol{y}_t|\boldsymbol{x}_t, S_t = k) = \prod_{j=1}^d f_{Y_{t,j}}(y_{t,j}|\boldsymbol{x}_t, S_t = k) \cdot c(u_1,\dots,u_d;\boldsymbol{\eta}_k) \quad (3.6)$$

where $u_j = F_{Y_{t,j}}(y_{t,j}|\boldsymbol{x}_t, S_t = k)$ and $c(\cdot;\boldsymbol{\eta}_k)$ is the copula density

$$c(\cdot;\boldsymbol{\eta}_k) = \frac{\partial^d C(\cdot;\boldsymbol{\eta}_k)}{\partial F_1 \cdots \partial F_d}. \quad (3.7)$$

Estimation of model parameters can be pursued using a ML approach. Specifically, to describe the conditional distribution of each response $Y_{t,j}$, $j = 1, \dots, d$, we use the density in (3.4) whose probability density function is now given by

$$f_{Y_{t,j}}(y_{t,j}|\boldsymbol{x}_t, S_t = k) = B_{l,\tau_j}(\sigma_{j,k}) \exp\left[-\omega_{l,\tau_j}\left(\frac{y_{t,j} - \mu_{t,j,k}}{\sigma_{j,k}}\right)\right], \tag{3.8}$$

where the location parameter $\mu_{t,j,k}$ is defined by the following linear model:

$$\mu_{t,j,k}(\tau_j) = \boldsymbol{X}_t \boldsymbol{\beta}_{j,k}(\tau_j), \quad j = 1, \dots, d, \tag{3.9}$$

with $\boldsymbol{\beta}_{j,k}(\tau_j)$ being the p-dimensional state-specific regression parameters that assumes one of the values $\{\boldsymbol{\beta}_{j,1}(\tau_j), \dots, \boldsymbol{\beta}_{j,K}(\tau_j)\}$ depending on the outcome of the Markov chain S_t. As described in Section 3.2, one of the advantages of using the distribution in (3.8) is that $\mu_{t,j,k}$ coincides with the τ_j-th conditional quantile of $Y_{t,j}$ when $l = 1$, which reduces to the AL distribution with $B_{l,\tau_j}(\sigma_{j,k}) = \frac{\tau_j(1-\tau_j)}{\sigma_{j,k}}$. Similarly, when $l = 2$, $\mu_{t,j,k}$ represents the τ_j-th conditional expectile of $Y_{t,j}$ and (3.8) corresponds to the AN distribution with $B_{l,\tau_j}(\sigma_{j,k}) = \frac{2\sqrt{\tau_j(1-\tau_j)}}{\sqrt{\pi\sigma^2_{j,k}}(\sqrt{\tau_j}+\sqrt{1-\tau_j})}$. These results hold true for all $k = 1, \dots, K$ and $j = 1, \dots, d$. In particular, (3.6)-(3.9) define the proposed Copula Quantile Hidden Markov Model (CQHMM) when $l = 1$ and the Copula Expectile Hidden Markov Model (CEHMM) when $l = 2$. With respect to the current literature, the proposed method reduces, respectively, to d separate linear quantile and expectile HMMs by Farcomeni (2012) and Foroni et al. (2023) for $l = 1$ and $l = 2$ when assuming conditional independence between the responses, i.e., under the independence copula. To allow for dependence, copulas have become a powerful tool in many applied settings and several families have been introduced over the years. Within the class of HMMs, copulas have been used by Lanchantin et al. (2011) for image analysis, by Härdle et al. (2015) in financial data and by Ötting et al. (2021) for investigating the dynamics of football matches. In the field of cryptocurrency market, copula dependence models have been analyzed in Hyun et al. (2019), Kim et al. (2020) and Gong & Huser (2022), who, in particular, manage to capture asymptotic dependence in lower and upper tails of cryptocurrency return distributions. In this work, to model the within state-dependence between the elements of $\boldsymbol{Y}_t$ we focus on the family of elliptical copulas which are derived from elliptically countered distributions, such as the Gaussian or the t distributions. Elliptical copulas have been frequently applied since the pioneering work of Embrechts et al. (2001), providing a convenient tool to characterize the dependence structure of financial returns through a correlation matrix with analytically tractable marginals. The Gaussian copula, due to its simplicity, is the most used copula in applied studies (see for example Joe 1997) and its distribution function can be written as

$$C^G(\boldsymbol{u}; \boldsymbol{\Omega}^\Phi) = \Phi_d(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d); \boldsymbol{\Omega}^\Phi) \tag{3.10}$$

where $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d$ is the generic vector of marginal distributions, Φ_d denotes the joint distribution function for the d-variate normal distribution with correlation matrix $\boldsymbol{\Omega}^\Phi$ and Φ^{-1} is the inverse distribution function of the univariate standard normal distribution. However, a number of papers have indicated the inability of the Gaussian copula to capture the dependence of extreme values, the so-called tail dependence, which is often observed in financial data (Cossin et al. 2010, Embrechts et al. 2011). To overcome this shortcoming, the t copula is symmetric like the Gaussian, but it incorporates the possibility of modeling tail dependence, which describes the behavior of copulas when the value of the marginal distribution function reaches its bounds of zero or one. The t copula is thus given by

$$C^t(\boldsymbol{u}; \boldsymbol{\Omega}^\Psi, \nu) = \Psi_d(\Psi^{-1}(u_1; \nu), \ldots, \Psi^{-1}(u_d, \nu); \boldsymbol{\Omega}^\Psi, \nu) \tag{3.11}$$

where Ψ_d denotes the joint distribution function of a d-variate t distribution with correlation matrix $\boldsymbol{\Omega}^\Psi$ and degrees of freedom $\nu > 2$. It is easy to see that as $\nu \to \infty$, the Gaussian copula in (3.10) may be thought of as a limiting case of the t copula.

3.3.1 Likelihood inference

In this section we consider a ML approach to make inference on model parameters. As is common for HMMs, and for latent variable models in general, we develop EM algorithms (Baum et al. 1970) to estimate the parameters of the methods proposed. As we will show in the following, the EM algorithms for fitting the proposed CQHMM and CEHMM have a similar structure and thus, so we present a general framework to avoid redundancies. To ease the notation, unless specified otherwise, hereinafter we also omit the vector $\boldsymbol{\tau}$ representing the quantile (expectile) indices, yet all model parameters are allowed to depend on it.

Let us denote with $\boldsymbol{\theta}_{\boldsymbol{\tau}} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K, \sigma_1, \ldots, \sigma_K, \boldsymbol{\pi}, \boldsymbol{\Pi}, \boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K)$ the vector of all model parameters. For the Gaussian copula, $\boldsymbol{\eta}_k$ comprises the elements of the state-specific correlation matrices $\boldsymbol{\Omega}_k^\Phi$, while for the t copula $\boldsymbol{\eta}_k = (\boldsymbol{\Omega}_k^\Psi, \nu_k)$, $k = 1, \ldots, K$.

Thus, for a given number of hidden states K, the EM algorithm runs on the complete log-likelihood function of the models introduced, which is defined as

$$\ell_c(\boldsymbol{\theta}_{\boldsymbol{\tau}}) = \sum_{k=1}^{K} \gamma_1(k) \log \pi_k + \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{j=1}^{K} \xi_t(j,k) \log \pi_{k|j} + \sum_{t=1}^{T} \sum_{k=1}^{K} \gamma_t(k) \log f_{\boldsymbol{Y}_t}(\boldsymbol{y}_t | \boldsymbol{x}_t, S_t = k), \tag{3.12}$$

where the joint density $f_{\boldsymbol{Y}_t}(\boldsymbol{y}_t | \boldsymbol{x}_t, S_t = k)$ is given in (3.6), $\gamma_t(k)$ denotes a dummy variable equal to 1 if the latent process is in state k at occasion t and 0 otherwise, and $\xi_t(j,k)$ is a dummy variable equal to 1 if the process is in state j in $t-1$ and in state k at time t and 0 otherwise.

To estimate $\boldsymbol{\theta}_\tau$, the algorithm iterates between two steps, the E- and M-steps, until convergence, as outlined below.

E-step:

In the E-step, at the generic $(h+1)$-th iteration, the unobservable indicator variables $\gamma_t(k)$ and $\xi_t(j,k)$ in (3.12) are replaced by their conditional expectations given the observed data and the current parameter estimates $\boldsymbol{\theta}_{\boldsymbol{\tau}}^{(h)}$. To compute such quantities we require the calculation of the probability of being in state k at time t given the observed sequence

$$\gamma_t^{(h)}(k) = P_{\boldsymbol{\theta}_{\boldsymbol{\tau}}^{(h)}}(S_t = k|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T) = \frac{a_{t,k} b_{t,k}}{\sum_{k'=1}^{K} a_{t,k'} b_{t,k'}} \tag{3.13}$$

and the probability that at time $t-1$ the process is in state j and then in state k at time t, given the observed sequence

$$\xi_t^{(h)}(j,k) = P_{\boldsymbol{\theta}_{\boldsymbol{\tau}}^{(h)}}(S_{t-1} = j, S_t = k|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T) = \frac{a_{t-1,k}\pi_{k|j} f(\cdot) b_{t,k}}{\sum_{k'=1}^{K} a_{T,k'}}, \tag{3.14}$$

where $f(\cdot)$ is the AL density for the CQHMM or the AN distribution for the CEHMM in (3.8), and $a_{t,k'}$ and $b_{t,k'}$ represent the forward and backward probabilities of each model. Such quantities can be efficiently obtained using the well-known Forward-Backward algorithm; see Baum et al. (1970) and Welch (2003). Then, we use these to calculate the conditional expectation of the complete log-likelihood function in (3.12) given the observed data and the current estimates:

$$\begin{aligned} Q(\boldsymbol{\theta}_{\boldsymbol{\tau}}|\boldsymbol{\theta}_{\boldsymbol{\tau}}^{(h)}) = &\sum_{k=1}^{K} \gamma_1^{(h)}(k) \log \pi_k + \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{j=1}^{K} \xi_t^{(h)}(j,k) \log \pi_{k|j} + \\ &\sum_{t=1}^{T}\sum_{k=1}^{K} \gamma_t^{(h)}(k) \log f_{\boldsymbol{Y}_t}(\boldsymbol{y}_t|\boldsymbol{x}_t, S_t = k). \end{aligned} \tag{3.15}$$

M-step:

In the M-step we maximize $Q(\boldsymbol{\theta}_{\boldsymbol{\tau}}|\boldsymbol{\theta}_{\boldsymbol{\tau}}^{(h)})$ in (3.15) with respect to $\boldsymbol{\theta}_{\boldsymbol{\tau}}$ to obtain the update parameter estimates $\boldsymbol{\theta}_{\boldsymbol{\tau}}^{(h+1)}$. Formally, the initial probabilities π_k and transition probabilities $\pi_{k|j}$ are updated using:

$$\pi_k^{(h+1)} = \gamma_1^{(h)}(k), \quad k = 1, \ldots, K \tag{3.16}$$

and

$$\pi_{k|j}^{(h+1)} = \frac{\sum_{t=1}^{T} \xi_t^{(h)}(j,k)}{\sum_{t=1}^{T}\sum_{k=1}^{K} \xi_t^{(h)}(j,k)}, \quad j,k = 1, \ldots, K. \tag{3.17}$$

To reduce the computational difficulty of the algorithm, we then update the state-dependent regression and copula parameters by adopting the Inference Functions

for Margins (IFM) method proposed in Joe & Xu (1996). This entails a two-stage estimation procedure that first estimates the regression parameters of each component of $\boldsymbol{Y}_t$, $(\boldsymbol{\beta}_{j,k}, \sigma_{j,k})$, and it then obtains the parameters of the copula function, $\boldsymbol{\eta}_k$. To implement this step, we write the last term in (3.15) as follows:

$$\begin{aligned}
&\sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_t^{(h)}(k)\log f_{\boldsymbol{Y}_t}(\boldsymbol{y}_t|\boldsymbol{x}_t, S_t=k) = \\
&\sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_t^{(h)}(k)\Big(\log c(F_{Y_{t,1}}(y_{t,1}|\boldsymbol{x}_t, S_t=k), \ldots, F_{Y_{t,d}}(y_{t,d}|\boldsymbol{x}_t, S_t=k); \boldsymbol{\eta}_k)\Big) + \\
&\sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_t^{(h)}(k)\Big(\sum_{j=1}^{d}\log f_{Y_{t,j}}(y_{t,j}|\boldsymbol{x}_t, S_t=k)\Big),
\end{aligned} \tag{3.18}$$

where $f_{Y_{t,j}}(y_{t,j}|\boldsymbol{x}_t, S_t=k)$ and $F_{Y_{t,j}}(y_{t,j}|\boldsymbol{x}_t, S_t=k)$ are the univariate conditional density and distribution function for each component j and each state k, respectively. In the CQHMM, these correspond to the density and distribution functions of the AL distribution, whereas, in the CEHMM, they represent the density and distribution functions of the AN distribution.

In the first step, we update the regression and scale parameters of each univariate conditional distribution by maximizing the second term in (3.18). Specifically, the estimate of the regression parameters are updated as follows

$$\boldsymbol{\beta}_{j,k}^{(h+1)} = \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{T}\gamma_t^{(h)}(k)\omega_{l,\tau_j}(y_{t,j} - \boldsymbol{x}_t'\boldsymbol{\beta}). \tag{3.19}$$

In the case of quantiles, a solution to (3.19) can be obtain by fitting a linear quantile regression with weights $\gamma_t^{(h)}(k)$ for $l=1$. As for expectiles, due to the differentiability of the loss function $\omega_{2,\tau_j}(\cdot)$, $\boldsymbol{\beta}_{j,k}^{(h+1)}$ can be efficiently computed using IRLS for cross-sectional data with appropriate weights. Similarly, the scale parameters of each marginal distribution can be obtained using the following M-step update formulas for the CQHMM

$$\sigma_{j,k}^{(h+1)} = \frac{1}{\sum_{t=1}^{T}\gamma_t^{(h)}(k)}\sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_t^{(h)}(k)\omega_{1,\tau_j}(y_{t,j} - \boldsymbol{x}_t'\boldsymbol{\beta}_{k,j}^{(h+1)}) \tag{3.20}$$

and for the CEHMM

$$\sigma_{j,k}^{2(h+1)} = \frac{2}{\sum_{t=1}^{T}\gamma_t^{(h)}(k)}\sum_{t=1}^{T}\gamma_t^{(h)}(k)|\tau_j - \mathbb{I}(y_{t,j} < \boldsymbol{x}_t'\boldsymbol{\beta}_{j,k}^{(h+1)})|(y_{t,j} - \boldsymbol{x}_t'\boldsymbol{\beta}_{j,k}^{(h+1)})^2. \tag{3.21}$$

In the second step, given the estimates of $\boldsymbol{\beta}_{j,k}^{(h+1)}$ and $\sigma_{j,k}^{(h+1)}$ for all components $j = 1, \ldots, d$ and for all hidden states, we compute the copula parameters for

$k = 1, \ldots, K$ as follows:

$$\boldsymbol{\eta}_k^{(h+1)} = \arg\min_{\boldsymbol{\eta}} \sum_{t=1}^{T} \gamma_t^{(h)}(k) \log c(F_{Y_{t,1}}^{(h+1)}(y_{t,1}|\boldsymbol{x}_t, S_t = k), \ldots, F_{Y_{t,d}}^{(h+1)}(y_{t,d}|\boldsymbol{x}_t, S_t = k); \boldsymbol{\eta}), \tag{3.22}$$

where $F_{Y_{t,j}}^{(h+1)}(y_{t,j}|\boldsymbol{x}_t, S_t = k)$ denotes the univariate conditional distribution of the j-th response evaluated at $(\boldsymbol{\beta}_{j,k}^{(h+1)}, \sigma_{j,k}^{(h+1)})$.

For the Gaussian copula, the update of the state-dependent correlation matrix $\boldsymbol{\Omega}_k^{\Phi}$ is obtained by maximizing (3.22), which has a closed form expression. This coincides with the weighted sample covariance matrix of the pseudo-observations $u_j = F_{Y_{t,j}}^{(h+1)}(y_{t,j}|\boldsymbol{x}_t, S_t = k)$, $j = 1, \ldots, d$, with weights $\gamma_t^{(h)}(k)$. In the case of the t copula, in order to reduce the dimensionality of the optimization problem in (3.22), we update the correlation matrix $\boldsymbol{\Omega}_k^{\Psi}$ for a fixed value of ν_k. This is equivalent to estimating the scale matrix of a multivariate t distribution with zero mean vector, which can be done in closed form exploiting the hierarchical representation of the t distribution (Liu 1997). Once we updated $\boldsymbol{\Omega}_k^{\Psi}$, the objective function in (3.22) is only a function of ν_k and, thus, the related computing is fast.

The E- and M- steps are alternated until convergence, that is when the observed likelihood between two consecutive iterations is smaller than a predetermined threshold. In this paper, we set this threshold criterion equal to 10^{-6}.
Following Maruotti et al. (2021) and Merlo et al. (2022), for fixed $\boldsymbol{\tau}$ and K we initialize the EM algorithm by providing the initial states partition, $\{S_t^{(0)}\}_{t=1}^T$, according to a Multinomial distribution with probabilities $1/K$. From the generated partition, the elements of $\boldsymbol{\Pi}^{(0)}$ are computed as proportions of transition, while we obtain $\boldsymbol{\beta}_{j,k}^{(0)}$ and $\sigma_{j,k}^{(0)}$ by fitting univariate mean and median regressions on the observations within state k for the CQHMM and CEHMM, respectively. The state-dependent correlation matrices of the copula are set equal to the empirical correlation matrices computed on observations in the k-th state, while the initial value for the degrees of freedom of the t copula is $\nu^{(0)} = \nu_k^{(0)} = 5$ for all $k = 1, \ldots, K$. To deal with the possibility of multiple roots of the likelihood equation and better explore the parameter space, we fit the proposed CQHMM and CEHMM using a multiple random starts strategy with different starting partitions and retain the solution corresponding to the maximum likelihood value.

Once we computed the ML estimates of the model parameters, to estimate the standard errors we employ the parametric bootstrap scheme of Visser et al. (2000).

3.4 Simulation Studies

In this section we evaluate the performance of the proposed CQHMM and CEHMM via simulated data under different scenarios. In particular, we assess the ability of recovering the true regression parameter values and the clustering accuracy. We consider a two-dimensional response variable ($d = 2$), two sample sizes ($T = 500$, $T = 1000$) and one explanatory variable $X_t \sim \mathcal{N}(0, 1)$. Observations are drawn from a bivariate two-states HMM using the following data generating process for $t = 1, \ldots, T$,

$$\boldsymbol{Y}_t = \boldsymbol{X}_t \boldsymbol{\beta}_k + \boldsymbol{\epsilon}_{t,k}, \quad S_t = k \tag{3.23}$$

where $\boldsymbol{X}_t = (1, X_t)'$ and the true values of the regression parameters are

$$\boldsymbol{\beta}_1 = \begin{pmatrix} -2 & 3 \\ 1 & -2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\beta}_2 = \begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix}.$$

We consider three distributions for the error term in (3.23). In the first case, we generate $\boldsymbol{\epsilon}_{t,k}$ from a multivariate Gaussian distribution with zero mean vector and variance-covariance matrix equal to $\boldsymbol{\Omega}_k$. In the second one, $\boldsymbol{\epsilon}_{t,k}$ is generated from a multivariate t distribution with 5 degrees of freedom, zero mean vector and scale matrix $\boldsymbol{\Omega}_k$. In the third scenario, $\boldsymbol{\epsilon}_{t,k}$ follows a multivariate skew-t distribution with 5 degrees of freedom, skewness parameters $\boldsymbol{\alpha} = (-2, 2)$ and scale matrix $\boldsymbol{\Omega}_k$. The state-specific covariance matrices for the two states are equal to:

$$\boldsymbol{\Omega}_1 = \begin{pmatrix} 1 & 0.2 \\ 0.2 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Omega}_2 = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}.$$

Finally, the matrix of transition probabilities is $\boldsymbol{\Pi} = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}$.

We fit the proposed CQHMM and CEHMM at three τ levels, i.e., $\tau = \{0.1, 0.5, 0.9\}$ with, $\tau = \tau_j$, $j = 1, 2$, by assuming the Gaussian and t copulas for each scenario considered. In order to assess the validity of the models, we compute the absolute bias and standard errors associated to the state-specific regression coefficients $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, averaged over 100 Monte Carlo replications, for each combination of sample size, error distribution and copula function.

In Tables 3.1, 3.2 and 3.3 we report the estimation results for the CQHMM regression parameters for each combination of copulas and error distributions. In this case we observe that the precision of the estimates is higher at the center of the distribution rather than on the tails, especially for the fat-tailed t and skew t distributions, mainly due to the reduced number of observations at extreme quantile levels. Some improvements can be observed when the parameters of the skew t case are estimated with the copula t rather than the Gaussian. Moving onto the

estimation results of the CEHMM (Tables 3.4, 3.5, 3.6), we observe that the recovery performance of the data generation parameters is always very satisfactory, no matter the sample size, the copula or the generating process considered. Even when we generate from a skew-t distribution we obtain good results on extreme expectiles, despite having a reduced amount of observations at the tails of the distribution. When we increase the sample size ($T = 1000$) we observe that standard deviation tends to decrease, as expected. Overall, the expectile model achieves lower bias and standard errors with respect to the quantile model, especially at the tails of non-Gaussian distributions.

In order to evaluate the ability in recovering the true states partition we consider the Adjusted Rand Index (ARI) of Hubert & Arabie (1985). The state partition provided by the fitted models is obtained by taking the maximum a posteriori probability, $\max_k \gamma_t(k)$, for every $t = 1, \ldots, T$. In Figure 3.1 is reported the box-plot of ARI for the posterior probabilities for all the settings considered for $T = 500$ and $T = 1000$ sample sizes. For the CQHMM, we obtain very high performances in estimating the true state partition. For both copula settings, we obtain a better clustering performance for the model with Gaussian and t errors with respect to the skew-t case. We also highlight that the cluster recovering ability depends on the specific quantile level, being the values slightly higher at the median ($\tau = 0.50$) than at the tails. Finally, when increasing the sample size to $T = 1000$, results slightly improve, reporting a lower variability for both error distributions.

Moving onto the CEHMM, we observe that in every setting we obtain very high performances in estimating the true state partition. For both copula settings, we obtain a better clustering performance for the model with Gaussian and t errors with respect to the skew-t case. Compare to the CQHMM, the goodness of the clustering obtained does not depend on the specific expectile level, being the values very alike along the centre and the tails of the distribution, probably due to the the major efficiency of the algorithm in the expectile case. Finally, when increasing the sample size to $T = 1000$, results slightly improve reporting a lower variability for all error distributions. As in the quantile framework, the proposed CEHMM is able to recover the true values of the parameters and the true state partition in a highly satisfactory way in all the distributions and copula settings examined.

3.5 Empirical Application

In this section we apply the proposed CQHMM and CEHMM to analyze daily returns of the five major cryptocurrencies as functions of global market indices. The goal of the analysis is to investigate their relationship with leading market indices and describe their dependence structure at different volatility states.

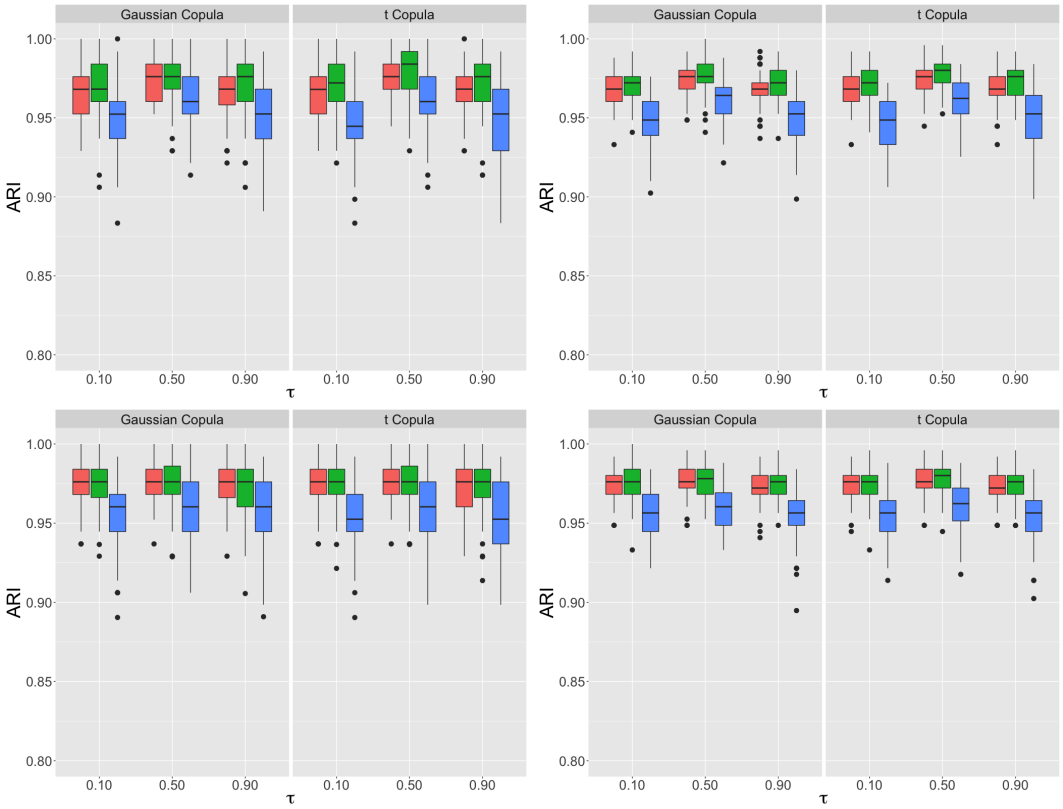

Figure 3.1. Box-plots of ARI for the posterior probabilities for CQHMM (first row) and CEHMM (second row) under the Gaussian (red), Student's t (green) and skew-t (blue) distributed errors with Gaussian and t copula, and sample sizes $T = 500$ (left column) and $T = 1000$ (right column).

CQHMM	τ	0.10	0.50	0.90		τ	0.10	0.50	0.90
Gaussian Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)	t Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)
Panel A: T=500					Panel A: T=500				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	-0.002 (0.076)	0.004 (0.058)	0.013 (0.079)	j=1	$\beta_{0,1} = -2$	-0.001 (0.083)	0.005 (0.057)	0.012 (0.082)
	$\beta_{1,1} = 1$	-0.013 (0.130)	-0.008 (0.078)	0.002 (0.123)		$\beta_{1,1} = 1$	-0.014 (0.133)	-0.009 (0.077)	-0.001 (0.123)
j=2	$\beta_{0,1} = 3$	-0.013 (0.088)	0.000 (0.055)	0.009 (0.086)	j=2	$\beta_{0,1} = 3$	-0.057 (0.455)	0.000 (0.055)	0.009 (0.087)
	$\beta_{1,1} = -2$	0.025 (0.108)	0.008 (0.079)	-0.009 (0.105)		$\beta_{1,1} = -2$	0.054 (0.300)	0.008 (0.079)	-0.006 (0.106)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	-0.008 (0.076)	0.001 (0.056)	0.000 (0.077)	j=1	$\beta_{0,2} = 3$	-0.020 (0.123)	0.000 (0.056)	-0.001 (0.078)
	$\beta_{1,2} = -2$	-0.002 (0.127)	0.003 (0.087)	0.017 (0.107)		$\beta_{1,2} = -2$	0.006 (0.176)	0.004 (0.088)	0.017 (0.106)
j=2	$\beta_{0,2} = -2$	0.003 (0.073)	-0.005 (0.050)	-0.007 (0.093)	j=2	$\beta_{0,2} = -2$	-0.005 (0.096)	-0.005 (0.051)	-0.005 (0.094)
	$\beta_{1,2} = 1$	-0.011 (0.110)	0.013 (0.082)	0.008 (0.123)		$\beta_{1,2} = 1$	0.015 (0.300)	0.014 (0.083)	0.010 (0.122)
Panel B: T=1000					Panel B: T=1000				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	-0.005 (0.064)	-0.005 (0.042)	-0.017 (0.054)	j=1	$\beta_{0,1} = -2$	-0.005 (0.064)	-0.004 (0.042)	-0.014 (0.054)
	$\beta_{1,1} = 1$	-0.013 (0.080)	-0.015 (0.045)	-0.002 (0.071)		$\beta_{1,1} = 1$	-0.013 (0.081)	-0.015 (0.045)	-0.002 (0.070)
j=2	$\beta_{0,1} = 3$	-0.008 (0.053)	-0.002 (0.038)	0.007 (0.066)	j=2	$\beta_{0,1} = 3$	-0.009 (0.052)	-0.001 (0.038)	0.007 (0.066)
	$\beta_{1,1} = -2$	0.001 (0.075)	0.006 (0.056)	0.013 (0.076)		$\beta_{1,1} = -2$	0.002 (0.073)	0.007 (0.056)	0.015 (0.076)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	0.007 (0.072)	0.003 (0.042)	0.018 (0.063)	j=1	$\beta_{0,2} = 3$	0.002 (0.071)	0.003 (0.042)	0.017 (0.064)
	$\beta_{1,2} = -2$	-0.001 (0.081)	0.007 (0.056)	0.005 (0.080)		$\beta_{1,2} = -2$	-0.005 (0.081)	0.006 (0.057)	0.004 (0.080)
j=2	$\beta_{0,2} = -2$	-0.006 (0.050)	0.000 (0.039)	-0.002 (0.062)	j=2	$\beta_{0,2} = -2$	-0.006 (0.050)	0.001 (0.040)	0.000 (0.062)
	$\beta_{1,2} = 1$	-0.003 (0.075)	0.000 (0.053)	-0.003 (0.077)		$\beta_{1,2} = 1$	-0.004 (0.074)	0.001 (0.053)	0.000 (0.077)

Table 3.1. Bias and standard error values of the state-regression parameter estimates for CQHMM with Gaussian distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B).

CQHMM	τ	0.10	0.50	0.90		τ	0.10	0.50	0.90
Gaussian Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)	t Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)
Panel A: T=500					Panel A: T=500				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	-0.066 (0.405)	-0.003 (0.069)	0.283 (1.004)	j=1	$\beta_{0,1} = -2$	-0.116 (0.715)	-0.003 (0.069)	0.360 (1.186)
	$\beta_{1,1} = 1$	-0.026 (0.239)	0.011 (0.094)	-0.202 (1.152)		$\beta_{1,1} = 1$	-0.097 (0.799)	0.011 (0.094)	-0.248 (0.869)
j=2	$\beta_{0,1} = 3$	-0.297 (1.155)	-0.001 (0.067)	-0.145 (0.884)	j=2	$\beta_{0,1} = 3$	-0.455 (1.438)	0.000 (0.068)	-0.094 (0.585)
	$\beta_{1,1} = -2$	0.176 (1.021)	-0.019 (0.079)	-0.108 (1.279)		$\beta_{1,1} = -2$	0.420 (1.331)	-0.018 (0.080)	0.012 (0.343)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	-0.343 (1.153)	0.010 (0.070)	0.013 (0.215)	j=1	$\beta_{0,2} = 3$	-0.497 (1.331)	0.011 (0.071)	0.071 (0.343)
	$\beta_{1,2} = -2$	0.185 (1.001)	0.002 (0.079)	0.005 (0.270)		$\beta_{1,2} = -2$	0.405 (1.268)	0.004 (0.079)	0.018 (0.191)
j=2	$\beta_{0,2} = -2$	0.042 (0.701)	-0.003 (0.066)	0.537 (1.553)	j=2	$\beta_{0,2} = -2$	0.110 (0.822)	-0.002 (0.065)	0.446 (1.464)
	$\beta_{1,2} = 1$	-0.001 (0.549)	0.004 (0.082)	-0.334 (1.034)		$\beta_{1,2} = 1$	-0.101 (0.584)	0.004 (0.082)	-0.199 (0.781)
Panel B: T=1000					Panel B: T=1000				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	-0.171 (0.727)	-0.002 (0.047)	0.270 (1.083)	j=1	$\beta_{0,1} = -2$	-0.133 (0.482)	-0.001 (0.047)	0.541 (1.396)
	$\beta_{1,1} = 1$	0.068 (1.074)	-0.002 (0.065)	-0.330 (1.309)		$\beta_{1,1} = 1$	-0.065 (0.268)	-0.002 (0.066)	-0.479 (1.225)
j=2	$\beta_{0,1} = 3$	-0.563 (1.560)	0.006 (0.043)	-0.006 (0.718)	j=2	$\beta_{0,1} = 3$	-1.006 (1.850)	0.006 (0.044)	0.022 (0.274)
	$\beta_{1,1} = -2$	0.576 (1.705)	-0.007 (0.065)	0.046 (0.963)		$\beta_{1,1} = -2$	0.716 (1.314)	-0.008 (0.066)	-0.015 (0.332)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	-0.412 (1.377)	0.002 (0.044)	0.044 (0.270)	j=1	$\beta_{0,2} = 3$	-0.936 (1.711)	0.001 (0.044)	0.083 (0.292)
	$\beta_{1,2} = -2$	0.554 (1.943)	0.002 (0.061)	0.004 (0.312)		$\beta_{1,2} = -2$	0.662 (1.228)	0.002 (0.061)	0.042 (0.170)
j=2	$\beta_{0,2} = -2$	0.184 (1.299)	-0.009 (0.041)	0.421 (1.327)	j=2	$\beta_{0,2} = -2$	0.002 (0.243)	-0.006 (0.041)	0.616 (1.500)
	$\beta_{1,2} = 1$	0.059 (2.221)	-0.004 (0.061)	-0.269 (1.005)		$\beta_{1,2} = 1$	-0.049 (0.151)	-0.002 (0.062)	-0.370 (1.003)

Table 3.2. Bias and standard error values of the state-regression parameter estimates for CQHMM with Student's t distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B).

CQHMM	τ	0.10	0.50	0.90	τ	0.10	0.50	0.90
Gaussian Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)	Student's t Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)
Panel A: T=500					Panel A: T=500				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	-0.117 (0.220)	-0.100 (0.057)	-1.047 (10.412)	j=1	$\beta_{0,1} = -2$	-0.117 (0.277)	-0.099 (0.059)	-0.038 (0.786)
	$\beta_{1,1} = 1$	-0.041 (0.185)	0.010 (0.073)	1.189 (13.284)		$\beta_{1,1} = 1$	-0.027 (0.239)	0.011 (0.074)	-0.087 (0.564)
j=2	$\beta_{0,1} = 3$	-0.253 (1.524)	0.099 (0.060)	0.472 (4.430)	j=2	$\beta_{0,1} = 3$	-0.314 (1.557)	0.099 (0.060)	0.035 (0.253)
	$\beta_{1,1} = -2$	0.206 (0.777)	0.007 (0.084)	-0.606 (6.110)		$\beta_{1,1} = -2$	0.258 (0.872)	0.008 (0.083)	-0.023 (0.347)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	-0.266 (1.189)	0.122 (0.070)	0.215 (0.281)	j=1	$\beta_{0,2} = 3$	-0.305 (1.230)	0.119 (0.068)	0.186 (0.148)
	$\beta_{1,2} = -2$	0.237 (0.853)	0.000 (0.067)	0.009 (0.152)		$\beta_{1,2} = -2$	0.291 (1.072)	-0.002 (0.066)	0.016 (0.159)
j=2	$\beta_{0,2} = -2$	0.056 (1.139)	-0.111 (0.065)	0.291 (1.418)	j=2	$\beta_{0,2} = -2$	-0.01 (0.919)	-0.110 (0.065)	0.349 (1.424)
	$\beta_{1,2} = 1$	-0.119 (0.616)	-0.001 (0.08)	-0.189 (0.683)		$\beta_{1,2} = 1$	-0.111 (0.566)	0.000 (0.081)	-0.212 (0.733)
Panel B: T=1000					Panel B: T=1000				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	-0.152 (0.487)	-0.101 (0.044)	-0.018 (1.094)	j=1	$\beta_{0,1} = -2$	-0.141 (0.301)	-0.101 (0.045)	0.168 (1.240)
	$\beta_{1,1} = 1$	0.070 (0.686)	-0.003 (0.056)	-0.266 (1.217)		$\beta_{1,1} = 1$	-0.009 (0.143)	-0.002 (0.056)	-0.244 (0.795)
j=2	$\beta_{0,1} = 3$	-0.274 (1.476)	0.103 (0.044)	0.167 (0.737)	j=2	$\beta_{0,1} = 3$	-0.318 (1.568)	0.103 (0.044)	0.051 (0.184)
	$\beta_{1,1} = -2$	0.285 (0.988)	-0.011 (0.048)	0.099 (0.909)		$\beta_{1,1} = -2$	0.298 (0.904)	-0.011 (0.048)	0.001 (0.126)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	-0.308 (1.239)	0.112 (0.046)	0.238 (0.410)	j=1	$\beta_{0,2} = 3$	-0.344 (1.281)	0.111 (0.045)	0.212 (0.221)
	$\beta_{1,2} = -2$	0.348 (1.138)	0.006 (0.060)	0.012 (0.275)		$\beta_{1,2} = -2$	0.284 (0.875)	0.004 (0.059)	0.092 (0.691)
j=2	$\beta_{0,2} = -2$	-0.099 (0.293)	-0.119 (0.047)	0.205 (1.117)	j=2	$\beta_{0,2} = -2$	-0.112 (0.589)	-0.118 (0.046)	0.385 (1.476)
	$\beta_{1,2} = 1$	-0.056 (0.671)	0.000 (0.065)	-0.244 (1.071)		$\beta_{1,2} = 1$	-0.056 (0.308)	0.001 (0.064)	-0.268 (0.932)

Table 3.3. Bias and standard error values of the state-regression parameter estimates for CQHMM with skew t distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B).

CEHMM	τ	0.10	0.50	0.90	τ	0.10	0.50	0.90
Gaussian Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)	Student's t Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)
Panel A: T=500					Panel A: T=500				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	0.000 (0.075)	0.000 (0.060)	-0.009 (0.081)	j=1	$\beta_{0,1} = -2$	0.003 (0.075)	0.001 (0.060)	-0.008 (0.081)
	$\beta_{1,1} = 1$	-0.009 (0.093)	-0.009 (0.070)	-0.007 (0.087)		$\beta_{1,1} = 1$	-0.009 (0.092)	-0.009 (0.070)	-0.007 (0.087)
j=2	$\beta_{0,1} = 3$	0.001 (0.072)	0.000 (0.057)	-0.002 (0.072)	j=2	$\beta_{0,1} = 3$	0.004 (0.072)	0.001 (0.057)	-0.001 (0.071)
	$\beta_{1,1} = -2$	0.012 (0.074)	0.005 (0.058)	-0.005 (0.073)		$\beta_{1,1} = -2$	0.012 (0.074)	0.005 (0.058)	-0.004 (0.073)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	-0.004 (0.067)	-0.005 (0.052)	-0.013 (0.072)	j=1	$\beta_{0,2} = 3$	-0.005 (0.067)	-0.005 (0.052)	-0.011 (0.071)
	$\beta_{1,2} = -2$	-0.004 (0.084)	0.003 (0.071)	0.008 (0.084)		$\beta_{1,2} = -2$	-0.005 (0.084)	0.003 (0.071)	0.008 (0.084)
j=2	$\beta_{0,2} = -2$	0.008 (0.063)	-0.003 (0.053)	-0.012 (0.079)	j=2	$\beta_{0,2} = -2$	0.007 (0.063)	-0.003 (0.053)	-0.012 (0.079)
	$\beta_{1,2} = 1$	0.001 (0.083)	0.008 (0.067)	0.012 (0.085)		$\beta_{1,2} = 1$	0.001 (0.083)	0.008 (0.067)	0.012 (0.085)
Panel B: T=1000					Panel B: T=1000				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	-0.004 (0.057)	-0.008 (0.040)	-0.018 (0.055)	j=1	$\beta_{0,1} = -2$	-0.001 (0.057)	-0.005 (0.040)	-0.014 (0.055)
	$\beta_{1,1} = 1$	-0.007 (0.059)	-0.010 (0.042)	-0.009 (0.050)		$\beta_{1,1} = 1$	-0.007 (0.059)	-0.010 (0.042)	-0.009 (0.050)
j=2	$\beta_{0,1} = 3$	0.010 (0.049)	0.006 (0.040)	0.002 (0.054)	j=2	$\beta_{0,1} = 3$	0.010 (0.049)	0.006 (0.040)	0.004 (0.054)
	$\beta_{1,1} = -2$	0.006 (0.049)	0.006 (0.043)	0.007 (0.053)		$\beta_{1,1} = -2$	0.006 (0.049)	0.006 (0.043)	0.007 (0.053)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	0.004 (0.059)	0.005 (0.048)	0.001 (0.061)	j=1	$\beta_{0,2} = 3$	-0.002 (0.059)	0.001 (0.048)	-0.002 (0.061)
	$\beta_{1,2} = -2$	-0.002 (0.060)	0.004 (0.048)	0.007 (0.056)		$\beta_{1,2} = -2$	-0.003 (0.060)	0.003 (0.048)	0.007 (0.056)
j=2	$\beta_{0,2} = -2$	0.008 (0.046)	0.004 (0.041)	-0.007 (0.059)	j=2	$\beta_{0,2} = -2$	0.008 (0.047)	0.003 (0.041)	-0.008 (0.059)
	$\beta_{1,2} = 1$	-0.007 (0.052)	-0.004 (0.041)	-0.003 (0.051)		$\beta_{1,2} = 1$	-0.007 (0.052)	-0.004 (0.041)	-0.003 (0.051)

Table 3.4. Bias and standard error values of the state-regression parameter estimates for CEHMM with Gaussian distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B).

CEHMM	τ	0.10	0.50	0.90		τ	0.10	0.50	0.90
Gaussian Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)	Student's t Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)
Panel A: T=500					Panel A: T=500				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	0.002 (0.138)	-0.002 (0.091)	-0.014 (0.152)	j=1	$\beta_{0,1} = -2$	0.003 (0.137)	0.001 (0.091)	-0.004 (0.156)
	$\beta_{1,1} = 1$	0.000 (0.116)	0.010 (0.082)	0.013 (0.141)		$\beta_{1,1} = 1$	-0.005 (0.119)	0.009 (0.084)	0.018 (0.144)
j=2	$\beta_{0,1} = 3$	0.023 (0.154)	0.006 (0.094)	-0.009 (0.143)	j=2	$\beta_{0,1} = 3$	0.023 (0.155)	0.007 (0.094)	-0.002 (0.146)
	$\beta_{1,1} = -2$	-0.007 (0.122)	-0.009 (0.079)	-0.001 (0.125)		$\beta_{1,1} = -2$	-0.018 (0.130)	-0.010 (0.080)	0.002 (0.125)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	0.014 (0.124)	0.007 (0.075)	0.000 (0.122)	j=1	$\beta_{0,2} = 3$	0.012 (0.129)	0.007 (0.075)	-0.001 (0.121)
	$\beta_{1,2} = -2$	-0.003 (0.128)	-0.002 (0.076)	0.003 (0.141)		$\beta_{1,2} = -2$	0.001 (0.131)	-0.002 (0.077)	-0.001 (0.141)
j=2	$\beta_{0,2} = -2$	-0.007 (0.146)	-0.002 (0.084)	0.001 (0.155)	j=2	$\beta_{0,2} = -2$	-0.002 (0.139)	-0.001 (0.084)	0.003 (0.146)
	$\beta_{1,2} = 1$	0.007 (0.127)	0.005 (0.072)	0.007 (0.118)		$\beta_{1,2} = 1$	0.015 (0.123)	0.007 (0.072)	0.007 (0.112)
Panel B: T=1000					Panel B: T=1000				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	0.006 (0.091)	-0.006 (0.060)	-0.033 (0.102)	j=1	$\beta_{0,1} = -2$	0.000 (0.094)	-0.008 (0.060)	-0.024 (0.117)
	$\beta_{1,1} = 1$	0.004 (0.093)	-0.004 (0.058)	-0.019 (0.098)		$\beta_{1,1} = 1$	0.000 (0.093)	-0.003 (0.058)	-0.012 (0.102)
j=2	$\beta_{0,1} = 3$	0.035 (0.093)	0.021 (0.058)	0.020 (0.095)	j=2	$\beta_{0,1} = 3$	0.029 (0.099)	0.016 (0.058)	0.013 (0.097)
	$\beta_{1,1} = -2$	-0.001 (0.089)	-0.009 (0.060)	-0.012 (0.100)		$\beta_{1,1} = -2$	-0.005 (0.094)	-0.011 (0.060)	-0.014 (0.095)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	0.009 (0.099)	0.005 (0.057)	0.016 (0.095)	j=1	$\beta_{0,2} = 3$	0.003 (0.100)	0.001 (0.057)	0.007 (0.098)
	$\beta_{1,2} = -2$	-0.001 (0.090)	0.002 (0.059)	0.010 (0.102)		$\beta_{1,2} = -2$	0.000 (0.086)	0.001 (0.060)	0.007 (0.098)
j=2	$\beta_{0,2} = -2$	0.010 (0.102)	0.003 (0.056)	0.012 (0.093)	j=2	$\beta_{0,2} = -2$	0.012 (0.103)	0.003 (0.057)	0.015 (0.100)
	$\beta_{1,2} = 1$	-0.002 (0.098)	0.005 (0.059)	0.013 (0.090)		$\beta_{1,2} = 1$	0.000 (0.099)	0.007 (0.059)	0.017 (0.091)

Table 3.5. Bias and standard error values of the state-regression parameter estimates for CEHMM with Student's t distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B).

CEHMM	τ	0.10	0.50	0.90		τ	0.10	0.50	0.90
Gaussian Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)	Student's t Copula		Bias (Std.Err)	Bias (Std.Err)	Bias (Std.Err)
Panel A: T=500					Panel A: T=500				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	0.012 (0.152)	0.008 (0.080)	0.008 (0.112)	j=1	$\beta_{0,1} = -2$	0.016 (0.151)	0.009 (0.081)	0.011 (0.109)
	$\beta_{1,1} = 1$	-0.006 (0.130)	0.006 (0.069)	0.015 (0.102)		$\beta_{1,1} = 1$	-0.003 (0.128)	0.007 (0.071)	0.017 (0.101)
j=2	$\beta_{0,1} = 3$	-0.001 (0.106)	-0.007 (0.083)	-0.013 (0.153)	j=2	$\beta_{0,1} = 3$	-0.003 (0.108)	-0.005 (0.082)	-0.015 (0.149)
	$\beta_{1,1} = -2$	-0.003 (0.099)	0.002 (0.082)	0.005 (0.155)		$\beta_{1,1} = -2$	-0.006 (0.104)	0.002 (0.083)	0.003 (0.148)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	0.012 (0.132)	0.010 (0.070)	-0.002 (0.102)	j=1	$\beta_{0,2} = 3$	0.010 (0.141)	0.011 (0.071)	0.001 (0.102)
	$\beta_{1,2} = -2$	-0.001 (0.143)	-0.001 (0.073)	0.001 (0.120)		$\beta_{1,2} = -2$	-0.010 (0.144)	-0.003 (0.072)	0.000 (0.119)
j=2	$\beta_{0,2} = -2$	0.000 (0.117)	-0.002 (0.077)	0.002 (0.151)	j=2	$\beta_{0,2} = -2$	0.007 (0.118)	0.001 (0.079)	0.006 (0.151)
	$\beta_{1,2} = 1$	0.001 (0.113)	0.002 (0.070)	0.007 (0.135)		$\beta_{1,2} = 1$	0.003 (0.113)	0.002 (0.070)	0.014 (0.132)
Panel B: T=1000					Panel B: T=1000				
State 1					State 1				
j=1	$\beta_{0,1} = -2$	0.003 (0.100)	0.003 (0.055)	0.006 (0.075)	j=1	$\beta_{0,1} = -2$	0.004 (0.101)	0.004 (0.056)	0.009 (0.082)
	$\beta_{1,1} = 1$	0.009 (0.100)	0.003 (0.055)	0.000 (0.078)		$\beta_{1,1} = 1$	0.012 (0.099)	0.003 (0.056)	0.002 (0.079)
j=2	$\beta_{0,1} = 3$	-0.001 (0.070)	-0.001 (0.054)	-0.005 (0.105)	j=2	$\beta_{0,1} = 3$	0.002 (0.069)	0.001 (0.054)	-0.001 (0.108)
	$\beta_{1,1} = -2$	-0.006 (0.065)	-0.010 (0.049)	-0.010 (0.109)		$\beta_{1,1} = -2$	-0.004 (0.065)	-0.011 (0.050)	-0.013 (0.107)
State 2					State 2				
j=1	$\beta_{0,2} = 3$	0.022 (0.102)	0.011 (0.055)	0.013 (0.086)	j=1	$\beta_{0,2} = 3$	0.009 (0.102)	0.007 (0.055)	0.007 (0.084)
	$\beta_{1,2} = -2$	0.006 (0.093)	0.007 (0.058)	0.014 (0.087)		$\beta_{1,2} = -2$	-0.004 (0.093)	0.005 (0.057)	0.011 (0.087)
j=2	$\beta_{0,2} = -2$	0.006 (0.085)	0.000 (0.054)	0.003 (0.098)	j=2	$\beta_{0,2} = -2$	0.005 (0.085)	0.001 (0.055)	0.010 (0.104)
	$\beta_{1,2} = 1$	-0.009 (0.087)	0.000 (0.059)	0.000 (0.099)		$\beta_{1,2} = 1$	-0.009 (0.086)	0.002 (0.058)	0.010 (0.105)

Table 3.6. Bias and standard error values of the state-regression parameter estimates for CEHMM with skew t distributed errors for $T = 500$ (Panel A) and $T = 1000$ (Panel B).

3.5.1 Descriptive Statistics

Following Pennoni et al. (2022), we choose crypto-assets that met some requirements of scarcity and tradability on reliable exchanges. In light of these decisions, we consider Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Ripple (XRP), Bitcoin cash (BCH) as dependent variables. To investigate for the interlinkages between the crypto markets and non-crypto global markets, we select the S&P500 (GSPC), S&P US Treasury Bond (SPUSBT), US dollar index (USDX), WTI crude oil and Gold as independent variables. The considered timespan extends from July, 25 2017 to December, 19 2022, including numerous crises that have impacted cross-market integration patterns, such as the crypto price bubbles of early 2018, the COVID-19 pandemic, Biden's election at the USA presidency in November 2020 and the Russian invasion of Ukraine at the beginning of 2022, which have caused unprecedented levels of uncertainty and risk. Returns are calculated daily for a total of $T = 1348$ observations. The crypto-assets have been downloaded from Coinbase, while the traditional ones from the SPGlobal.com (for S&P500 and S&P US Treasury Bond), Investing.com (for the US dollar index) and Yahoo finance database for the remaining assets. Figure 3.2 shows the daily prices (top) and log-returns (bottom) of the five cryptocurrencies over the entire period, where the vertical dotted lines indicate globally relevant events occurred during the study period. We immediately recognise the typical characteristics of this market, i.e. high volatility and sudden waves of exponential price increases. Daily log-returns of the five cryptocurrencies confirm their high volatility, a strong degree of comovement, and show the typical volatility clustering in common with other traditional financial assets. We observe volatility jumps not only during the first crypto bubble, but also during the financial market crash caused by COVID-19 pandemic and right after Biden's election at the end of 2020, confirming that crypto investors reactions do not differ from the behavior of the investors in the traditional financial markets. In Table 3.7 we report the list of examined variables and the summary statistics for the whole sample. The high levels of volatility of cryptocurrencies are noticeable, where Ethereum and Bitcoin Cash in particular stand out, having the highest standard deviation. Crypto assets returns also show very high negative skewness and very high kurtosis, as well as S&P500. The highest level of kurtosis is reported by Crude Oil, which was likely determined by prices fluctuations after the COVID-19 outbreak. On the other hand, the positive skewness of S&P Treasury Bond indicates longer and fatter tails on the right side of the distribution, highlighting an inverse relationship with the S&P500. In concluding, the Augmented Dickey-Fuller (ADF) test Dickey & Fuller (1979) shows that all daily returns are stationary at the 1% level of significance. The bottom part of Table 3.7 also reports the empirical correlations among the

dependent variables. As expected, the five cryptocurrencies are all highly correlated, justifying the multivariate approach of the paper. Following these considerations, the proposed copula-based QHMM and EHMM are able to provide useful insights on the evolution of the relationships within crypto assets and between non-crypto markets under different market conditions.

	Min	Mean	Max	Stdev	Skewness	Kurtosis	Jarque-Bera test	ADF test
BTC	-46.47	0.13	22.51	4.89	-0.77	8.79	**4471.41**	**-9.44**
ETH	-55.07	0.12	34.35	6.28	-0.66	7.42	**3186.99**	**-9.57**
LTC	-44.91	0.03	53.98	6.62	-0.03	9.05	**4602.24**	**-9.89**
XRP	-55.05	0.04	62.67	7.60	0.74	15.10	**12927.1**	**-9.49**
BCH	-56.13	-0.11	43.16	8.06	0.00	8.08	**3670.89**	**-9.89**
GSPC	-12.77	0.04	8.97	1.33	-0.83	14.11	**11340.81**	**-9.87**
SPUSBT	-1.69	0.00	1.79	0.28	0.15	5.10	**1467.35**	**-10.14**
USDX	-2.17	0.01	2.10	0.43	0.02	2.01	**226.45**	**-11.38**
WTI	-28.22	0.08	31.96	3.29	0.04	24.70	**34272.07**	**-8.09**
GOLD	-5.11	0.03	5.78	0.94	-0.20	4.95	**1387.37**	**-10.76**

Correlation matrix

	ETH	LTC	XRP	BCH
BTC	0.76	0.74	0.52	0.63
ETH		0.80	0.62	0.68
LTC			0.59	0.66
XRP				0.53

Table 3.7. Descriptive statistics for the whole sample. The Jarque-Bera test and the ADF test statistics are displayed in boldface when the null hypothesis is rejected at the 1% significance level.

3.5.2 Main Results

In order to apply the aforementioned methods to cryptocurrency returns we consider the model in (3.9) as:

$$\boldsymbol{\mu}_{t,k}^{Crypto} = \boldsymbol{X}_{t-1}\boldsymbol{\beta}_k(\boldsymbol{\tau}), \tag{3.24}$$

where $\boldsymbol{\mu}_{tk}^{Crypto} = (\mu_{tk}^{BTC}, \mu_{tk}^{ETH}, \mu_{tk}^{LTC}, \mu_{tk}^{XRP}, \mu_{tk}^{BCH})$ denotes the vector of all component-wise quantile (expectile) regression functions of the cryptos considered (BTC, ETH, LTC, XRP, BCH) at time t for a given state $k = 1, \ldots, K$, $\boldsymbol{X}_{t-1}$ is the vector of returns at the previous date for the S&P500, SPUSBT, USDX, WTI and Gold, with the first element being the intercept, and $\boldsymbol{\beta}_k(\boldsymbol{\tau})$ is the $p \times d$ matrix of state-dependent regression coefficients defined as $\boldsymbol{\beta}_k(\boldsymbol{\tau}) = (\boldsymbol{\beta}_{1,k}(\tau_1), \ldots, \boldsymbol{\beta}_{d,k}(\tau_d))$. The first step of the empirical analysis requires the choice of the appropriate copula and number of states. In order to do so, we fit the proposed CQHMM and CEHMM using the Gaussian and t copulas described in Section 3.3 for a grid of different values of K, spanning from 1 to 4. For ease of comparison between the two models, we select the copula function and K for $\tau = \tau_j = 0.50$, $j = 1, \ldots, d$, and we keep them fixed at

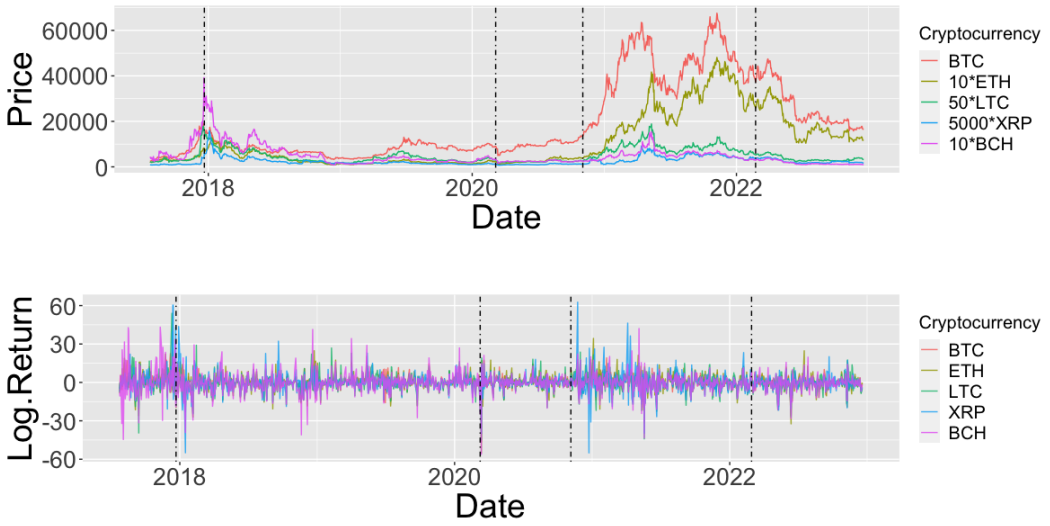

Figure 3.2. Cryptocurrencies daily prices (top) and log return (bottom) series. Vertical dashed lines indicate globally relevant events in the financial markets that occurred in 2017,12; 2020,03; 2020,11; and 2022,02. Prices are multiplied by a constant to have a similar scale.

the other values of $\tau_j \neq 0.50$, $j = 1, \ldots, d$. We consider three widely used penalized likelihood selection criteria, namely the AIC (Akaike 1998), the BIC (Schwarz et al. 1978) and the ICL (Biernacki et al. 2000), and report the results in Table 3.8. In order to clearly identify high and low volatility market conditions, we use $K = 2$ which is supported by the parsimonious ICL criteria for both CQHMM and CEHMM, together with a t copula. From a graphical perspective, in Figure 3.3 we report the scatterplots and the marginal densities colored according to the estimated posterior probability of class membership, $\max_k \gamma_t(k)$. We thus fit the CQHMM and CEHMM under the t copula for $K = 2$ hidden states at three levels $\tau = \{0.05, 0.5, 0.95\}$, with $\tau = \tau_j$ for all $j = 1, \ldots, d$. The estimates of the state-specific parameters are gathered in Tables 3.9, 3.10 and 3.11 for the CQHMM and in Tables 3.12, 3.13 and 3.14 for the CEHMM, along with the standard errors (in brackets), computed by using the parametric bootstrap technique illustrated in Section 3.3 over $R = 1000$ resamples. We start the discussion by looking at the estimated state-dependent scale parameters, σ_1 and σ_2, and the degrees of freedom of the t copula, ν_1 and ν_2 for all the models considered. As regards the scale parameters, the first reflects more stable periods and represents the so-called low-volatility state, meanwhile σ_2 contemplates rapid (positive and negative) peak and burst returns, defining the second state as the high-volatility state, confirming the graphical analysis conducted in Figure 3.3. By looking at the state-dependent degrees of freedom, the results show the necessity

to forgo the Gaussian copula in favor of the more robust, fat-tailed t copula for dependence modeling in financial data.

Taking the impact of covariates into account, we first comment on the parameter estimates of the CQHMM (see Tables 3.9, 3.10 and 3.11). As could be expected, the state-specific intercepts are increasing somewhat with τ, with State 1 having lower absolute values than State 2 for all τ's. For $\tau = 0.5$ (see Table 3.10), we observe that at low volatility periods S&P500 is the only statistically significant asset, negatively influencing almost all the cryptocurrencies for both quantile and expectile models, which implies that during tranquil periods crypto assets can be considered as weak hedges (Bouri, Jalkh, Molnár & Roubaud 2017, Bouri, Molnár, Azzi, Roubaud & Hagfors 2017). During high volatility periods we observe that S&P500, Gold and Crude Oil influence some cryptos, especially Bitcoin and Litecoin, while in the CEHMM (Table 3.13) there is no statistical association among cryptocurrencies and the assets considered. The number of significantly parameters increases by moving to the extreme tails of cryptocurrencies returns distribution, exposing a connection during bearish and bullish market periods between traditional financial markets and the crypto market. In particular, for the CQHMM at $\tau = 0.05$ (Table 3.9) we observe that at low volatility periods Gold negatively influences all the cryptos considered. It can also be seen that the negative impact of S&P Treasury Bond is statistically significant with respect to all crypto-assets considered, with the exception of Ethereum. In the second state, most of the regression parameters are significantly different from zero, with few exceptions. We highlight in particular the strong positive influence of S&P Treasury Bond and the negative one of US dollar index. The CEHMM shows similar results at $\tau = 0.05$ for the second state (Table 3.12), where we highlight the strong positive influences of S&P500 and S&P Treasury Bond, with the exception of Ripple and Bitcoin Cash. Finally, at $\tau = 0.95$ for the CQHMM (Table 3.11) for both states considered we note strong influences of S&P Treasury Bond, US dollar index and Gold. Similar results can be highlighted for the CEHMM (Table 3.14) in the high volatility state, especially regarding the strong negative influences of US dollar index and Gold. Overall, these results are consistent with the recent works of Bouri et al. (2020*a*), Conlon & McGee (2020), Corbet et al. (2020), Caferra & Vidal-Tomás (2021) and Yousaf & Ali (2021) and highlight that: (i) the relationship among these two different markets is rather complex, and it is more pronounced in the tails of the returns distributions; (ii) the dependence within the crypto market varies over time according to the market conditions.

Finally, Figure 3.4 reports the estimated pair-wise correlations of the t copula for both states under the CQHMM and CEHMM models at $\tau = \{0.05, 0.50, 0.95\}$. Overall we observe higher correlation estimates during low volatility states, especially at $\tau = 0.50$. Looking at specific values, we highlight the high correlation levels

between BTC, ETH and LTC and, on the other hand it is visible the divergent behavior of XRP with respect to the other cryptocurrencies (Pennoni et al. 2022).

	Gaussian copula			t copula		
	AIC	BIC	ICL	AIC	BIC	ICL
CQHMM						
$K = 1$	37322.29	37556.55	37556.55	35970.66	36210.11	36210.11
$K = 2$	35326.37	**35810.50**	**36137.40**	35160.71	**35655.25**	**35821.08**
$K = 3$	35078.00	35822.41	36609.45	34988.46	35748.48	36655.90
$K = 4$	**34871.65**	35886.75	36908.26	**34802.28**	35838.20	36657.36
CEHMM						
$K = 1$	35157.65	35362.76	35362.76	34455.43	34665.67	34665.67
$K = 2$	32354.36	32779.96	**33075.53**	32308.68	32744.53	**33067.44**
$K = 3$	31869.51	**32525.86**	33140.76	31871.03	**32542.76**	33203.10
$K = 4$	**31635.38**	32532.72	33390.67	**31642.18**	32560.04	33387.36

Table 3.8. AIC, BIC and ICL values with varying number of states for the CQHMM and CEHMM under the Gaussian and t copulas. Bold font highlights the best values for the considered criteria (lower-is-better).

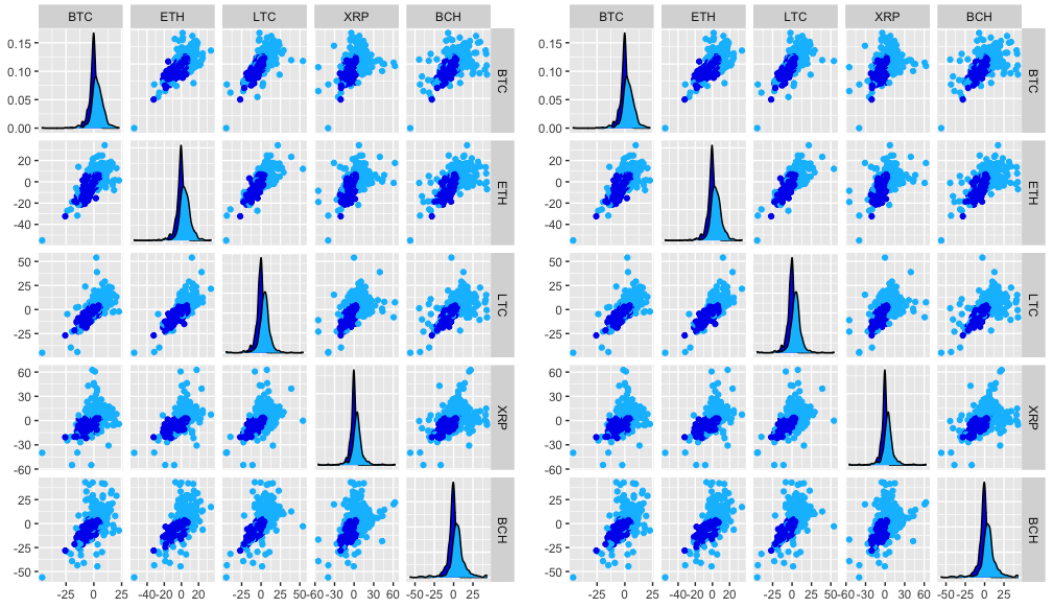

Figure 3.3. Cryptocurrencies marginal distributions and scatterplots classified according to the estimated posterior probability for CQHMM (left) and CEHMM (right) of class membership for $\tau = 0.5$. Dark-blue points (State 1) identify low volatility periods, while light-blue ones (State 2) identify high volatility periods.

CQHMM	BTC	ETH	LTC	XRP	BCH
State 1					
Intercept	**-2.65 (0.038)**	**-3.575 (0.051)**	**-3.957 (0.055)**	**-3.926 (0.051)**	**-4.205 (0.056)**
GSPC	**-0.166 (0.035)**	-0.075 (0.047)	**-0.149 (0.051)**	-0.095 (0.051)	**-0.148 (0.053)**
SPUSBT	**-0.716 (0.166)**	-0.061 (0.218)	**-1.897 (0.234)**	**-1.078 (0.24)**	**-1.117 (0.244)**
USDX	0.109 (0.108)	-0.261 (0.15)	**-0.609 (0.159)**	**-0.732 (0.154)**	**-0.714 (0.167)**
WTI	-0.008 (0.015)	**-0.079 (0.02)**	**-0.084 (0.021)**	**-0.148 (0.02)**	**-0.089 (0.023)**
GOLD	**-0.138 (0.051)**	**-0.428 (0.066)**	**-0.175 (0.072)**	**-0.377 (0.071)**	**-0.44 (0.077)**
σ_1	0.237 (0.008)	0.313 (0.011)	0.34 (0.011)	0.328 (0.011)	0.353 (0.012)
ν_1	5.022 (0.505)				
State 2					
Intercept	**-14.552 (0.224)**	**-19.013 (0.297)**	**-18.956 (0.294)**	**-19.268 (0.361)**	**-22.306 (0.401)**
GSPC	**0.809 (0.23)**	0.567 (0.292)	**1.765 (0.296)**	**1.172 (0.364)**	**2.639 (0.43)**
SPUSBT	**7.348 (1.031)**	**7.462 (1.429)**	**3.938 (1.356)**	1.588 (1.704)	-1.034 (1.934)
USDX	**1.402 (0.685)**	**-3.617 (0.866)**	-1.315 (0.895)	**-4.287 (1.108)**	**-2.4 (1.202)**
WTI	**0.472 (0.097)**	**0.908 (0.129)**	**0.403 (0.129)**	**0.757 (0.155)**	-0.132 (0.181)
GOLD	0.088 (0.326)	**-0.872 (0.425)**	**1.223 (0.433)**	-0.546 (0.552)	0.414 (0.589)
σ_2	0.692 (0.04)	0.909 (0.056)	0.898 (0.054)	1.102 (0.064)	1.24 (0.073)
ν_2	7.867 (2.485)				

Table 3.9. CQHMM state-specific parameter estimates for $\tau = 0.05$, with bootstrapped standard errors (in brackets) obtained over 1000 replications. Point estimates are displayed in boldface when significant at the standard 5% level. σ_k and ν_k represent the state-specific scale parameter and degrees of freedom, respectively.

CQHMM	BTC	ETH	LTC	XRP	BCH
State 1					
Intercept	0.029 (0.096)	-0.125 (0.125)	-0.187 (0.125)	**-0.249 (0.117)**	**-0.383 (0.132)**
GSPC	**-0.197 (0.083)**	**-0.307 (0.108)**	-0.143 (0.114)	**-0.227 (0.098)**	**-0.374 (0.11)**
SPUSBT	-0.011 (0.405)	-0.54 (0.546)	-0.341 (0.524)	-0.123 (0.511)	-0.275 (0.538)
USDX	0.093 (0.264)	0.038 (0.342)	0.036 (0.344)	0.151 (0.332)	-0.103 (0.349)
WTI	-0.005 (0.034)	-0.029 (0.044)	-0.07 (0.047)	0.006 (0.041)	0.003 (0.044)
GOLD	-0.057 (0.121)	-0.09 (0.158)	-0.175 (0.158)	-0.072 (0.152)	-0.072 (0.162)
σ_1	1.485 (0.046)	1.975 (0.061)	2 (0.061)	1.842 (0.056)	2.014 (0.06)
ν_1	5.73 (0.625)				
State 2					
Intercept	**1.438 (0.314)**	**1.198 (0.399)**	**1.556 (0.457)**	1.209 (0.651)	**2.263 (0.723)**
GSPC	**0.58 (0.292)**	**0.922 (0.376)**	**1.077 (0.413)**	0.099 (0.605)	-0.122 (0.658)
SPUSBT	0.192 (1.447)	1.311 (1.731)	-0.286 (1.88)	-1.741 (2.822)	0.738 (3.045)
USDX	0.163 (0.945)	1.04 (1.13)	1.823 (1.271)	-3.139 (1.876)	0.566 (2.027)
WTI	**0.262 (0.133)**	0.148 (0.152)	**0.497 (0.178)**	0.232 (0.256)	**0.649 (0.283)**
GOLD	-0.364 (0.429)	-0.06 (0.542)	**1.196 (0.601)**	-1.509 (0.812)	-0.901 (0.934)
σ_2	2.231 (0.145)	2.739 (0.174)	3.06 (0.19)	4.456 (0.284)	4.911 (0.312)
ν_2	8.276 (2.823)				

Table 3.10. CQHMM state-specific parameter estimates for $\tau = 0.50$, with bootstrapped standard errors (in brackets) obtained over 1000 replications. Point estimates are displayed in boldface when significant at the standard 5% level. σ_k and ν_k represent the state-specific scale parameter and degrees of freedom, respectively.

CQHMM	BTC	ETH	LTC	XRP	BCH
State 1					
Intercept	**3.062 (0.051)**	**4.19 (0.064)**	**3.008 (0.057)**	**2.703 (0.054)**	**2.805 (0.058)**
GSPC	**0.103 (0.047)**	**0.139 (0.063)**	**-0.313 (0.053)**	**-0.321 (0.05)**	**-0.342 (0.055)**
SPUSBT	**1.607 (0.219)**	**2.133 (0.287)**	-0.152 (0.255)	-0.107 (0.243)	0.387 (0.273)
USDX	**0.396 (0.145)**	**0.584 (0.18)**	**-0.747 (0.163)**	**-0.409 (0.146)**	-0.115 (0.17)
WTI	**0.056 (0.02)**	-0.033 (0.025)	0.004 (0.022)	-0.001 (0.021)	-0.013 (0.023)
GOLD	**-0.378 (0.065)**	**-0.557 (0.082)**	0.047 (0.076)	-0.112 (0.07)	-0.125 (0.081)
σ_1	0.268 (0.009)	0.354 (0.013)	0.31 (0.011)	0.294 (0.01)	0.324 (0.011)
ν_1	3.382 (0.311)				
State 2					
Intercept	**10.69 (0.142)**	**13.979 (0.182)**	**14.249 (0.203)**	**17.824 (0.305)**	**21.306 (0.3)**
GSPC	-0.034 (0.138)	0.182 (0.176)	0.034 (0.203)	**-2.569 (0.313)**	-0.236 (0.289)
SPUSBT	-0.064 (0.672)	1.051 (0.807)	**1.908 (0.895)**	**-11.908 (1.414)**	**-9.343 (1.356)**
USDX	-0.79 (0.405)	**3.829 (0.538)**	**-1.667 (0.615)**	**-4.206 (0.906)**	**-4.33 (0.896)**
WTI	**-0.181 (0.058)**	0.116 (0.071)	0.017 (0.078)	-0.028 (0.131)	0.184 (0.124)
GOLD	**-0.697 (0.196)**	-0.458 (0.242)	**-2.435 (0.287)**	**1.839 (0.43)**	**-2.912 (0.422)**
σ_2	0.565 (0.026)	0.707 (0.033)	0.808 (0.038)	1.228 (0.055)	1.195 (0.053)
ν_2	6.265 (0.956)				

Table 3.11. CQHMM state-specific parameter estimates for $\tau = 0.95$, with bootstrapped standard errors (in brackets) obtained over 1000 replications. Point estimates are displayed in boldface when significant at the standard 5% level. σ_k and ν_k represent the state-specific scale parameter and degrees of freedom, respectively.

CEHMM	BTC	ETH	LTC	XRP	BCH
State 1					
Intercept	**-2.666 (0.085)**	**-3.448 (0.106)**	**-3.755 (0.108)**	**-3.674 (0.106)**	**-3.894 (0.109)**
GSPC	**-0.164 (0.072)**	-0.134 (0.093)	-0.051 (0.094)	-0.097 (0.092)	-0.14 (0.092)
SPUSBT	**-0.675 (0.34)**	-0.79 (0.425)	-0.567 (0.412)	-0.6 (0.416)	-0.132 (0.434)
USDX	0.344 (0.229)	-0.254 (0.28)	0.124 (0.291)	-0.384 (0.278)	-0.336 (0.299)
WTI	-0.019 (0.029)	-0.041 (0.036)	-0.05 (0.037)	**-0.076 (0.035)**	-0.059 (0.037)
GOLD	**0.223 (0.104)**	0.181 (0.13)	0.173 (0.131)	0.065 (0.127)	0.157 (0.137)
σ_1	1.812 (0.041)	2.247 (0.051)	2.33 (0.053)	2.225 (0.052)	2.322 (0.052)
ν_1	6.985 (0.833)				
State 2					
Intercept	**-12.027 (0.495)**	**-15.603 (0.626)**	**-15.652 (0.67)**	**-16.562 (0.848)**	**-19.073 (0.844)**
GSPC	**1.743 (0.416)**	**1.719 (0.546)**	**2.097 (0.594)**	1.47 (0.775)	**3.203 (0.769)**
SPUSBT	**8.728 (1.938)**	**10.294 (2.685)**	**6.881 (2.717)**	2.749 (3.616)	2.861 (3.713)
USDX	-0.612 (1.325)	-1.18 (1.696)	0.402 (1.864)	0.016 (2.361)	3.631 (2.448)
WTI	**0.684 (0.177)**	**0.879 (0.227)**	**0.621 (0.251)**	**0.818 (0.317)**	0.116 (0.319)
GOLD	0.641 (0.595)	0.755 (0.798)	1.652 (0.847)	1.057 (1.095)	1.775 (1.132)
σ_2	4.66 (0.217)	6.13 (0.294)	6.461 (0.302)	8.436 (0.388)	8.442 (0.395)
ν_2	14.816 (93.272)				

Table 3.12. CEHMM state-specific parameter estimates for $\tau = 0.05$, with bootstrapped standard errors (in brackets) obtained over 1000 replications. Point estimates are displayed in boldface when significant at the standard 5% level. σ_k and ν_k represent the state-specific scale parameter and degrees of freedom, respectively.

CEHMM	BTC	ETH	LTC	XRP	BCH
State 1					
Intercept	0.054 (0.109)	-0.019 (0.137)	**-0.358 (0.132)**	**-0.402 (0.127)**	**-0.547 (0.136)**
GSPC	**-0.228 (0.091)**	**-0.243 (0.117)**	**-0.232 (0.111)**	-0.199 (0.106)	**-0.261 (0.117)**
SPUSBT	-0.134 (0.454)	-0.173 (0.576)	-0.677 (0.545)	-0.339 (0.525)	0.093 (0.563)
USDX	-0.007 (0.284)	-0.123 (0.369)	-0.113 (0.351)	-0.061 (0.338)	-0.16 (0.355)
WTI	0.029 (0.036)	0.008 (0.047)	0.001 (0.044)	-0.023 (0.043)	0.008 (0.045)
GOLD	-0.051 (0.131)	-0.101 (0.166)	-0.052 (0.161)	0.005 (0.154)	-0.162 (0.162)
σ_1	3.388 (0.081)	4.351 (0.101)	4.136 (0.097)	3.981 (0.095)	4.227 (0.098)
ν_1	11.101 (2.09)				
State 2					
Intercept	0.467 (0.439)	0.678 (0.57)	**1.339 (0.613)**	**1.46 (0.731)**	1.4 (0.79)
GSPC	0.294 (0.381)	0.31 (0.497)	0.326 (0.547)	-0.137 (0.656)	0.61 (0.689)
SPUSBT	1.355 (1.816)	3.534 (2.264)	1.874 (2.603)	-1.448 (3.076)	-2.041 (3.31)
USDX	0.767 (1.182)	1.861 (1.55)	1.493 (1.748)	-1.25 (2.062)	0.698 (2.12)
WTI	-0.016 (0.153)	0.043 (0.195)	0.156 (0.219)	0.129 (0.262)	0.106 (0.282)
GOLD	-0.233 (0.56)	-0.121 (0.714)	-0.37 (0.775)	-0.305 (0.911)	-0.81 (0.995)
σ_2	7.822 (0.318)	10.026 (0.41)	11.074 (0.459)	13.522 (0.527)	14.285 (0.576)
ν_2	12.806 (9.207)				

Table 3.13. CEHMM state-specific parameter estimates for $\tau = 0.50$, with bootstrapped standard errors (in brackets) obtained over 1000 replications. Point estimates are displayed in boldface when significant at the standard 5% level. σ_k and ν_k represent the state-specific scale parameter and degrees of freedom, respectively.

CEHMM	BTC	ETH	LTC	XRP	BCH
State 1					
Intercept	**2.466 (0.09)**	**3.443 (0.121)**	**2.499 (0.108)**	**2.132 (0.102)**	**2.388 (0.114)**
GSPC	**-0.153 (0.073)**	-0.177 (0.099)	**-0.296 (0.091)**	**-0.327 (0.084)**	**-0.247 (0.096)**
SPUSBT	**0.924 (0.359)**	**1.275 (0.483)**	-0.087 (0.434)	0.008 (0.394)	0.298 (0.447)
USDX	-0.056 (0.229)	-0.572 (0.318)	-0.378 (0.282)	-0.432 (0.259)	-0.532 (0.291)
WTI	0.035 (0.029)	-0.029 (0.039)	0.016 (0.035)	-0.001 (0.033)	-0.014 (0.036)
GOLD	-0.124 (0.109)	-0.205 (0.145)	-0.037 (0.132)	-0.093 (0.122)	-0.248 (0.137)
σ_1	1.731 (0.041)	2.383 (0.055)	2.152 (0.049)	1.981 (0.045)	2.234 (0.052)
ν_1	5.753 (0.673)				
State 2					
Intercept	**9.344 (0.291)**	**11.905 (0.345)**	**13.356 (0.448)**	**17.59 (0.621)**	**17.766 (0.593)**
GSPC	0.023 (0.253)	0.236 (0.329)	0.401 (0.385)	-1.03 (0.608)	**-0.991 (0.49)**
SPUSBT	-0.692 (1.199)	**4.962 (1.553)**	3.312 (1.865)	-3.591 (2.733)	-2.229 (2.527)
USDX	-0.206 (0.807)	**3.267 (0.999)**	-0.244 (1.264)	**-6.721 (1.789)**	-2.002 (1.674)
WTI	-0.137 (0.109)	0.034 (0.127)	0.105 (0.16)	0.074 (0.228)	0.322 (0.197)
GOLD	-0.343 (0.363)	-0.707 (0.47)	**-2.874 (0.573)**	0.831 (0.845)	**-1.957 (0.751)**
σ_2	3.636 (0.137)	4.6 (0.173)	5.645 (0.206)	8.079 (0.293)	7.335 (0.269)
ν_2	6.102 (1.1)				

Table 3.14. CEHMM state-specific parameter estimates for $\tau = 0.95$, with bootstrapped standard errors (in brackets) obtained over 1000 replications. Point estimates are displayed in boldface when significant at the standard 5% level. σ_k and ν_k represent the state-specific scale parameter and degrees of freedom, respectively.

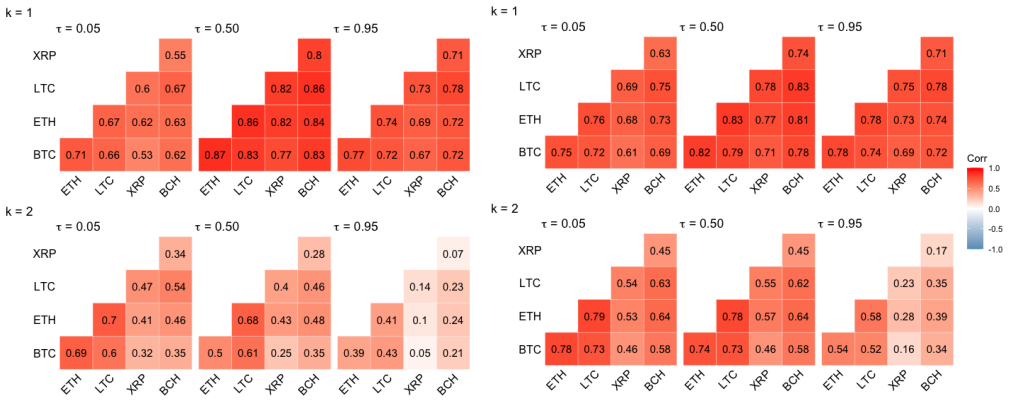

Figure 3.4. CQHMM (left) and CEHMM (right) state-specific correlation estimates for the two states at $\tau = \{0.05, 0.50, 0.95\}$.

3.6 Conclusions

Since cryptocurrencies were first introduced in 2008 as a form of payment and later turned into high yield assets, scholars and professionals have questioned their extreme volatility and fluctuations over time. The early 2018 crypto-bubble and the COVID-19 market crash in 2020 have contributed to strengthening the focus on the relationship between digital currencies and traditional financial assets. In this context, our contribution to the existing literature is twofold. From a theoretical standpoint, we develop hidden Markov regression models for joint estimation of conditional quantiles or expectiles of multivariate time series. To grasp unobserved serial heterogeneity and rapid volatility jumps, the regression parameters vary according to an unobservable homogeneous Marov chain. At the same time, the method proposed allows us to model the time-varying dependence structure of cryptocurrency returns using state-dependent elliptical copula functions. From a practical point of view, we jointly investigate the impact of global market indices on daily returns of the five most important cryptocurrencies from 2017 to 2022, while taking into account for their association under different market conditions. We found that interrelations between crypto and stocks increase while moving to the extreme tails of returns distributions. In particular, a weak relationship between cryptocurrencies and stock markets and commodities occurr at the centre of the returns' distributions, concurring with the consensus that cryptocurrencies are good diversifiers from stocks and commodities during periods of tranquillity of financial markets (Bouri, Molnár, Azzi, Roubaud & Hagfors 2017, Bouri, Jalkh, Molnár & Roubaud 2017, Cremaschini et al. 2022). Conversely, we highlight an important influence of S&P500, S&P Treasury Bond and Gold during both bearish and bullish periods. Overall, these results are consistent the existing strands of literature on the subject (Bouri et al. 2020*a*, Caferra & Vidal-Tomás 2021, Yousaf & Ali 2021). Finally, as regards the dependence analysis among the crypto, our results seem to adhere to the ones in Pennoni et al. (2022).

Future topics of research could extend the proposed methods to the hidden semi-Markov model setting where the sojourn-distributions, that is, the distributions of the number of consecutive time points that the chain spends in each state, are modeled by the researcher using either parametric or non-parametric approaches instead of assuming geometric sojourn densities as in HMMs. Moreover, even tough the application focused on five cryptocurrencies and five predictors, in high-dimensional settings with a greater number of response variables and/or hidden states, the described models can be easily over-parameterized. This is frequently the case due to the large number of regression parameters and unique parameters in the correlation matrices of the copulas to be estimated, meaning a loss in interpretability

as well as numerically ill-conditioned estimators. In these situations, one could specify a parsimonious parametrization of correlation matrices of copula functions or, following Maruotti et al. (2017), consider mixtures of factor analysis models, whose parameters evolve according to the latent Markov chain.

Bibliography

Adam, T., Langrock, R. & Kneib, T. (2019), Model-based clustering of time series data: a flexible approach using nonparametric state-switching quantile regression models, *in* 'Proceedings of the 12th Scientific Meeting on Classification and Data Analysis', pp. 8–11.

Adcock, C., Eling, M. & Loperfido, N. (2015), 'Skewed distributions in finance and actuarial science: A review', *The European Journal of Finance* **21**(13-14), 1253–1281.

Agosto, A. & Cafferata, A. (2020), 'Financial bubbles: a study of co-explosivity in the cryptocurrency market', *Risks* **8**(2), 34.

Ajanovic, A. (2011), 'Biofuels versus food production: Does biofuels production increase food prices?', *Energy* **36**(4), 2070–2076.

Akaike, H. (1998), Information theory and an extension of the maximum likelihood principle, *in* 'Selected papers of Hirotugu Akaike', Springer, pp. 199–213.

Al-Maadid, A., Caporale, G. M., Spagnolo, F. & Spagnolo, N. (2017), 'Spillovers between food and energy prices and structural breaks', *International Economics* **150**, 1–18.

Alfò, M., Salvati, N. & Ranallli, M. G. (2017), 'Finite mixtures of quantile and M-quantile regression models', *Statistics and Computing* **27**(2), 547–570.

Algieri, B. & Leccadito, A. (2017), 'Assessing contagion risk from energy and non-energy commodity markets', *Energy Economics* **62**, 312–322.

Aloui, C. & Mabrouk, S. (2010), 'Value-at-Risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models', *Energy Policy* **38**(5), 2326–2339.

Andriosopoulos, K. & Nomikos, N. (2014), 'Performance replication of the Spot Energy Index with optimal equity portfolio selection: Evidence from the UK, US

and Brazilian markets', *European Journal of Operational Research* **234**(2), 571–582.

Ardia, D., Bluteau, K., Boudt, K. & Catania, L. (2018), 'Forecasting risk with Markov-switching GARCH models: A large-scale performance study', *International Journal of Forecasting* **34**(4), 733–747.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), 'Coherent measures of risk', *Math. Finance* **9**(3), 203–228.

Assaf, A., Bhandari, A., Charif, H. & Demir, E. (2022), 'Multivariate long memory structure in the cryptocurrency market: The impact of covid-19', *International Review of Financial Analysis* **82**, 102132.

Aven, T. (2016), 'Risk assessment and risk management: Review of recent advances on their foundation', *European Journal of Operational Research* **253**(1), 1–13.

Awokuse, T. O. & Yang, J. (2003), 'The informational role of commodity prices in formulating monetary policy: A reexamination', *Economics Letters* **79**(2), 219–224.

Bae, G. I., Kim, W. C. & Mulvey, J. M. (2014), 'Dynamic asset allocation for varied financial markets under regime switching framework', *European Journal of Operational Research* **234**(2), 450–458.

Baffes, J. (2007), *Oil spills on other commodities*, The World Bank.

Balcilar, M., Gabauer, D. & Umar, Z. (2021), 'Crude Oil futures contracts and commodity markets: New evidence from a TVP-VAR extended joint connectedness approach', *Resources Policy* **73**, 102219.

Balli, F., Naeem, M. A., Shahzad, S. J. H. & de Bruin, A. (2019), 'Spillover network of commodity uncertainties', *Energy Economics* **81**, 914–927.

Barbaglia, L., Wilms, I. & Croux, C. (2016), 'Commodity dynamics: A sparse multi-class approach', *Energy Economics* **60**, 62–72.

Barber, R. F., Drton, M. et al. (2015), 'High-dimensional Ising model selection with Bayesian information criteria', *Electronic Journal of Statistics* **9**(1), 567–607.

Barry, A., Oualkacha, K. & Charpentier, A. (2021), 'A new GEE method to account for heteroscedasticity using asymmetric least-square regressions', *Journal of Applied Statistics* pp. 1–27.

Basak, S. & Pavlova, A. (2016), 'A model of financialization of commodities', *The Journal of Finance* **71**(4), 1511–1556.

Battiston, S., Mandel, A., Monasterolo, I., Schütze, F. & Visentin, G. (2017), 'A climate stress-test of the financial system', *Nature Climate Change* **7**(4), 283–288.

Baum, L. E. & Petrie, T. (1966), 'Statistical inference for probabilistic functions of finite state markov chains', *The annals of mathematical statistics* **37**(6), 1554–1563.

Baum, L. E., Petrie, T., Soules, G. & Weiss, N. (1970), 'A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains', *The Annals of Mathematical Statistics* **41**(1), 164–171.

Baumeister, C. & Kilian, L. (2016), 'Forty years of oil price fluctuations: Why the price of oil may still surprise us', *Journal of Economic Perspectives* **30**(1), 139–60.

Baur, D. G., Hong, K. & Lee, A. D. (2018), 'Bitcoin: Medium of exchange or speculative assets?', *Journal of International Financial Markets, Institutions and Money* **54**, 177–189.

Bellini, F. (2012), 'Isotonicity properties of generalized quantiles', *Statistics & Probability Letters* **82**(11), 2017–2024.

Bellini, F. & Di Bernardino, E. (2017), 'Risk management with expectiles', *The European Journal of Finance* **23**(6), 487–506.

Bellini, F., Klar, B., Müller, A. & Rosazza Gianin, E. (2014), 'Generalized quantiles as risk measures', *Insurance Math. Econom.* **54**, 41–48.

Bernardi, M., Maruotti, A. & Petrella, L. (2017), 'Multiple risk measures for multivariate dynamic heavy–tailed models', *Journal of Empirical Finance* **43**, 1–32.

Biernacki, C., Celeux, G. & Govaert, G. (2000), 'Assessing a mixture model for clustering with the integrated completed likelihood', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(7), 719–725.

Billio, M., Getmansky, M., Lo, A. W. & Pelizzon, L. (2012), 'Econometric measures of connectedness and systemic risk in the finance and insurance sectors', *Journal of Financial Economics* **104**(3), 535–559.

Bollerslev, T., Chou, R. Y. & Kroner, K. F. (1992), 'ARCH modeling in finance: A review of the theory and empirical evidence', *Journal of Econometrics* **52**(1-2), 5–59.

Borri, N. (2019), 'Conditional tail-risk in cryptocurrency markets', *Journal of Empirical Finance* **50**, 1–19.

Bottone, M., Petrella, L. & Bernardi, M. (2021), 'Unified Bayesian conditional autoregressive risk measures using the skew exponential power distribution', *Statistical Methods & Applications* **30**(3), 1079–1107.

Bouri, E., Jalkh, N., Molnár, P. & Roubaud, D. (2017), 'Bitcoin for energy commodities before and after the december 2013 crash: diversifier, hedge or safe haven?', *Applied Economics* **49**(50), 5063–5073.

Bouri, E., Lucey, B. & Roubaud, D. (2020*a*), 'Cryptocurrencies and the downside risk in equity investments', *Finance Research Letters* **33**, 101211.

Bouri, E., Lucey, B. & Roubaud, D. (2020*b*), 'The volatility surprise of leading cryptocurrencies: Transitory and permanent linkages', *Finance Research Letters* **33**, 101188.

Bouri, E., Molnár, P., Azzi, G., Roubaud, D. & Hagfors, L. I. (2017), 'On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?', *Finance Research Letters* **20**, 192–198.

Browne, F. & Cronin, D. (2010), 'Commodity prices, money and inflation', *Journal of Economics and Business* **62**(4), 331–345.

Bulla, J. & Bulla, I. (2006), 'Stylized facts of financial time series and hidden semi-Markov models', *Computational Statistics & Data Analysis* **51**(4), 2192–2209.

Caferra, R. & Vidal-Tomás, D. (2021), 'Who raised from the abyss? a comparison between cryptocurrency and stock market dynamics during the covid-19 pandemic', *Finance Research Letters* **43**, 101954.

Cashin, P. & McDermott, C. J. (2002), 'The long-run behavior of commodity prices: Small trends and big variability', *IMF staff Papers* **49**(2), 175–199.

Cevik, S. & Sedik, T. S. (2014), 'A barrel of oil or a bottle of wine: How do global growth dynamics affect commodity prices?', *Journal of Wine Economics* **9**(1), 34–50.

Charfeddine, L., Klein, T. & Walther, T. (2020), 'Reviewing the oil price–GDP growth relationship: A replication study', *Energy Economics* p. 104786.

Cheah, E.-T. & Fry, J. (2015), 'Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin', *Economics Letters* **130**, 32–36.

Chen, J. & Chen, Z. (2008), 'Extended Bayesian information criteria for model selection with large model spaces', *Biometrika* **95**(3), 759–771.

Chen, Y.-c., Turnovsky, S. J. & Zivot, E. (2014), 'Forecasting inflation using commodity price aggregates', *Journal of Econometrics* **183**(1), 117–134.

Cheng, I.-H. & Xiong, W. (2014), 'Financialization of commodity markets', *Annual Review of Financial Economics* **6**(1), 419–441.

Cheung, A., Roca, E. & Su, J.-J. (2015), 'Crypto-currency bubbles: an application of the phillips–shi–yu (2013) methodology on mt. gox bitcoin prices', *Applied Economics* **47**(23), 2348–2358.

Chiou-Wei, S.-Z., Chen, S.-H. & Zhu, Z. (2019), 'Energy and agricultural commodity markets interaction: An analysis of crude oil, natural gas, corn, soybean, and ethanol prices', *The Energy Journal* **40**(2).

Christodoulakis, G. (2020), 'Estimating the term structure of commodity market preferences', *European Journal of Operational Research* **282**(3), 1146–1163.

Christoffersen, P. (2010), 'Backtesting', *Encyclopedia of Quantitative Finance* .

Christoffersen, P. F. (1998), 'Evaluating interval forecasts', *International Economic Review* **39**(4), 841–862.

Christoffersen, P. F. & Diebold, F. X. (2000), 'How relevant is volatility forecasting for financial risk management?', *Review of Economics and Statistics* **82**(1), 12–22.

Chuliá, H., Guillén, M. & Uribe, J. M. (2017), 'Measuring uncertainty in the stock market', *International Review of Economics & Finance* **48**, 18–33.

Ciner, C. (2001), 'On the long run relationship between gold and silver prices. A note', *Global Finance Journal* **12**(2), 299–303.

Ciner, C., Lucey, B. & Yarovaya, L. (2022), 'Determinants of cryptocurrency returns: A lasso quantile regression approach', *Finance Research Letters* **49**, 102990.

Conlon, T. & McGee, R. (2020), 'Safe haven or risky hazard? bitcoin during the covid-19 bear market', *Finance Research Letters* **35**, 101607.

Corbet, S., Larkin, C. & Lucey, B. (2020), 'The contagion effects of the covid-19 pandemic: Evidence from gold and cryptocurrencies', *Finance Research Letters* **35**, 101554.

Corbet, S., Lucey, B., Urquhart, A. & Yarovaya, L. (2019), 'Cryptocurrencies as a financial asset: A systematic analysis', *International Review of Financial Analysis* **62**, 182–199.

Corbet, S., Meegan, A., Larkin, C., Lucey, B. & Yarovaya, L. (2018), 'Exploring the dynamic relationships between cryptocurrencies and other financial assets', *Economics Letters* **165**, 28–34.

Cossin, D., Schellhorn, H., Song, N. & Tungsong, S. (2010), 'A theoretical argument why the t-copula explains credit risk contagion better than the gaussian copula', *Advances in Decision Sciences* **2010**.

Cremaschini, A., Punzo, A., Martellucci, E. & Maruotti, A. (2022), 'On stylized facts of cryptocurrencies returns and their relationship with other assets, with a focus on the impact of covid-19', *Applied Economics* pp. 1–14.

Danielsson, J. (2011), Risk and crises: How models failed and are failing. <<voxeu.org>>, Technical report.

Das, D., Le Roux, C. L., Jana, R. K. & Dutta, A. (2020), 'Does bitcoin hedge crude oil implied volatility and structural shocks? a comparison with gold, commodity and the us dollar', *Finance Research Letters* **36**, 101335.

Das, S. R. & Sundaram, R. K. (1997), 'Taming the skew: Higher-order moments in modeling asset price processes in finance', *National Bureau of Economic Research* .

De Angelis, L. & Paas, L. J. (2013), 'A dynamic analysis of stock markets using a hidden Markov model', *Journal of Applied Statistics* **40**(8), 1682–1700.

De Luca, G., Genton, M. G. & Loperfido, N. (2006), 'A multivariate skew-GARCH model', *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series, Part A (Special volume in honor of Robert Engle and Clive Granger, the 2003 winners of the Nobel Prize in Economics)* pp. 33–57.

de Nicola, F., De Pace, P. & Hernandez, M. A. (2016), 'Co-movement of major energy, agricultural, and food commodity price returns: A time-series assessment', *Energy Economics* **57**, 28–41.

Deaton, A. (1999), 'Commodity prices and growth in Africa', *Journal of Economic Perspectives* **13**(3), 23–40.

Demir, E., Bilgin, M. H., Karabulut, G. & Doker, A. C. (2020), 'The relationship between cryptocurrencies and covid-19 pandemic', *Eurasian Economic Review* **10**(3), 349–360.

Dickey, D. A. & Fuller, W. A. (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**(366a), 427–431.

Diebold, F. X., Liu, L. & Yilmaz, K. (2017), Commodity connectedness, Technical report, National Bureau of Economic Research.

Diebold, F. X. & Yılmaz, K. (2014), 'On the network topology of variance decompositions: Measuring the connectedness of financial firms', *Journal of Econometrics* **182**(1), 119–134.

Du, X., Cindy, L. Y. & Hayes, D. J. (2011), 'Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis', *Energy Economics* **33**(3), 497–503.

Duan, K., Li, Z., Urquhart, A. & Ye, J. (2021), 'Dynamic efficiency and arbitrage potential in bitcoin: A long-memory approach', *International Review of Financial Analysis* **75**, 101725.

Dunis, C. L., Laws, J. & Evans, B. (2009), Modelling and trading the soybean-oil crush spread with recurrent and higher order networks: A comparative analysis, *in* 'Artificial Higher Order Neural Networks for Economics and Business', IGI Global, pp. 348–366.

Dyhrberg, A. H. (2016), 'Hedging capabilities of Bitcoin. Is it the virtual gold?', *Finance Research Letters* **16**, 139–144.

Embrechts, P., Frey, R. & McNeil, A. (2011), 'Quantitative risk management.'.

Embrechts, P., Lindskog, F. & McNeil, A. (2001), 'Modelling dependence with copulas', *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich* **14**, 1–50.

Engle, R. F. & Manganelli, S. (2004), 'CAViaR: Conditional Autoregressive Value at Risk by Regression quantiles', *Journal of Business & Economic Statistics* **22**(4), 367–381.

Ergen, I. & Rizvanoghlu, I. (2016), 'Asymmetric impacts of fundamentals on the natural gas futures volatility: An augmented GARCH approach', *Energy Economics* **56**, 64–74.

Ewing, B. T. & Malik, F. (2013), 'Volatility transmission between gold and oil futures under structural breaks', *International Review of Economics & Finance* **25**, 113–121.

Farcomeni, A. (2012), 'Quantile regression for longitudinal data based on latent markov subject-specific parameters', *Statistics and Computing* **22**(1), 141–152.

Fernández, C. & Steel, M. F. (1998), 'On Bayesian modeling of fat tails and skewness', *Journal of the American Statistical Association* **93**(441), 359–371.

Ferrer, R., Shahzad, S. J. H., López, R. & Jareño, F. (2018), 'Time and frequency dynamics of connectedness between renewable energy stocks and crude oil prices', *Energy Economics* **76**, 1–20.

Flori, A., Pammolli, F. & Spelta, A. (2021), 'Commodity prices co-movements and financial stability: A multidimensional visibility nexus with climate conditions', *Journal of Financial Stability* p. 100876.

Fong, W. M. & See, K. H. (2002), 'A Markov switching model of the conditional volatility of crude oil futures prices', *Energy Economics* **24**(1), 71–95.

Fong, W. M. & See, K. H. (2003), 'Basis variations and regime shifts in the oil futures market', *The European Journal of Finance* **9**(5), 499–513.

Foroni, B., Merlo, L. & Petrella, L. (2023), 'Expectile hidden markov regression models for analyzing cryptocurrency returns', *arXiv preprint arXiv:2301.09722* .

Foroni, B., Morelli, G. & Petrella, L. (2022), 'The network of commodity risk', *Energy Systems* pp. 1–47.

Foygel, R. & Drton, M. (2010), Extended Bayesian information criteria for Gaussian graphical models, *in* 'Proceedings of the 23rd International Conference on Neural Information Processing Systems-Volume 1', pp. 604–612.

Friedman, J., Hastie, T. & Tibshirani, R. (2008*a*), 'Sparse inverse covariance estimation with the graphical lasso', *Biostatistics* **9**(3), 432–441.

Friedman, J., Hastie, T. & Tibshirani, R. (2008*b*), 'Sparse inverse covariance estimation with the graphical LASSO', *Biostatistics (Oxford, England)* **9**, 432–41.

Gabrel, V., Murat, C. & Thiele, A. (2014), 'Recent advances in robust optimization: An overview', *European Journal of Operational Research* **235**(3), 471–483.

Garman, M. B. & Klass, M. J. (1980), 'On the estimation of security price volatilities from historical data', *Journal of Business* pp. 67–78.

Gelos, G. & Ustyugova, Y. (2017), 'Inflation responses to commodity price shocks–how and why do countries differ?', *Journal of International Money and Finance* **72**, 28–47.

Gerlach, R. & Chen, C. W. (2015), 'Bayesian expected shortfall forecasting incorporating the intraday range', *Journal of Financial Econometrics* **14**(1), 128–158.

Giampietro, M., Guidolin, M. & Pedio, M. (2018), 'Estimating stochastic discount factor models with hidden regimes: Applications to commodity pricing', *European Journal of Operational Research* **265**(2), 685–702.

Giot, P. & Laurent, S. (2003), 'Market risk in commodity markets: a VaR approach', *Energy Economics* **25**(5), 435–457.

Giudici, P. & Abu Hashish, I. (2020), 'A hidden markov model to detect regime changes in cryptoasset markets', *Quality and Reliability Engineering International* **36**(6), 2057–2065.

Gong, Y. & Huser, R. (2022), 'Asymmetric tail dependence modeling, with application to cryptocurrency market data', *The Annals of Applied Statistics* **16**(3), 1822–1847.

Gould, P. R. (1967), 'On the geographical interpretation of eigenvalues', *Transactions of the Institute of British Geographers* pp. 53–86.

Guesmi, K., Saadi, S., Abid, I. & Ftiti, Z. (2019), 'Portfolio diversification with virtual currency: Evidence from Bitcoin', *International Review of Financial Analysis* **63**, 431–437.

Haas, M., Mittnik, S. & Paolella, M. S. (2004), 'A new approach to Markov-switching GARCH models', *Journal of Financial Econometrics* **2**(4), 493–530.

Hamilton, J. D. (1983), 'Oil and the macroeconomy since World War II', *Journal of Political Economy* **91**(2), 228–248.

Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica* pp. 357–384.

Hamilton, J. D. (1990), 'Analysis of time series subject to changes in regime', *Journal of econometrics* **45**(1-2), 39–70.

Hansen, P. R. & Lunde, A. (2003), Does anything beat a GARCH (1,1)? A comparison based on test for superior predictive ability, *in* '2003 IEEE International Conference on Computational Intelligence for Financial Engineering, 2003. Proceedings.', IEEE, pp. 301–307.

Härdle, W. K., Okhrin, O. & Wang, W. (2015), 'Hidden markov structures for dynamic copulae', *Econometric Theory* **31**(5), 981–1015.

Hastie, T., Tibshirani, R. & Wainwright, M. (2015), *Statistical learning with sparsity: the LASSO and generalizations*, Chapman and Hall/CRC.

Henderson, B. J., Pearson, N. D. & Wang, L. (2015), 'New evidence on the financialization of commodity markets', *The Review of Financial Studies* **28**(5), 1285–1311.

Hess, D., Huang, H. & Niessen, A. (2008), 'How do commodity futures respond to macroeconomic news?', *Financial Markets and Portfolio Management* **22**(2), 127–146.

Huang, J.-J., Lee, K.-J., Liang, H. & Lin, W.-F. (2009), 'Estimating value at risk of portfolio by conditional copula-garch method', *Insurance: Mathematics and economics* **45**(3), 315–324.

Hubert, L. & Arabie, P. (1985), 'Comparing partitions', *Journal of Classification* **2**(1), 193–218.

Hyun, S., Lee, J., Kim, J.-M. & Jun, C. (2019), 'What coins lead in the cryptocurrency market: using copula and neural networks models', *Journal of Risk and Financial Management* **12**(3), 132.

Ji, Q., Bouri, E., Gupta, R. & Roubaud, D. (2018), 'Network causality structures among bitcoin and other financial assets: A directed acyclic graph approach', *The Quarterly Review of Economics and Finance* **70**, 203–213.

Ji, Q., Bouri, E., Roubaud, D. & Shahzad, S. J. H. (2018), 'Risk spillover between energy and agricultural commodity markets: A dependence-switching CoVaR-copula model', *Energy Economics* **75**, 14–27.

Ji, Q. & Fan, Y. (2016), 'Evolution of the world crude oil market integration: A graph theory analysis', *Energy Economics* **53**, 90–100.

Joe, H. (1997), *Multivariate models and multivariate dependence concepts*, CRC press.

Joe, H. & Xu, J. J. (1996), 'The estimation method of inference functions for margins for multivariate models'.

Joëts, M. (2015), 'Heterogeneous beliefs, regret, and uncertainty: The role of speculation in energy price dynamics', *European Journal of Operational Research* **247**(1), 204–215.

Jorion, P. (2006), *Value at Risk, 3rd Ed.: The New Benchmark for Managing Financial Risk*, McGraw-Hill Education.

Kang, S. H., McIver, R. & Yoon, S.-M. (2017), 'Dynamic spillover effects among crude oil, precious metal, and agricultural commodity futures markets', *Energy Economics* **62**, 19–32.

Kat, H. M. & Oomen, R. C. (2006), 'What every investor should know about commodities, Part II: Multivariate return analysis', *Alternative Investment Research Centre Working Paper* (33).

Katsiampa, P., Corbet, S. & Lucey, B. (2019), 'Volatility spillover effects in leading cryptocurrencies: A BEKK-MGARCH analysis', *Finance Research Letters* **29**, 68–74.

Kilian, L. & Vigfusson, R. J. (2011), 'Are the responses of the US economy asymmetric in energy price increases and decreases?', *Quantitative Economics* **2**(3), 419–453.

Kim, J.-M., Kim, S.-T. & Kim, S. (2020), 'On the relationship of cryptocurrency price with us stock and gold price using copula models', *Mathematics* **8**(11), 1859.

Kim, M. & Lee, S. (2016), 'Nonlinear expectile regression with application to Value-at-Risk and Expected Shortfall estimation', *Computational Statistics & Data Analysis* **94**, 1–19.

Koenker, R. (2005), *Quantile Regression*, Cambridge University Press.

Koenker, R. & Bassett, G. (1978), 'Regression quantiles', *Econometrica: Journal of the Econometric Society* **46**(1), 33–50.

Koenker, R., Chernozhukov, V., He, X. & Peng, L. (2017), 'Handbook of quantile regression'.

Kristjanpoller, W., Bouri, E. & Takaishi, T. (2020), 'Cryptocurrencies and equity funds: Evidence from an asymmetric multifractal analysis', *Physica A: Statistical Mechanics and Its Applications* **545**, 123711.

Kupiec, P. H. (1995), 'Techniques for verifying the accuracy of risk measurement models', *The Journal of Derivatives* **3**(2), 73–84.

Lambert, N. S., Pennock, D. M. & Shoham, Y. (2008), Eliciting properties of probability distributions, *in* 'Proceedings of the 9th ACM Conference on Electronic Commerce', ACM, pp. 129–138.

Lanchantin, P., Lapuyade-Lahorgue, J. & Pieczynski, W. (2011), 'Unsupervised segmentation of randomly switching data hidden with non-gaussian correlated noise', *Signal Processing* **91**(2), 163–175.

Laporta, A. G., Merlo, L. & Petrella, L. (2018), 'Selection of Value at Risk models for energy commodities', *Energy Economics* **74**, 628–643.

Lauritzen, S. L. (1996), *Graphical models*, Vol. 17, Clarendon Press.

Lauritzen, S. L. & Wermuth, N. (1989), 'Graphical models for associations between variables, some of which are qualitative and some quantitative', *The annals of Statistics* pp. 31–57.

Liu, C. (1997), 'ML estimation of the multivariate t distribution and the EM algorithm', *Journal of Multivariate Analysis* **63**(2), 296–312.

Liu, H., Song, X., Tang, Y. & Zhang, B. (2021), 'Bayesian quantile nonhomogeneous hidden Markov models', *Statistical Methods in Medical Research* **30**(1), 112–128.

Liu, X. (2016), 'Markov switching quantile autoregression', *Statistica Neerlandica* **70**(4), 356–395.

López-Martín, C., Benito Muela, S. & Arguedas, R. (2021), 'Efficiency in cryptocurrency markets: New evidence', *Eurasian Economic Review* **11**(3), 403–431.

MacDonald, I. L. & Zucchini, W. (1997), *Hidden Markov and other models for discrete-valued time series*, Vol. 110, CRC Press.

Malik, F. & Umar, Z. (2019), 'Dynamic connectedness of oil price shocks and exchange rates', *Energy Economics* **84**, 104501.

Mariana, C. D., Ekaputra, I. A. & Husodo, Z. A. (2021), 'Are bitcoin and ethereum safe-havens for stocks during the covid-19 pandemic?', *Finance research letters* **38**, 101798.

Marimoutou, V., Raggad, B. & Trabelsi, A. (2009), 'Extreme value theory and Value at Risk: Application to oil market', *Energy Economics* **31**(4), 519–530.

Marino, M. F., Tzavidis, N. & Alfò, M. (2018), 'Mixed hidden Markov quantile regression models for longitudinal data with possibly incomplete sequences', *Statistical Methods in Medical Research* **27**(7), 2231–2246.

Maruotti, A., Bulla, J., Lagona, F., Picone, M. & Martella, F. (2017), 'Dynamic mixtures of factor analyzers to characterize multivariate air pollutant exposures', *The Annals of Applied Statistics* **11**(3), 1617–1648.

Maruotti, A., Petrella, L. & Sposito, L. (2021), 'Hidden semi-Markov-switching quantile regression for time series', *Computational Statistics & Data Analysis* **159**, 107208.

Maruotti, A., Punzo, A. & Bagnato, L. (2019), 'Hidden Markov and semi-Markov models with multivariate leptokurtic-normal components for robust modeling of daily returns series', *Journal of Financial Econometrics* **17**(1), 91–117.

Marvasti, A. & Lamberte, A. (2016), 'Commodity price volatility under regulatory changes and disaster', *Journal of Empirical Finance* **38**, 355–361.

Masala, G. (2021), 'Backtesting energy portfolio with copula dependence structure', *Energy Systems* **12**(2), 393–410.

Mensi, W., Hammoudeh, S., Shahzad, S. J. H. & Shahbaz, M. (2017), 'Modeling systemic risk and dependence structure between oil and stock markets using a variational mode decomposition-based copula method', *Journal of Banking & Finance* **75**, 258–279.

Mergner, S. & Bulla, J. (2008), 'Time-varying beta risk of Pan-European industry portfolios: A comparison of alternative modeling techniques', *The European Journal of Finance* **14**(8), 771–802.

Merlo, L., Maruotti, A., Petrella, L. & Punzo, A. (2022), 'Quantile hidden semi-Markov models for multivariate time series', *Statistics and Computing* **32**(4), 1–22.

Merlo, L., Petrella, L. & Raponi, V. (2021), 'Forecasting VaR and ES using a joint quantile regression and its implications in portfolio allocation', *Journal of Banking & Finance* p. 106248.

Myint, L. L. & El-Halwagi, M. M. (2009), 'Process analysis and optimization of biodiesel production from soybean oil', *Clean Technologies and Environmental Policy* **11**(3), 263–276.

Naeem, M. A., Bouri, E., Peng, Z., Shahzad, S. J. H. & Vo, X. V. (2021), 'Asymmetric efficiency of cryptocurrencies during COVID19', *Physica A: Statistical Mechanics and its Applications* **565**, 125562.

Naeem, M., Umar, Z., Ahmed, S. & Ferrouhi, E. M. (2020), 'Dynamic dependence between ETFs and crude oil prices by using EGARCH-Copula approach', *Physica A: Statistical Mechanics and its Applications* **557**, 124885.

Nakamoto, S. (2008), 'Bitcoin: A peer-to-peer electronic cash system', *Decentralized Business Review* p. 21260.

Nazlioglu, S. (2011), 'World oil and agricultural commodity prices: Evidence from nonlinear causality', *Energy Policy* **39**(5), 2935–2943.

Nazlioglu, S., Erdem, C. & Soytas, U. (2013), 'Volatility spillover between oil and agricultural commodity markets', *Energy Economics* **36**, 658–665.

Nazlioglu, S. & Soytas, U. (2011), 'World oil prices and agricultural commodity prices: evidence from an emerging market', *Energy Economics* **33**(3), 488–496.

Newey, W. K. & Powell, J. L. (1987), 'Asymmetric least squares estimation and testing', *Econometrica: Journal of the Econometric Society* pp. 819–847.

Newman, M. E. (2006), 'Modularity and community structure in networks', *Proceedings of the National Academy of Sciences* **103**(23), 8577–8582.

Nguyen, Q. N., Aboura, S., Chevallier, J., Zhang, L. & Zhu, B. (2020), 'Local Gaussian correlations in financial and commodity markets', *European Journal of Operational Research* **285**(1), 306–323.

Nigri, A., Barbi, E. & Levantesi, S. (2022), 'The relationship between longevity and lifespan variation', *Statistical Methods & Applications* **31**(3), 481–493.

Nystrup, P., Hansen, B. W., Madsen, H. & Lindström, E. (2015), 'Regime-based versus static asset allocation: Letting the data speak', *The Journal of Portfolio Management* **42**(1), 103–109.

Nystrup, P., Madsen, H. & Lindström, E. (2017), 'Long memory of financial time series and hidden Markov models with time-varying parameters', *Journal of Forecasting* **36**(8), 989–1002.

Ötting, M., Langrock, R. & Maruotti, A. (2021), 'A copula-based multivariate hidden markov model for modelling momentum in football', *AStA Advances in Statistical Analysis* pp. 1–19.

Pennoni, F., Bartolucci, F., Forte, G. & Ametrano, F. (2022), 'Exploring the dependencies among main cryptocurrency log-returns: A hidden markov model', *Economic Notes* **51**(1), e12193.

Pilipovic, D. (2007), *Energy risk: Valuing and managing energy derivatives*, McGraw Hill Professional.

Pindyck, R. S. & Rotemberg, J. J. (1990), 'The excess co-movement of commodity prices', *The Economic Journal* **100**(403), 1173–1189.

Ramiah, V., Wallace, D., Veron, J. F., Reddy, K. & Elliott, R. (2019), 'The effects of recent terrorist attacks on risk and return in commodity markets', *Energy Economics* **77**, 13–22.

Reboredo, J. C. (2015), 'Is there dependence and systemic risk between oil and renewable energy stock prices?', *Energy Economics* **48**, 32–45.

Rehman, M. U., Shahzad, S. J. H., Uddin, G. S. & Hedström, A. (2018), 'Precious metal returns and oil shocks: A time varying connectedness approach', *Resources Policy* **58**, 77–89.

Rezitis, A. N. (2015), 'The relationship between agricultural commodity prices, crude oil prices and US dollar exchange rates: A panel VAR approach and causality analysis', *International Review of Applied Economics* **29**(3), 403–434.

Roache, S. K. (2008), 'Commodities and the market price of risk', *IMF Working Paper No. 08/221* .

Sadorsky, P. (2006), 'Modeling and forecasting petroleum futures volatility', *Energy Economics* **28**(4), 467–488.

Schwarz, G. et al. (1978), 'Estimating the dimension of a model', *The Annals of Statistics* **6**(2), 461–464.

Shahzad, S. J. H., Bouri, E., Ahmad, T. & Naeem, M. A. (2022), 'Extreme tail network analysis of cryptocurrencies and trading strategies', *Finance Research Letters* **44**, 102106.

Shen, Y., Shi, X. & Variam, H. M. P. (2018), 'Risk transmission mechanism between energy markets: A VAR for VaR approach', *Energy Economics* **75**, 377–388.

Sklar, M. (1959), 'Fonctions de repartition an dimensions et leurs marges', *Publ. inst. statist. univ. Paris* **8**, 229–231.

Smales, L. A. (2014), 'News sentiment in the gold futures market', *Journal of Banking & Finance* **49**, 275–286.

Sobotka, F. & Kneib, T. (2012), 'Geoadditive expectile regression', *Computational Statistics & Data Analysis* **56**(4), 755–767.

Spiegel, E., Kneib, T. & Otto-Sobotka, F. (2020), 'Spatio-temporal expectile regression models', *Statistical Modelling* **20**(4), 386–409.

Spierdijk, L. & Umar, Z. (2013), 'Are commodity futures a good hedge against inflation?'.

Tang, K. & Xiong, W. (2012), 'Index investment and the financialization of commodities', *Financial Analysts Journal* **68**(6), 54–74.

Taylor, J. W. (2008), 'Estimating Value at Risk and Expected Shortfall using expectiles', *Journal of Financial Econometrics* **6**(2), 231–252.

Taylor, J. W. (2019), 'Forecasting Value at Risk and Expected Shortfall using a semiparametric approach based on the asymmetric Laplace distribution', *Journal of Business & Economic Statistics* **37**(1), 121–133.

Basel Committee on Banking Supervision (2016), Minimum capital requirements for market risk, Technical report.

Tibshirani, R. (1996), 'Regression shrinkage and selection via the LASSO', *Journal of the Royal Statistical Society: Series B (Methodological)* **58**(1), 267–288.

Tiwari, A. K., Mukherjee, Z., Gupta, R. & Balcilar, M. (2019), 'A wavelet analysis of the relationship between oil and natural gas prices', *Resources Policy* **60**, 118–124.

Tiwari, A. K., Umar, Z. & Alqahtani, F. (2021), 'Existence of long memory in crude oil and petroleum products: Generalised Hurst exponent approach', *Research in International Business and Finance* **57**, 101403.

Tyner, W. E. (2010), 'The integration of energy and agricultural markets', *Agricultural Economics* **41**, 193–201.

Tzavidis, N., Salvati, N., Schmid, T., Flouri, E. & Midouhas, E. (2016), 'Longitudinal analysis of the strengths and difficulties questionnaire scores of the Millennium Cohort Study children in England using M-quantile random-effects regression', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **179**(2), 427–452.

Umar, Z., Gubareva, M. & Teplova, T. (2021), 'The impact of Covid-19 on commodity markets volatility: Analyzing time-frequency relations between commodity prices and Coronavirus panic levels', *Resources Policy* **73**, 102164.

Umar, Z., Jareño, F. & Escribano, A. (2021), 'Agricultural commodity markets and oil prices: An analysis of the dynamic return and volatility connectedness', *Resources Policy* **73**, 102147.

Umar, Z., Jareño, F. & Escribano, A. (2022), 'Dynamic return and volatility connectedness for dominant agricultural commodity markets during the COVID-19 pandemic era', *Applied Economics* **54**(9), 1030–1054.

Umar, Z. & Spierdijk, L. (2011), Are commodities a good hedge against inflation? A comparative approach, Technical report, Netspar discussion paper, <<rug.nl>>.

UNCTAD (2019), UN conference on Trade and Development (UNCTAD), Technical report.

Visser, I., Raijmakers, M. E. & Molenaar, P. C. (2000), 'Confidence intervals for hidden Markov model parameters', *British Journal of Mathematical and Statistical Psychology* **53**(2), 317–327.

Waldmann, E., Sobotka, F. & Kneib, T. (2013), 'Bayesian geoadditive expectile regression', *arXiv preprint arXiv:1312.5054* .

Waldmann, E., Sobotka, F. & Kneib, T. (2017), 'Bayesian regularisation in geoadditive expectile regression', *Statistics and Computing* **27**(6), 1539–1553.

Welch, L. R. (2003), 'Hidden Markov models and the Baum-Welch algorithm', *IEEE Information Theory Society Newsletter* **53**(4), 10–13.

White, H., Kim, T.-H. & Manganelli, S. (2015), 'VAR for VaR: Measuring tail dependence using multivariate regression quantiles', *Journal of Econometrics* **187**(1), 169–188.

Whittaker, J. (2009), *Graphical models in applied multivariate statistics*, Wiley Publishing.

Wilhelmsson, A. (2006), 'GARCH forecasting performance under different distribution assumptions', *Journal of Forecasting* **25**(8), 561–578.

Xiong, J., Liu, Q. & Zhao, L. (2020), 'A new method to verify bitcoin bubbles: Based on the production cost', *The North American Journal of Economics and Finance* **51**, 101095.

Yarovaya, L., Brzeszczyński, J. & Lau, C. K. M. (2016), 'Intra-and inter-regional return and volatility spillovers across emerging and developed markets: Evidence from stock indices and stock index futures', *International Review of Financial Analysis* **43**, 96–114.

Ye, W., Zhu, Y., Wu, Y. & Miao, B. (2016), 'Markov regime-switching quantile regression models and financial contagion detection', *Insurance: Mathematics and Economics* **67**, 21–26.

Yi, S., Xu, Z. & Wang, G.-J. (2018), 'Volatility connectedness in the cryptocurrency market: Is Bitcoin a dominant cryptocurrency?', *International Review of Financial Analysis* **60**, 98–114.

Yousaf, I. & Ali, S. (2021), 'Linkages between stock and cryptocurrency markets during the covid-19 outbreak: An intraday analysis', *The Singapore Economic Review* pp. 1–20.

Yu, K. & Moyeed, R. A. (2001), 'Bayesian quantile regression', *Statistics & Probability Letters* **54**(4), 437–447.

Yuan, M. & Lin, Y. (2007), 'Model selection and estimation in the Gaussian graphical model', *Biometrika* **94**(1), 19–35.

Zaremba, A., Umar, Z. & Mikutowski, M. (2019), 'Inflation hedging with commodities: A wavelet analysis of seven centuries worth of data', *Economics Letters* **181**, 90–94.

Zaremba, A., Umar, Z. & Mikutowski, M. (2021), 'Commodity financialisation and price co-movement: Lessons from two centuries of evidence', *Finance Research Letters* **38**, 101492.

Zaremba, A., Umar, Z., Mikutowski, M. et al. (2021), 'Practical Applications of Inflation Hedging in the Long Run: Perspectives from Seven Centuries of Commodity Prices', *The Journal of Alternative Investments* **24**(Supplement1), 1–5.

Zhang, D. & Broadstock, D. C. (2018), 'Global financial crisis and rising connectedness in the international commodity markets', *International Review of Financial Analysis* p. 101239.

Zhang, D. & Broadstock, D. C. (2020), 'Global financial crisis and rising connectedness in the international commodity markets', *International Review of Financial Analysis* **68**, 101239.

Zhang, Y.-J., Bouri, E., Gupta, R. & Ma, S.-J. (2021), 'Risk spillover between bitcoin and conventional financial markets: An expectile-based approach', *The North American Journal of Economics and Finance* **55**, 101296.

Ziegel, J. F. (2016), 'Coherence and elicitability', *Mathematical Finance* **26**(4), 901–918.

Zucchini, W., MacDonald, I. L. & Langrock, R. (2016), *Hidden Markov models for time series: an introduction using R*, Chapman and Hall/CRC.