NTNU
Handelshøyskolen

Stine Marie Reese
Margrethe Ringkjøb Skjelstad

Portfolio Stress Testing Applied to Commodity Futures

Portefølje stresstesting anvendt på råvarefutures

MASTEROPPGAVE - Økonomi og administrasjon/siviløkonom
Trondheim, dato: 24.05.2018

Hovedprofil: Finansiering og investering

Veileder: Florentina Paraschiv

Preface

This master thesis is the final product of the master degree in Economics and Business Administration at NTNU Business school. It is a master thesis in the field of Finance and investment spring 2018. This work has given us valuable insights in the portfolio risk management and developed our abilities in econometric modelling.

We want to thank our supervisor Professor Florentina Paraschiv for her constructive feedback and for supporting us throughout the semester. We would also like to thank Denis Becker and Valeriy Kunst for useful input and comments.

The authors take full responsibility for the content of this thesis.

Trondheim

24.05.2018

Stine Marie Reese

Margrethe R. Skjelstad

Abstract

In this thesis, we construct a portfolio of commodity futures, which mimics the Dow Jones Commodity Index, and perform an extensive stress testing exercise with a focus on hybrid scenarios. Limitations to the risk management practices became clear during the recent financial crisis, and the increased volume of investments in commodity markets over the last decades underline the importance of a more thorough framework for stress testing of related portfolios. Our study is the first to show the marginal impact of the model choice for portfolio components versus the marginal role of tail dependency in stress testing exercises to assess the portfolio risk profile. We model the distribution of portfolio components with an asymmetric GARCH model combined with Extreme Value Theory for extreme tails. We then implement a copula function to model the time-varying joint dependency structure. Our study reveals that indeed, for an accurate stress test, a special attention should be given to the tail risk in individual commodity returns as well as to tail correlations. Furthermore, we find that the parameter risk in the model for the individual portfolio components impact the portfolio profit and loss profile most in stress testing exercises. Finally yet importantly, in line with Basel III, the study highlights the importance of using forward-looking hybrid and hypothetical scenarios over historical scenarios, for a comprehensive stress testing.

Sammendrag

I denne masteroppgaven bruker vi råvarefutures til å konstruere en portefølje som etterligner Dow Jones Commodity Index. Vi gjennomfører en omfattende stresstestingsøvelse på denne porteføljen, med fokus på hybride scenarioer. Begrensningene i risikostyringpraksis ble tydelige under den siste finanskrisen, og det økte volumet av råvareinvesteringer over de siste tiår understreker viktigheten av et grundigere rammeverk for stresstesting av porteføljer. Vår studie er den første til å vise den marginale påvirkningen fra valg av modell for komponentene i en portefølje versus den marginale påvirkningen fra haleavhengighet i stresstesting øvelser for å undersøke porteføljens risikoprofil. Vi modellerer fordelingen til komponentene i porteføljen med en asymmetrisk GARCH-modell kombinert med Extreme Value Theory for ekstreme haler. Vi implementerer så en copula-funksjon for å modellere den felles tidsvarierende avhengighetsstrukturen. Vår studie viser at for en korrekt stresstestingsøvelse bør et spesielt fokus rettes mot halerisiko for de individuelle råvareavkastningene, samt til halekorrelasjoner. Videre finner vi at parameterrisiko i modellen for de individuelle komponentene påvirker porteføljens gevinst- og tapsfordeling mest i stresstestingsøvelser. Sist, men ikke minst, fremhever studien viktigheten av å bruke framtidsrettede hybride og hypotetiske scenarioer, snarere enn historiske scenarioer, for en omfattende stresstesting. Dette er sammenfallende med Basel III rammeverket.

Contents

List of Figures

List of Tables

1 Introduction

Financial investments in commodities have grown rapidly over the last decades. It is estimated that the total value of commodity index related instruments purchased by institutional investors increased from $15 billion in 2003 to $200 billion in mid-2008 (Commodity Futures Trading Commission, 2008). Research such as Silvennoinen and Thorp (2013) and Daskalaki and Skiadopoulos (2011) show increased integration of commodity and financial markets, with higher correlation, especially in bearish times (Cheung and Miu, 2010). Adams and Glück (2015) show that the risk spillovers to commodities from the financial crisis continue to be present today. In addition, volatility in commodity markets increased during the past decade (Tang and Xiong, 2010, Basak and Pavlova, 2016). These changes in commodity characteristics is often referred to as the financialization of commodity markets (Cheng and Xiong, 2014). This leads to the need for an instrument to measure and manage the associated risks in commodity futures investments.

A common tool for risk management is stress testing. European Banking Authority (2017, p.28) point out in their new guidelines under development that ”Institutions should ensure that the scenario analysis is a core part of their stress testing programme”. Implementing stress testing is now mandatory for banks, due to regulations from Basel III formed in the post crisis environment (Basel Committee on Banking Supervision, 2009). However, this stress testing framework does not provide an understanding of which underlying factors to focus on for the stress testing to be optimal. Aepli (2011) suggests a stress testing framework with focus on realistic models for individual risk factor distributions and dynamic dependence structure, with a special focus on tail risk, which we will adapt in this thesis.

In this thesis, we focus on stress testing of a portfolio of commodity futures. The existing literature on stress testing of commodity portfolios is scarce, despite the massive investment growth in commodities. We update the analysis in Paraschiv, Mudry and Andries (2015), keeping the same procedure for constructing the stress portfolio as the original study. However, we innovated in several directions. We extended the data set by including several new shocks, among these the oil price drop in 2014. Secondly, we enriched the spectrum of stress testing scenarios, focusing more on the forward-looking ones. Paraschiv, Mudry and Andries (2015) limit their study to show the effects of a reoccurring financial crisis on the portfolio. Our study shows the importance of combining historical estimations with more flexible forward-looking scenario construction. Furthermore, this is the first study in the literature where we disentangle the effect of individual model components on the portfolio profit and loss.

For our portfolio construction we mimic the Dow Jones Commodity Index, a well known and much traded index. The DJCI is a broad commodity index consisting of 24 commodities in three major sectors: energy, metals and agriculture & livestocks. The weights are based on traded volume, ensuring a liquid index. The index restrains the maximum and minimum weight each commodity and sector can constitute (S&P Dow Jones Indices, 2017), providing diversification and continuity for the potential investor, making it a good proxy for our study.

We will apply modern econometric techniques in our modelling. For the marginal distributions of commodity returns, we use an asymmetric AR-GARCH process, and model the tails by applying Extreme Value Theory. For the joint dependency we use a copula function. Finally, we simulate the profit and loss distributions for the portfolio under different scenarios in a stress testing framework.

Our results are twofold. First we find that the simulated profit and loss distribution of the portfolio is highly sensitive to the choice of modelling approach for the marginal distribution of portfolio components. A close focus on estimating the marginal distributions, especially for the tails, is of importance for the stress testing purpose. In addition, we found that dependence modelling is less crucial for the stress testing to be informative. Secondly, we find the construction of hybrid scenarios to be a relevant tool to combine both historical information and the flexibility of forward looking approaches in line with the requirements from Basel III (Basel Committee on Banking Supervision, 2009).

1.1 Aim and structure

The remainder of this thesis is structured as follows. Chapter two is an overview of the most relevant literature for our study. In chapter three we will provide an introduction of our data, focusing on the characteristics of the portfolio. In chapter four we will move on to the theoretical background of the different methodologies applied. Chapter five provides the implementation of the methodology for our data set. Finally, in chapter six we will explain and apply stress testing and display our analysis.

For the accomplishment of this thesis we base the calculations on a protocol developed by The Matworks Inc. We modify the code to fit our scope and data. The code is freely available online.

2 Literature review

In this chapter we will discuss three literature streams relevant for our thesis. First, recent literature on stress testing is presented, before an overview of the economics of commodity markets is provided. Finally, the financialization of commodities is investigated.

2.1 Stress testing

Stress testing can according to Lopez (2005) be defined as a risk management tool used to evaluate the potential impact on portfolio values of unlikely, although plausible events or movements in a set of financial variables. The recent financial crisis led the attention of banks and authorities to the insufficient methods of risk management, and the need for more accurate stress testing became obvious, since financial institutions were not prepared to deal with the crisis. One main concern was that the scenario selection and simulation was carried out by separate units for each business line and for particular risk types (Basel Committee on Banking Supervision, 2009). This indicates that the stress testing was isolated and did not provide a complete picture on the firm level.

Seemingly the most recent development in methodology for stress testing of portfolios is the use of Extreme Value Theory (EVT) and copulas as input to the analysis. EVT was first introduced in Embrechts, Mikosch and Klüppelberg (1997) to better model the tail distribution of risk factors. Extreme Value Theory focuses on shaping the tails rather than the whole distribution of returns, providing potentially better estimates of risk for financial portfolios. McNeil, Frey and Embrechts (2015) suggest using a combination of GARCH and EVT where the GARCH residuals are used as input to EVT. This methodology is somewhat adapted in recent literature, with the largest proportion of new studies focusing on stock markets or single commodities (Ghorbel and Souilmi, 2014, Liu, 2011, Wang et al., 2010, Aepli, 2011).

Koliai (2016) analyses existing risk models for stress testing purposes. The study presents a semi-parametric copula-GARCH risk model for equity indices, exchange rates and commodity prices, to perform stress testing on hypothetical portfolios, where the marginal distributions of returns are specified using EVT. Findings are that different risk models produce significantly different results in terms of corresponding stress scenarios and impact on the portfolios. Further on, the proposed model in the paper ensures the credibility of the stress scenario and the usefulness of the stress testing results, because of considering flexible and consistent specifications.

Paraschiv, Mudry and Andries (2015) is to our knowledge the only example of stress testing with the GARCH-EVT-Copula methodology for a broad portfolio of commodities. They find that implementing Extreme Value Theory and a t copula for a commodity portfolio improves the capture of potential losses. They also point out the importance of using forward-looking scenarios to enable the simulations of extreme quantiles, providing a better understanding of risk.

Stress tests can be conducted with several methodologies. One can firstly differentiate between univariate and multivariate stress tests. Univariate stress tests aim to identify the isolated influence of stressing or shocking one single risk factor of a portfolio (Aepli, 2011, p. 4). This makes the univariate stress tests simple to apply, but very limited, since they do not take dependencies between risk factors in a portfolio into account, like a multivariate stress test would. Basel Committee on Banking Supervision (2009) classifies stress test methodologies for financial institutions.

Further on, one can separate between different scenarios when running stress tests. The need for hypothetical scenarios were highlighted after the crisis, since risk managers mostly performed historical stress testing under Basel II (Basel Committee on Banking Supervision, 2006). The European Banking Authority (2017, p.28) point out that ”the design of the stress test scenarios should not only be based on historical events, but should also consider hypothetical scenarios, also based on non-historical events”. Forward-looking scenarios are now required for European banks according to Basel Committee on Banking Supervision (2009). Aepli (2011) presents an extensive framework for complex stress testing for portfolios of futures. He develops a foundation that is in agreement with the regulations from Basel III formed in the post crisis environment. We will give a short description of the scenario categories, following Aepli (2011, p. 5-7).

Historical Scenario

Historical scenarios are based on actual, realised data steaming from a historical episode of financial stress. This makes them realistic and easy to access. The profit and loss distribution in the historical scenario is simply given by the realised empirical distributions. Lopez (2005) points out that historical scenarios are developed more fully than other scenarios since they reflect an actual stressed market environment that can be studied in great detail, therefore requiring fewer judgements by risk managers.

One major drawback with historical scenarios is the assumption that passed financial crises will reoccur with the same consequences on portfolio losses. This makes them unable to capture risks linked to new products that may have significant impact on the outcome of a crisis. The worst observed loss in the past might not reflect the worst possible outcome in the future. This drawback was proven to be essential in the financial crisis of 2007 and resulted in the underestimation of the risk level and interaction between risks (Basel Committee on Banking Supervision, 2009, p. 5).

Another drawback in historical scenarios is the sample size. Due to the limited number of observations, computing risk metrics in the higher confidence levels becomes problematic. This is a considerable drawback as the most extreme losses are of great interest in stress testing exercises.

Hypothetical Scenario

Hypothetical scenarios are, unlike the historical scenarios, forward looking. Scenarios can be constructed in multiple ways, for example by shocking model parameters arbitrarily, based on own experiences of market movements. Hypothetical scenarios have the advantage of being more flexible and forward looking, making them more informative if conducted correctly. More focus on hypothetical stress testing scenarios allows the institution to be both well prepared for potential extreme unexpected outcomes, and lay the foundation to overcome these potential losses.

An extensive analysis has to be in place before constructing hypothetical scenarios, which can be both time consuming and difficult. Basel Committee on Banking Supervision (2009, p. 5) point out that banks had implemented hypothetical scenarios prior to the crisis, but it was difficult for risk managers to obtain the support of the senior management, since the scenarios were extreme or innovative, and often were considered as implausible. Extremes that have not yet been experienced are often difficult to imagine and to be taken seriously into consideration by risk managers.

Hybrid Scenario

Hybrid scenarios combine the knowledge found in historical scenarios with the flexibility of the more hypothetical scenarios, making them a suitable alternative in stress testing. Hybrid scenarios are also easier to implement than more extensive forward-looking scenarios, as they are anchored in actual experienced market conditions. Hybrid scenarios are constructed by using historical data during times of financial distress to calibrate the process of risk factor evolution, but allow extrapolation beyond experienced events.

Even though hybrid scenarios allow the construction of new possible scenarios, they are still somewhat backward looking in the sense that they do not fully explore the risk of shifting market conditions or risk associated with new products. However, Lopez (2005) points out that risk managers always face a trade-off between scenario realism and comprehensibility; that is, more fully developed scenarios generate results that are more difficult to interpret. The benefits from implementing hybrid scenarios should not be neglected as they balance this trade-off.

2.2 Economics of Commodity markets

Commodities are a broad diversified non-homogeneous asset group. Commodities can roughly be divided into energy, metals (industrial and precious) and livestocks & agriculture (grains and soft). Each commodity is driven by specific supply and demand conditions, so their risk characteristics differ from each other (Tyner, 2010, He, Wang and Lai, 2010). However, all commodities are affected somewhat by global activity since most humans consume commodities in their everyday life, whether it is grains for food, natural gas for heating or copper for electronics (Delle Chiaie, Ferrara and Giannone, 2017, Pindyck, 2004). The supply and demand of commodities are mainly affected by liquidity, weather and natural disasters, geopolitics and global activity. Typically, commodities also show, like most financial assets, fat tails and skewness (Bhardwaj, Gorton and Rouwenhorst, 2015). Another important factor affecting prices and risk of commodities is the cost of storage. While most commodities can be stored, the costs attached to this are varying. Some commodities, like electricity, cannot be stored. See for example Pindyck (2004) for a deeper discussion.

Delle Chiaie, Ferrara and Giannone (2017) study the co-integration of commodities of different sectors. They examine the drivers of commodity prices, and bring evidence to the increased impact of the global activity on the price developments of commodities, especially for oil. Their price model reveals that global activity explains much of the fluctuations in commodity prices during episodes linked to changes in global demand conditions (Delle Chiaie, Ferrara and Giannone, 2017, p. 4), while price fluctuations during episodes with changes in supply mostly are explained by local commodity market conditions.

Despite the differences in price drivers among commodities one can observe a common increase in prices since the beginning of the twenty-first century. Furthermore, one can observe an identical drop in prices during the financial crisis. Cheng and Xiong (2014) call it a common boom and bust cycle. They point out two main reasons for the simultaneous price boom; the first is the increased demand for oil and other commodities from emerging markets, one example here is China (see also Kilian and Zhou, 2018). The second reason is the entry of new technology transforming agricultural commodities into substitutions for oil and energy commodities.

2.3 Financialization of commodity markets

The risk associated with weather, storage etc. led to the rise of commodity indices in the early 2000s, providing a hedge opportunity for commodity producers. The commodity indices also allowed for easy-access investments in a diversified basket of commodities for all types of investors, effectively reducing the risks of each individual commodity. This increased investment in commodities is often referred to as the financialization of commodity markets (Tang and Xiong, 2010). The main derivative for financial investors in the commodity market is the future contract. The use of futures allows the investors to expose themselves to the return and risk of commodities without storing the physical product.

Investing in commodity indices is now common for big institutions such as pension funds and banks. The diversification effect of including commodities in a stock and bond portfolio is driven by the fact that commodity returns correlation with stocks and bonds have historically been low and there is still a diversification potential for portfolios found in recent literature (Bhardwaj, Gorton and Rouwenhorst, 2015, Gorton and Rouwenhorst, 2006, Öztek and Öcal, 2017). Commodities are known to be volatile, but have historically been less volatile than the stock market (Bhardwaj, Gorton and Rouwenhorst, 2015). Stoll and Whaley (2015) argue that the commodities' role as inflation hedge is one important reason for the increased popularity of the commodity indices as a risk management tool for institutions.

Basak and Pavlova (2016) examine how the presence of institutional investors may affect the commodity prices and their dynamics. They find that futures prices, volatility and correlations of all commodities rise, especially for index commodities, in the presence of financial or investing institutions. They also find that demand and supply shocks to index commodities get transmitted to the price of other commodities.

Tang and Xiong (2010) show that the increasing presence of index investors has exposed commodity prices to market-wide shocks, such as shocks to the world equity index, the US dollar exchange rate, and shocks to other commodities, such as oil. They extend their research by arguing that the financialization with investors trading and linking different commodity prices is the main reason of the overall increase in commodity correlations. Another important driver of correlation is the linked demand for energy and agricultural commodities used for energy purposes, like biofuel from corn, as the demand for green energy has spiked (Tang and Xiong, 2012). Bhardwaj, Gorton and Rouwenhorst (2015) show that correlation among commodities also increase during periods of intensive financial stress. They also show that correlations with other assets, such as stocks and bonds, temporarily increase during times of extreme market turmoil. Pointing to the financialization of commodity markets they highlight an extreme growth in commodity investments, despite the stable composition of traders. This urges the need to stress test commodity portfolios to prevent extreme, but plausible losses.

3 Data selection and description

In this chapter we will present our data set and show the construction of our test portfolio. We will then analyse the descriptive statistics for the commodity returns and do preliminary tests.

3.1 Commodity indices

Commodity indices have become quite popular in the last decades (see Section 2.3) and several commodity indices exist. Two big commodity indices are the S&P Goldman Sachs Commodity Index (S&P GSCI) and the Dow Jones Commodity Index (DJCI). The S&P GSCI consists of 24 commodities and the weights are based on trading volume. It is therefore often seen as a benchmark for investment performance in commodities. The trading volume in energy commodities is higher than any other commodity sector so this index is heavily based in energy (60% of the total weight in 2017 (S&P Dow Jones Indices, 2018)). To get a more balanced portfolio across different commodity sectors the DJCI will be the focus in this thesis. It consists of 24 commodities divided into three major sectors; metals, energy and livestock & agriculture. The weights are based on the total volume traded, but unlike the S&P GSCI, the DJCI has constraints on total weight allocated in each sector and commodity. By not allowing any of the three sectors to obtain more than 35% of the weight, and no single commodity to constitute less then 2% or more than 17% of the total index, the DJCI becomes diversified. These restrictions also provide continuity and high liquidity for potential investors. The weights are rebalanced annually. See S&P Dow Jones Indices (2017) for a detailed methodology.

3.1.1 Risk factor selection

To select the risk factors (commodities) for our analysis we apply the method introduced in Mudry (2013). We take the ten commodities with the largest weights for 2017 in the DJCI, and form our test portfolio. This is done to get a more time efficient portfolio and to make the analysis more practical. The ten commodities add up to 76% of the DJCI, providing a good proxy for the movements of the entire index. To form our test portfolio we scale up the weights, proportionally to 100%. The weights of the ten commodities can be seen in Table 3.1. Our selection leaves us with three portfolio components in energy, three in metals and four in agriculture & livestocks.

Commodity	Weight in index	Weight in test portfolio
Wheat	3.208%	4.217%
Corn	7.114%	9.351%
Soybeans	11.91%	15.656%
Live Cattle	2.730%	3.589%
Copper	10.39%	13.658%
Gold	10.23%	13.447%
Aluminium	4.613%	6.064%
WTI	9.718%	12.774%
Brent	8.849%	11.632%
Natural Gas	7.313%	9.613%
Sum	76.075%	100%

Table 3.1: Portfolio weights scaled up from the weights in DJCI 2017. Source: S&P Dow Jones Indices (2016).

3.2 Data Extraction

We extracted daily data from 1996 - 2017 from Thomson Reuters Eikon for continuous series of futures with approximately one year to maturity for the ten selected commodities. This leaves us with 5741 observations for each commodity. Details about the data extraction are found in Table 3.2.

When working with futures the concept of rolling over is essential. This is done to adjust for the increased volatility as a contract gets closer to maturity, and also to adjust for the sudden differences in price when going from one contract to the next. There are many different methods to adjust the time series dependent on the purpose etc. As for our data, the rolling over was done by Eikon. Their methodology can be found in Thomson Reuters (2012). Shortly explained; for monthly futures data roll over is done by jumping to the nearest future contract with a switch over following the last trading day. In other words, they use the nearest contract month to form the first values of the continuous series, and when the contract expires the next point of data is the next one year to maturity contract. They do not adjust for price differentials when adjusting the data, but we found this methodology to be sufficient for our analysis especially as our futures have one year to maturity.

Commodity	Ticker	Classification	Units	Price Quote	Exchange
WTI	CL	Energy	1000 barrels	USD/barrel	NYMEX
Brent	LCO	Energy	1000 barrels	USD/barrel	ICE
Natural Gas	NG	Energy	10.000 Million BTU	USD/MMBtu	NYMEX
Corn	C	Grains	5000 bushels	US cent/bushels	CBoT
Wheat	W	Grains	5000 bushels	US cent/bushels	CBoT
Soybeans	ISF	Grains	5000 bushels	US cent/bushels	CBoT
Live Cattle	LC	Livestock	40.000 Pounds	US cent/pound	CME
Gold	GC	Precious Metal	100 Ounces	USD/Troy oz	COMEX
Aluminium	MAL	Industrial Metal	25 Metric Tonne	USD/MT	LME
Copper	HG	Industrial Metal	25.000 pounce	USD/pounce	COMEX

Table 3.2: Data extraction details. Source: Thomson Reuters Eikon

Figure 3.1 shows the historical price movements for the commodities, measured in a relative index value. We observe a co-movement of many commodities. Especially in the 2000s the commodity markets experienced a uniform rise in prices until the financial crisis (see Section 2.2 for discussion).

We observe several structural breaks across the commodities during our time period. Especially the financial crisis in 2007-2009 heavily affected commodity markets. Delle Chiaie, Ferrara and Giannone (2017) bring evidence supporting that global activity has clear implications for the commodity markets. Their analysis shows that since the year 2000 the price drivers of oil have fundamentally changed, and during the time of the financial crisis global activity strongly affected the oil price.

The acute drop in oil price in 2014 was driven by several factors, among them the increased supply of unconventional oil and a significant shift in OPEC policy (Baffes et al., 2015). What differentiates the price drop in 2014 from previous collapses in oil price is, according to Baffes et al. (2015), that the fluctuation could not be explained by a weakened demand or expansion of supply in isolation, but rather a combination of the two.

While during the financial crisis all commodity sectors were affected, the price drop in 2014 to a lesser degree showed spillover to non-energy sectors. This indicates the decoupling of oil price from other commodities in agriculture and metals. According to Erdős (2012) the co-integration of oil and natural gas ended in 2009 after an increase in shale gas production. We observe from the graph that the commodities in non-energy sectors in more recent years do not necessarily follow the oil price as closely as in the past decade, potentially affecting the dynamics of the commodity markets.

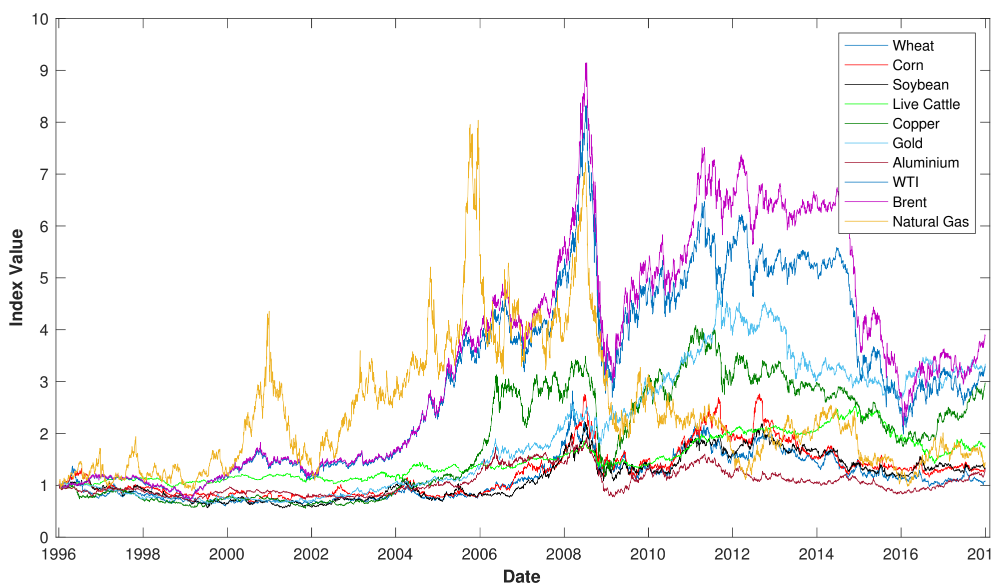

Figure 3.1: Historical daily price movements from 1996-2017 for the ten commodities in relative value.

3.3 Descriptive statistics

McNeil, Frey and Embrechts (2015, p. 117) present six stylized facts of financial returns that can be observed, especially when looking at monthly, weekly and daily data:

1. Return series are not i.i.d. although they show little serial correlation.

2. Series of absolute or squared returns show profound serial correlation.

3. Conditional expected returns are close to zero.

4. Volatility appears to vary over time.

5. Return series are leptokurtic or heavy-tailed.

6. Extreme returns appear in clusters.

Commodity	Mean	Std. Dev.	Max.	Min.	Skewness	Kurtosis	Jarque-Bera (p-value)
Wheat	0.0014%	1.529%	8.15%	-12.17%	-0.1123	6.86	3573 (0.001)
Corn	0.0046%	1.447%	9.74%	-14.48%	-0.2157	8.19	6467 (0.001)
Soybeans	0.0049%	1.364%	7.04%	-8.11%	-0.3039	6.48	2984 (0.001)
Live Cattle	0.0098%	0.722%	6.61%	-6.89%	-0.6774	12.98	24 210 (0.001)
Copper	0.019%	1.545%	11.41%	-11.26%	-0.1425	7.80	5512 (0.001)
Gold	0.021%	1.062%	8.62%	-9.87%	-0.1347	9.89	11 351 (0.001)
Aluminium	0.0046%	1.131%	6.37%	-7.60%	-0.2470	6.09	2335 (0.001)
WTI	0.0208%	1.623%	10.00%	-9.19%	-0.1655	6.33	2670 (0.001)
Brent	0.02367%	1.582%	9.26%	-9.26%	-0.1055	6.00	2162 (0.001)
Natural Gas	0.0064%	2.752%	21.64%	-31.12%	-0.0807	10.30	12 716 (0.001)

Table 3.3: Daily descriptive statistics of commodity returns for years 1996-2017.

Table 3.3 shows the daily descriptive statistics of our time series. Referring to the stylized facts above, we can see that they apply for our data. All ten commodity returns show positive mean close to zero, in line with stylized fact 3. Natural gas is the most volatile commodity, while live cattle is the least volatile. The commodity returns exhibit negative skewness and excess kurtosis, which implies they have significant fat tails. The Jarqe-Bera test rejects the null hypothesis of a normal distribution for all the commodity returns.

Figure 3.2 shows the sample autocorrelation plot of the returns and squared returns for WTI, as well as the daily logarithmic returns and a quantile-quantile plot. Plots for the other commodities can be viewed in Appendix A1, A2, A3, A4, which show similar results. We can now also see graphically that the returns show autocorrelation, in line with stylized fact 2. The commodity returns furthermore show volatility clustering and substantiate fact 4 and 6. The QQ-plot in Figure 3.2 clearly substantiates the rejection of the normal distribution by the Jarque-Bera test.

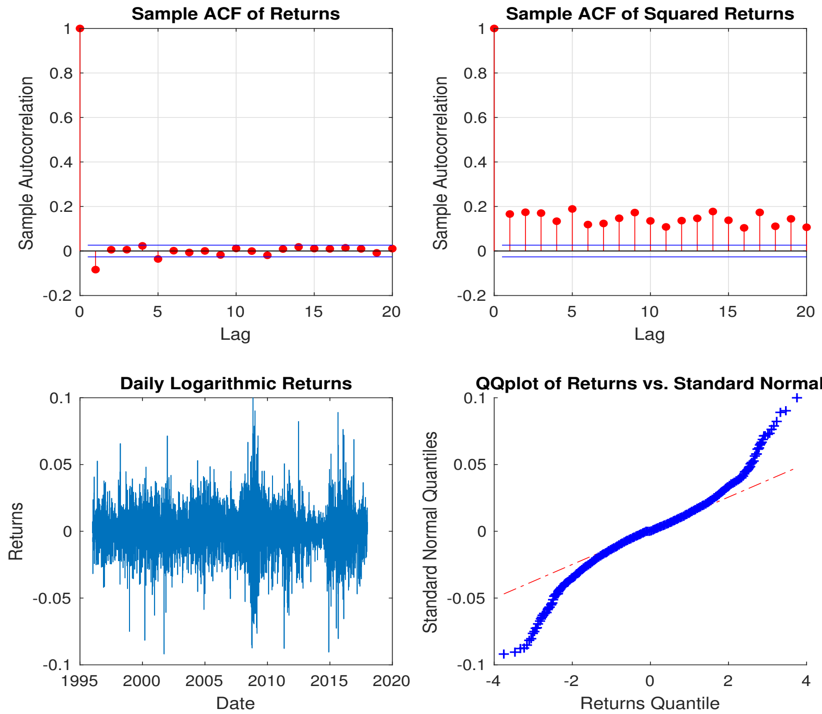

Figure 3.2: Sample autocorrelation plot of the returns and squared returns for WTI, as well as the daily logarithmic returns and a quantile-quantile plot. Corresponding graphs for all commodities are found in Appendix A1, A2, A3, A4.

Figure 3.3 displays the probability plot of WTI returns, which shows that the returns follow the t distribution better than the Normal distribution, but the data deviates from the t distribution in the tails. This means that we have heavy tails, in line with stylized fact 5. The Student t distribution is a symmetrical distribution. This indicates that it is not able to capture heavy tail asymmetry of the risk factors. This argument strongly points toward using a different distribution estimation to describe the tails.

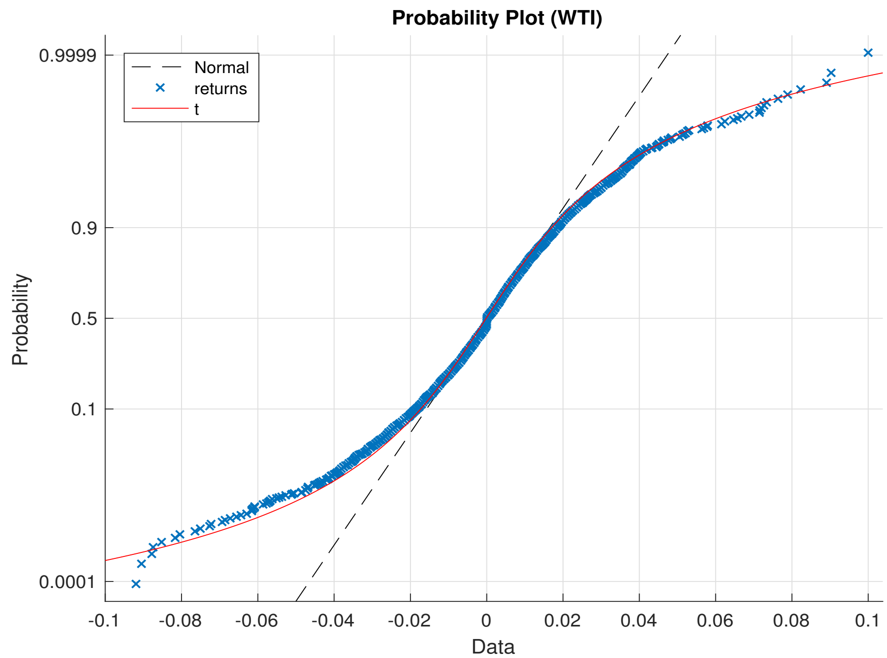

Figure 3.3: Probability Plot of WTI returns vs Standard Normal vs t distributions. Corresponding graphs for the other commodities are found in Appendix A5.

Commodity	ARCH(L5)	ARCH(L10)	ARCH(L15)	ARCH(L20)
Wheat	237.06 (0.00)	315.09 (0.00)	360.67 (0.00)	396.00 (0.00)
Corn	195.32 (0.00)	261.28 (0.00)	281.01 (0.00)	296.43 (0.00)
Soybeans	326.98 (0.00)	419.83 (0.00)	449.80 (0.00)	466.46 (0.00)
Live Cattle	63.49 (0.00)	86.39 (0.00)	103.79 (0.00)	124.39 (0.00)
Copper	908.33 (0.00)	1025.04 (0.00)	1070.35 (0.00)	1099.46 (0.00)
Gold	241.25 (0.00)	274.14 (0.00)	323.52 (0.00)	339.13 (0.00)
Aluminium	306.52 (0.00)	410.75 (0.00)	487.77 (0.00)	515.77 (0.00)
WTI	491.58 (0.00)	596.96 (0.00)	671.05 (0.00)	708.36 (0.00)
Brent	497.78 (0.00)	579.76 (0.00)	641.59 (0.00)	662.43 (0.00)
Natural Gas	57.17 (0.00)	74.71 (0.00)	106.23 (0.00)	116.06 (0.00)

Table 3.4: Results from ARCH tests for conditional heteroscedastisity for lags 5, 10, 15 and 20 (p-values in brackets). All the commodities show conditional heteroscedastisity.

The volatility clustering and serial correlation can be tested with a ARCH test for significant conditional heteroscedastisity, following Engle (1982). Results from the ARCH tests are found in Table

3.4. We performed the test for lags 5, 10, 15 and 20. The test rejects the null hypothesis of no ARCH effects for all the lags for all ten commodities. This implicates that we have significant conditional heteroscedastisity, and that we need to model volatility using a GARCH process.

Before we can proceed with the GARCH process the assumption of stationary data needs to be tested. A stationary process is a stochastic process whose expected value and variance do not change over time, and the unconditional joint probability distribution does not change when shifted in time. We will perform both Augmented Dickey Fuller test (ADF), Kwiatkowski–Phillips–Schmidt–Shin test (KPSS) and Phillips Perron test (PP). We follow Alexander (2008*b*) for methodology and refer the reader to this author for further details about stationary processes and the tests.

The results from the stationarity tests can be found in Table 3.5. Both Augmented Dickey-Fuller test, Phillips Perron test and KPSS test show that the returns of all the commodities are stationary, which allows to proceed with a GARCH process.

Commodity	ADF test	PP test	KPSS test
Wheat	-76.54 (0.001)	-76.54 (0.001)	0.083 (0.10)
Corn	-74.71 (0.001)	-74.71 (0.001)	0.073 (0.10)
Soybeans	-76.76 (0.001)	-76.76 (0.001)	0.083 (0.10)
Live Cattle	-75.90 (0.001)	-75.90 (0.001)	0.063 (0.10)
Copper	-81.80 (0.001)	-81.80 (0.001)	0.139 (0.063)
Gold	-76.08 (0.001)	-76.08 (0.001)	0.201 (0.016)
Aluminium	-78.57 (0.001)	-78.57 (0.001)	0.065 (0.10)
WTI	-82.35 (0.001)	-82.35 (0.001)	0.079 (0.10)
Brent	-83.10 (0.001)	-83.10 (0.001)	0.093 (0.10)
Natural Gas	-78.23 (0.001)	-78.23 (0.001)	0.023 (0.10)

Table 3.5: Stationarity tests statistics (p-value between brackets). Augmented Dickey Fuller (ADF) and Phillips Perron (PP) test for non-stationarity and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for stationarity. The tests conclude that the returns of the ten commodities are stationary.

As a result of the return characteristics for the ten commodities we model the conditional volatility with a GARCH process. A GARCH process can be extended in various ways, depending on the purpose. For commoditiy markets it has been shown that volatility tends to increase more after large negative returns, than after large positive returns (Nyström and Skoglund, 2002*b*). We therefore see it as an appropriate fit to extend to a GARCH-GJR model, which includes a leverage parameter to capture this asymmetry.

Since our focus is on stress testing extreme returns are of special interest. We have shown the deviation from the normal and student t distribution for the returns, especially in the tails, so these distributions are not satisfying for modelling the individual return time series. Extreme Value Theory, with the Peak over Threshold method, has in earlier studies (Aepli, 2011, Mudry, 2013, Wang et al., 2010) shown to be a good fit for modelling the tails accurately, and will therefore also be applied in our analysis.

Due to the common bust and boom cycles and co-integration of commodity markets, the dependency between the risk factors is important to be modelled realistically. Aepli et al. (2017) bring evidence to the importance of modelling time-variation and asymmetries in the dependence structure of a commodity futures portfolio. In support to Basel III critics on over-reliance on historical correlation, they introduce multivariate dynamic copula models as the superior alternative. There exists a numerous amount of copulas to choose from, and the best choice is dependent on the aim of the analysis and the data. Studies such as Mudry (2013), Aepli (2011) and McNeil, Frey and Embrechts (2015) find the t copula to be superior over the Gaussian copula in the context of modelling multivariate financial return data. For our purpose we therefore prefer a t copula over the more common Gaussian copula. The asymmetry of our data would probably be better modelled by an asymmetric copula. However, for our analysis we find the t copula suitable as it keeps the analysis traceable and allows a direct comparison across stress tests. The theoretical background for the methodology is presented in the next chapter.

4 Methodology

In this chapter we will present the theoretical background for modelling the distributions of the individual risk factors and the dependencies between them step by step. We will present an introduction to GARCH processes to obtain standardised residuals, before introducing the empirical background of Extreme Value Theory and copulas. The application for our data and analysis will be presented in the next chapter.

4.1 AR models

The following sections are based on Alexander (2008*b*). Autocorrelation, also called serial correlation, is when the return today is dependent on its passed history, the lagged returns. A simple autoregressive AR(p) process is a simple way to capture the autocorrelation between the individual commodity returns:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$$

where ϵ_t is i.i.d. with mean zero and variance σ^2.

The AR(p) process above can be extended to an ARMA(p,q) process when we include a MA(q) process to capture movements of the long term average. An ARMA(p,q) is given by:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \gamma_j \epsilon_{t-i} + \epsilon_t$$

ARMA models are typically used to account for linear serial dependence. However, ARMA models fail to take into account the asymmetry of returns, fat tails and volatility clustering, which has important implications for risk measurement. Therefore, an ARMA model is not optimal for our study, and we need other specifications for the variance to take the mentioned problems into account.

4.2 GARCH

The residuals in the previous AR(q) model can be decomposed such that:

$$\epsilon_t = z_t \sigma_t$$

where z_t is i.i.d. and σ_t is the conditional variance.

The generalized autocorrelation conditional heteroscedasticity model (GARCH) is then used to capture the volatility change and clustering of returns over time.

The symmetric normal GARCH assumes that the dynamic behaviour of the conditional variance is given by:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2\,, \quad \epsilon_t|I_{t-1} \sim N(0, \sigma_t^2).$$

The parameters of the GARCH model are estimated by maximising the value of the log likelihood function (see Alexander, 2008*b*, p. 137) .

4.2.1 GARCH-GJR model

It has been empirically shown that for commodity markets volatility tends to increase more after large negative returns than after large positive returns (Nyström and Skoglund, 2002*b*, p. 5). This is called a leverage effect and is captured by extending with one extra parameter, the leverage parameter. The GARCH-GJR model was originally designed for stocks where the higher impact on the volatility from the negative shocks than from positive shocks, is captured by the leverage parameter. Since commodities also often show negative skewness it can also be applied to our data. The GARCH-GJR can be written from the GARCH(1,1) model above, including the extra parameter (Alexander, 2008*b*, p. 150):

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \lambda 1_{(\epsilon_{t-1}<0)}\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

where the indicator function $1_{(\epsilon_t<0)} = 1$ if $\epsilon_t < 0$.

The estimation of the GARCH-GJR model is based on the likelihood function, same as the normal GARCH, but with the extra leverage parameter λ to account for the asymmetry.

Nyström and Skoglund (2002*a*, p. 10-12) discuss which distribution should be assumed for the filtered residuals z_t for financial data, and find that there is no empirical support for the symmetry in normal distributions, nor is there support for the exponentially decaying tails. They therefore point out that using normal distribution approximation for the high quantiles might lead to significant underestimation. They suggest the t distribution as an alternative distribution assumption, which might be more accurate in capturing fat tails, but unable to capture asymmetry. Nyström and Skoglund (2002*a*) use Extreme Value Theory to account for both the fat tails and the skewness and asymmetry of financial data. In this thesis we will apply this combined method. The distribution of the filtered residuals from the GARCH-GJR process is a t distribution, and the tails will be modelled based on Extreme Value Theory with the generalized Pareto distribution. The theoretical background will be presented in the following section.

4.3 Extreme Value Theory

Extreme value theory (EVT) is the study of improbable, but extreme events. EVT is more commonly used in weather and insurance, but has over the past decade become more popular also in financial studies. Embrechts, Mikosch and Klüppelberg (1997) introduced a full framework for the analysis, and they argue that EVT should be given more attention in risk management for financial institutions. McNeil, Frey and Embrechts (2015) introduced a combination of GARCH-EVT models where the GARCH residuals are used as input to EVT, since EVT requires the residuals to be i.i.d.

The theoretical framework for Extreme Value Theory is extensively shown in Nyström and Skoglund (2002*b*) and Embrechts, Mikosch and Klüppelberg (1997).

Let $X_1, ..., X_n$ be n observations from n independent and identically distributed random variables with the distribution function F. The focus is now to understand the distribution function with a closer look on its upper and lower tails. From this point of view we consider:

$$M_n = max(X_1, ..., X_n)$$

$$m_n = min(X_1, ..., X_n)$$

Both M_n and m_n are random variables that depend on the size of the sample n. We are focusing on understanding asymptotic behaviour of these random variables as $n \to \infty$. From this it is important to notice that $m_n = (-X_1, ..., -X_n)$, as this allows us to only concentrate on the upper tail for the underlying distribution by describing the theory for M_n in the following.

EVT provides us with the structure of the asymptotic limit of the random variable M_n that we defined above.

To clarify the theory we introduce an example, showed by Nyström and Skoglund (2002*b*, p. 9). Let $F(x) = (1 - exp(-x))X_{[0,\infty)}(x)$, and if we assume independence we have

$$P(M_n \leq x) = (1 - exp(-x))^n$$

If we now let $n \to 0$ the right hand side of the equation will, for every positive value of x, tend to zero. Hence, we cannot get a non-degenerated limit without a normalisation. One should therefore redo the calculation in such a way that

$$P(M_n \leq x + log\, n) = (1 - exp(-(x + log\, n)))^n$$

$$= (1 - \tfrac{exp(-x)}{n})^n \to exp(-exp(-x)) =: \Gamma(x)$$

From this the limit indicates that we let n tend to infinity. And from this we can prove that the convergence is uniform, and so for large numbers we have

$$P(M_n \leq x) \sim \Gamma(x - log(n))$$

Related to this derivation, the generalised form of extreme value distribution introduced by Nyström and Skoglund (2002*b*) is

$$\Gamma_{\xi,\mu,\sigma}(x) = exp(-(1 + \xi \tfrac{(x-\mu)}{\sigma})_+^{-1/\xi}), x \in R$$

$\frac{1}{\xi}$ is known as the tail index. It indicates the heaviness of the upper tail of the underlying distribution F. If we let $\xi \to 0$ the tail index will go towards infinity and $\Gamma_{\xi,\mu,\sigma}(x) \to \Gamma((x-\mu)/_{\sigma})$ with the parameters μ and σ as translation and scaling.

$\xi = 0$ is the Gumbel distribution

$\xi > 0$ is the Fréchet distribution

$\xi < 0$ is the Weibull distribution

The Gumbel distribution indicates that the distribution will extend along the entire real axis, while the Weibull distribution indicates that the distribution has a upper bound. The Fréchet distribution has a lower bound, and t distributions empirically tends to converge to the Fréchet distribution. The different cases are illustrated in Figure 4.1.

The generalised Pareto distribution (GPD) is introduced for any $\xi \in \mathbf{R}, \beta \in \mathbf{R}_{+}$:

$$GP_{\xi,\beta}(x) = 1 - (1 + \xi \tfrac{x}{\beta})_{+}^{-\frac{1}{\xi}}, x \in \mathbf{R}$$

The link between the generalised Pareto distribution and extreme value theory can be expressed as the following:

$$1 - GP_{\xi,\beta}(x) = -\ln \Gamma_{\xi,0,\beta}(x)$$

From the literature there are two practical methods for applying EVT to our data. The first is the block maxima method. In this approach we define blocks in the data, and then extract the maxima (maximum loss) in each block. There are several drawbacks to this approach. The local maxima in a block might not capture the actual maximas in the time series, and the second and third maxima in a block might be of significance to the investor but will not be captured by the block maxima approach.

The second method, the Peak over threshold, focuses on the events that exceed a specified, high threshold. Here the observations over the threshold are asymptotically described by the generalised Pareto distribution. The Peak over threshold is the preferred method for practitioners as it makes better use of the data, and so we will use this approach in our thesis.

The central point here is how the GPD can be used in the tail estimation for the unknown distribution $F(x)$. The distribution function of the excess distribution over the selected threshold u is:

$$F_u(z) = P(Z \leq z | Z \geq u) = \tfrac{F(z)-F(u)}{1-F(u)}$$

And so,

$$1 - F(z) = (1 - F(u))(1 - F_u(z))$$

The goal is to apply EVT to the estimation of the tail for large values of z. To do this we have to define where the tail starts, in other words we need to select the threshold u. For our sample with n point sorted according to sixe: $Z_{n,n} \leq ... \leq Z_{1,n}$ we can define the upper tail by the integer $k < n$. The observations in the upper tail for the distribution are then $Z_{k,n} \leq ... \leq Z_{1,n}$. The threshold is then $u = Z_{k+1,n}$, and ${}^{k}/_{n}$ is an estimator for $1 - F(u)$.

For any $z > u$ the estimation of the tail distribution below is obtained by the GPD.

$$\tilde{F}(z) = 1 - \frac{k}{n}(1 + \tilde{\xi}\frac{(z - Z_{k+1,n})}{\tilde{\beta}})_{+}^{-\frac{1}{\tilde{\xi}}}$$

$\tilde{\xi}$ and $\tilde{\beta}$ are parameter estimates of the generalised Pareto distribution. To estimate the parameters in the GPD we use maximum likelihood. This is the preferred method as it provides estimates of the parameters that are consistent and asymptotically normal as $n \rightarrow \infty$ given that $\xi > -1/2$. When using the maximum likelihood it is nearly invariant to the level of the threshold given that the threshold is within a reasonable limit. Determining the optimal threshold is challenging, and there are several methods which can be used. However, Nyström and Skoglund (2002*b*) argue that the threshold should be between 5-13% of the data.

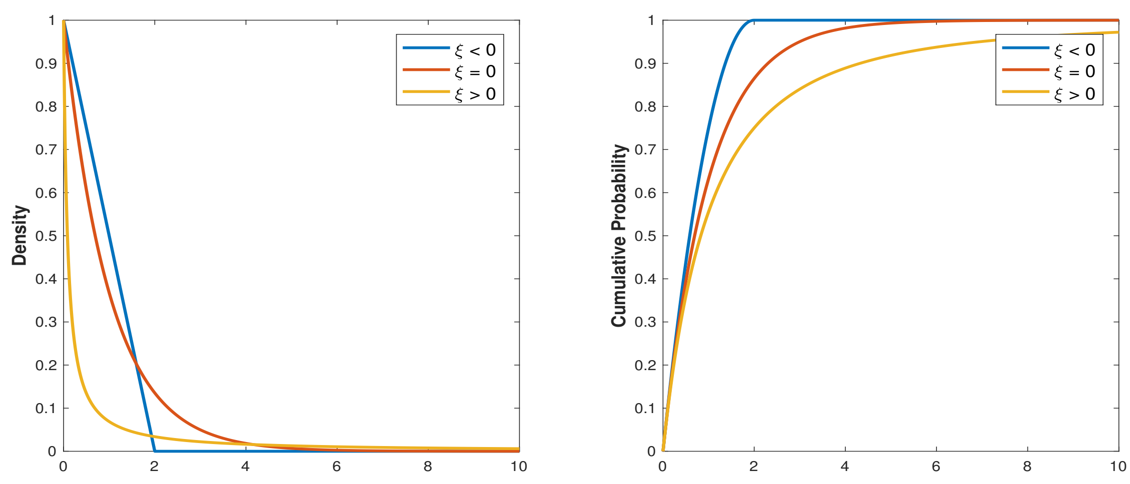

Figure 4.1: Probability density of the generalized Pareto distribution with β = 1 in all cases

4.3.1 Kernel smoothed interior

The data between the thresholds are fitted by a Gaussian Kernel estimator. A kernel estimator is a function that derives a smooth curve from the observed data that is the best possible representation of the probability density. Alexander (2008a, p. 106) defines the kernel approximation to the density of X as:

$$\hat{f}(x) = (nh)^{-1} \sum_{i=1}^{n} K \frac{(x - x_i)}{h}$$

where $(x_1, x_2, ..., x_n)$ is a random sample on a random variable X, K is the kernel function, that satisfies $\int K(x)dx = 1$, and h is the bandwidth parameter, which is strictly positive. K is normally selected from a unimodel probability density function which needs to be symmetric around zero. In most cases the choice of density function is of little importance. In this thesis we will use the Gaussian one, such that

$$K(x) = \tfrac{1}{\sqrt{2\pi}} exp(-\tfrac{x^2}{2})$$

4.4 Dependence structure

The GARCH-GJR-EVT process focuses on modelling the distribution of individual risk factors by modelling the conditional volatility, asymmetric adjustment and fat tails. However, this is done by modelling each risk factor in isolation, and tells us nothing about the dependence structure, which is a very important part of stress testing. The most popular and simple measurement of dependence is Pearson's linear correlation. In the following section we discuss drawbacks of employing simple correlations and motivate the use of copula functions.

4.5 Pearson's linear Correlation

Pearson's linear Correlation can be defined by (Alexander, 2008a, p. 112):

Let X and Y be vectors of random variables with non zero finite variances, then Pearson's linear correlation is given by

$$\rho(X, Y) = \tfrac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

where $Cov(X, Y)$ denotes the covariance of X and Y, while $Var(X)$ and $Var(Y)$ denote the variances, respectively. ρ can take on values in $[-1, 1]$. If $\rho = 0$, $X$and Y are independent.

Embrechts, Mcneil and Straumann (1999) discuss that if risks have an elliptical distribution, such as a jointly multivariate normal distribution, then the use of standard correlation is unproblematic. This is however very often not the case. With elliptical distributions we constitute an ideal world where nothing can go wrong. When moving forward with that assumption it can lead to a wide misjudgement of risk. This means that outside the elliptical world simple correlation must be used with care.

Another major fallacy pointed out by Embrechts, Mcneil and Straumann (1999) is that all correlations between -1 and 1 can be attained. When Pearson's linear correlation coefficient is 1, it is interpreted as perfect positive dependence, and perfect negative dependence for -1. For that to be possible two risk factors must have identical distributions up to a change of location and scale. This is again only true in the elliptical world. Further on, a correlation of zero does not indicate independence of risks. They also correctly point out that correlation is not invariant under transformations of the risks. For example, log(X) and log(Y) generally do not have the same correlation as X and Y (Embrechts, Mcneil and Straumann, 1999, p. 5) .

4.6 Spearman's rho and Kendall's tau (concordance)

The Spearmans's rho, and also Kendall's tau, are perhaps the best alternatives to the linear correlation coefficient to measure dependence for non-elliptical distributions like ours (Embrechts, Lindskog and J. Mcneil, 2001). Spearman's rho is called the correlation of ranks, and is a measure of concordance. Concordance is a definition of association between two random variables.

Consider two pairs of observations of continuous random variables X and Y, and denote them (x_1, y_1) and (x_2, y_2). For the pairs to be concordant $x_1 - x_2$ has to have the same sign as $y_1 - y_2$. And, for the pairs to be discordant $x_1 - x_2$ has to have the opposite sign of $y_1 - y_2$. If the proportion of concordance in a sample is increasing, then the probability that a large value of X is paired with a large number of Y is also increasing.

Rank correlations focus on the rank of the data point instead of the data itself. The smallest return is given the rank of 1, and the second smallest return is ranked 2 and so on. Spearman's Rho is estimated for a data set of two variables, by ranking the data in each of the two time series, and then taking the sum of the squared differences in rank between the two data series. It is defined by:

$$\rho_s(X_1, X_2) = \rho(F_1(X_1), F_2(X_2))$$

For a vector $(X_1, X_2)^T$ of continous random variables with a copula C, the Spearman's rho is given by:

$$\rho_s(X_1, X_2) = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3$$

4.7 Copulas

A copula allows for modelling the dependence structure without knowledge about the underlying joint distributions. It bases on the joint distribution of two or more assets by only specifying the marginals. Alexander (2008*b*) points out that one of the advantages of using a copula is that it isolates the dependence structure from the structure of the different marginal distributions. A copula can therefore be modelled for any marginal distribution, and the marginal distributions for the different risk factors do not have to be the same. The theory and notation behind copulas was already introduced in 1959 by Sklar. The use of copulas to measure dependence became more popular in the literature in the end of the 1990s, but only in the recent decade have copulas become a much used financial tool.

4.7.1 Sklar's theorem

There exists a very large number of copulas, some more advanced than others. We will base our theoretical framework of copulas on Alexander (2008*b*). To be able to understand copulas we introduce The Sklar's theorem (Sklar, 1959):

Theorem 1 *Consider a n-dimensional distribution function with marginal distributions* $F_1, ...F_n$. *Then there will exist a copula* $C : [0, 1]^n \to [0, 1]$, *such that*

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n))$$

for all $x_i, ..., x_n$ in $[-\infty, \infty]$. C is unique if the marginal distributions are continuous.

If it exists the copula density from the theorem is the function

$$c(F_1(x_1), ..., F_n(x_n)) = \frac{\partial^n C(F_1(x_1),...,F_n(x_n))}{\partial F_1(x_1).....\partial F_n(x_n)}$$

If we have the marginal densities $f_i(x) = F_i^{'}(x)$ we will have access to the joint density of the original variables from

$$f(x_1, ..., x_n) = f_1(x_1)...f_n(x_n)c(F_1(x_1), ..., F_n(x_n)).$$

Since the values of the marginal distribution function $F_i(x_i))$ are uniformly distributed we can give the copula distribution an alternative notation by using uniformly distributed variables $u_i \in [0, 1]$ as representation of the values of the marginal distributions at the realisations x_i (Alexander, 2008*b*).

The copula density with $u_i = F_i(x_i)$ might be written as

$$c(u_1, ..., u_n) = \frac{\partial^n C(u_1,...,u_n)}{\partial u_1..-\partial u_n}$$

Often the density function rather than the copula distribution is used in practise as they are easier to interpret.

4.7.2 Tail dependence

Tail dependence for copulas tells us about the concordance in the tails for the joint distribution. In other words, the upper tail dependence for two random variables implies that large positive values of U_1 tend to go together with large positive values for U_2. For financial data the copulas tend to have higher dependency in the tails. The upper tail dependence is defined by:

$$\lambda_{ij^u} = lim_{q\to[1]}\ P(X_i > F_i^{-1}(q)|X_j > F_j^{-1}(q))$$

given that the limit exist. The coefficient λ is a conditional probability, and so $\lambda_{ij}^u \in [0, 1]$. The copula shows upper tail dependence if $\lambda_{ij}^u > 0$.

Similarly, we have the lower tail dependence defined by:

$$\lambda_{ij^l} = lim_{q\to[0]}\ P(X_i < F_i^{-1}(q)|X_j > F_j^{-1}(q))$$

given that the limit exist. We have that $\lambda^l_{ij} \in [0, 1]$. And the copula has a lower tail dependency if $\lambda^l_{ij} > 0$.

If $\lambda^l_{ij} = \lambda_{ij^u}$ the tail dependence is symmetrical, and asymmetric tail dependence if the coefficients are different.

4.8 Student t copula

We will only focus the theoretical background of the symmetric t copula. For a more detailed and complementary background of copulas we refer the reader to Alexander (2008*b*) or McNeil, Frey and Embrechts (2015).

4.8.1 Theoretical background

The multivariate t copula can be derived from the multivariate t distribution, and is defined as (Alexander, 2008*b*, p. 268):

$$C_v(u_1, ..., u_n; \mathbf{\Sigma}) = \boldsymbol{t}_v(t_v^{-1}(u_1), ..., t_v^{-1}(u_n)),$$

where $\boldsymbol{t}_v$ and t_v are multivariate and univariate Student t distribution functions. v is the degrees of freedom, and $\mathbf{\Sigma}$ is the correlation matrix.

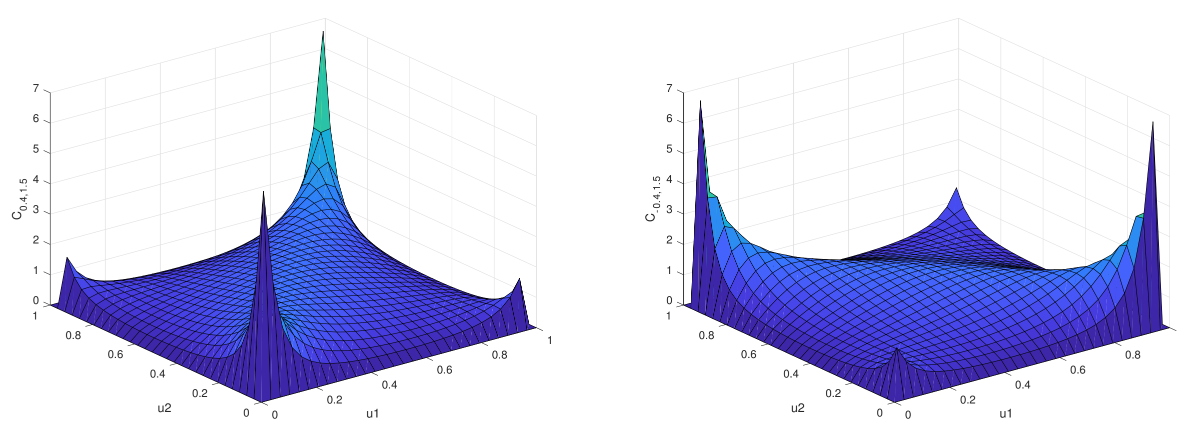

Figure 4.2: t copula with ρ = +/-0.4 and DoF 1.5. Source: Mudry (2013)

4.8.2 Estimation

To use the copula one will need to estimate the different parameters. Maximum likelihood estimation is a common tool in this process. The maximum likelihood estimator is applied to the theoretical joint distribution function. This is a difficult process, especially for the multivariate case in our analysis. This study will use the less time consuming approach of an approximate maximum likelihood function. It is known as a canonical maximum likelihood (CML), or a semi-parametric maximum likelihood estimation (see Bouyé et al., 2000, for parametric estimation). This method does not imply any assumptions of the marginal distributions, but is based on the empirical distributions. After transforming the standardized residuals to uniform variates the copula is fitted to the transformed data.

5 Application on the portfolio

In this chapter we will apply the presented methodology on our data. First we model the conditional volatility with a GARCH-GJR model to get the filtered residuals and the filtered conditional standard deviations. Further on, we standardise these residuals to model the tails with EVT. The tails are modelled with a generalized Pareto distribution, by the Peak over threshold method. By combining the tails with a Kernel smoothing for the interior, the entire distribution of the individual risk factors is modelled. Next on, a multivariate t copula is applied to model the dependencies, before moving on to the simulations and stress testing in the next chapter.

5.1 Application of the GARCH-GJR

To find the appropriate lag structure for the GARCH(p,q) process we estimate models with q and p ranging from 1 to 6. To select the best model for the data we perform Akaike (AIC) and Bayesian (BIC) information criteria (Box et al., 2015, p. 193). These criterias are the preferred ones for selecting the best GARCH fit for the data because it penalises models for additional parameters estimated.

$$AIC = -2(log\hat{L}) + 2NumParams$$
$$BIC = -2(log\hat{L}) + NumParams * log(n)$$

The process which minimises these criteria is considered to be the best specification. Table 5.1 shows the results from the AIC and BIC criterion tests for lags from 1 to 2. We tested up to 6 lags, but the results show insignificant parameters, and higher AIC and BIC criterion than the displayed models. The table indicates that the GARCH(1,1) is the optimal choice overall, and we will continue with this specification. A similar approach has been applied in Mudry (2013) and Aepli (2011).

GARCH(p,q) / Commodity	AIC criterion				BIC criterion			
	(1,1)	(1,2)	(2,1)	(2,2)	(1,1)	(1,2)	(2,1)	(2,2)
Wheat	-32810	-32809	-32805	-32808	-32764	-32749	-32752	-32741
Corn	-33868	-33865	-33866	-33863	-33821	-33805	-33812	-33796
Soybeans	-34272	-34271	-34270	-34268	-34225	-34211	-34217	-34202
Live Cattle	-42571	-42569	-42568	-42569	-42524	-42509	-42515	-42503
Copper	-33360	-33357	-33358	-33357	-33313	-33297	-33305	-33290
Gold	-37650	-37649	-37648	-37644	-37603	-37589	-37594	-37578
Aluminium	-36445	-36444	-36439	-36446	-36399	-36384	-36385	-36379
WTI	-32432	-32428	-32434	-32431	-32385	-32368	-32381	-32364
Brent	-32637	-32634	-32642	-32638	-32591	-32574	-32589	-32572
Natural Gas	-25884	-25891	-25882	-25889	-25838	-25831	-25829	-25823

Table 5.1: AIC & BIC criterion for t distributed residuals for GARCH process with various (p,q) lag structures. Preferred model highlighted in blue.

Commodity	$\hat{\omega}$ (SE)	$\hat{\alpha}$ (SE)	$\hat{\beta}$ (SE)	$\hat{\lambda}$ (SE)
Wheat	1.36e-06** (5.63e-07)	0.043*** (0.005)	0.966*** (0.004)	-0.030*** (0.006)
Corn	1.43e-06** (5.80e-07)	0.097*** (0.010)	0.913*** (0.007)	-0.021* (0.012)
Soybeans	2.36e-06** (7.30e-07)	0.071*** (0.009)	0.937*** (0.007)	-0.039*** (0.009)
Live Cattle	1.67e-06** (4.21e-07)	0.043** (0.011)	0.907*** (0.009)	0.076*** (0.016)
Copper	1.05e-06* (5.30e-07)	0.034*** (0.006)	0.955*** (0.005)	0.017** (0.007)
Gold	3.80e-07 (3.01e-07)	0.056*** (0.007)	0.956*** (0.004)	-0.024** (0.008)
Aluminium	8.35e-07** (4.13e-07)	0.053*** (0.007)	0.946*** (0.006)	-0.007 (0.009)
WTI	6.96e-07 (5.19e-07)	0.034*** (0.006)	0.955*** (0.005)	0.019** (0.008)
Brent	8.07e-07* (5.30e-07)	0.036*** (0.006)	0.955*** (0.005)	0.015** (0.008)
Natural Gas	1.28e-05*** (2.74e-06)	0.040*** (0.007)	0.944*** (0.008)	-0.002 (0.008)

Table 5.2: Estimated GARCH-GJR(1,1) parameters for the variance equation. Standard errors (SE) in brackets. ω is the constant, α is the reaction, β is the persistens, λ is the leverage.
***significant on 1% level, **significant on 5% level, *significant on 10% level.

Table 5.2 displays the estimated GARCH-GJR parameters. The parameters closely align with previous empirical results for financial assets. Following Alexander (2008*b*, p. 137), the β is a measurement of the persistence in conditional volatility regardless of what happens in the market. Large β, above 0.9, indicates that high volatility following market stress will persist for a long time which is true for all of our commodities. α measures the reaction of conditional volatility to shocks in the market. The sum of the two parameters is the rate of convergence, for our risk factors the sum is close to 1, indicating high persistence and a relatively flat term structure of volatility forecasts. From the table we see that the estimated ARCH and GARCH coefficients, $\hat{\alpha}$ and $\hat{\beta}$, are significant different from zero for all commodities. The estimated leverage parameter $\hat{\lambda}$ indicates that not all commodities show significant asymmetry.

The residuals from the GARCH process can be decomposed in two parts: $\epsilon_t = z_t \sigma_t$, where z_t is i.i.d. and t distributed, and σ_t is the conditional variance of each observation. Figure 5.1 displays the filtered residuals and the filtered conditional standard deviation of WTI, corresponding figures are shown in Appendix A6 and A7. We observe that the GARCH process models realistically the volatility clustering pattern in commodity returns.

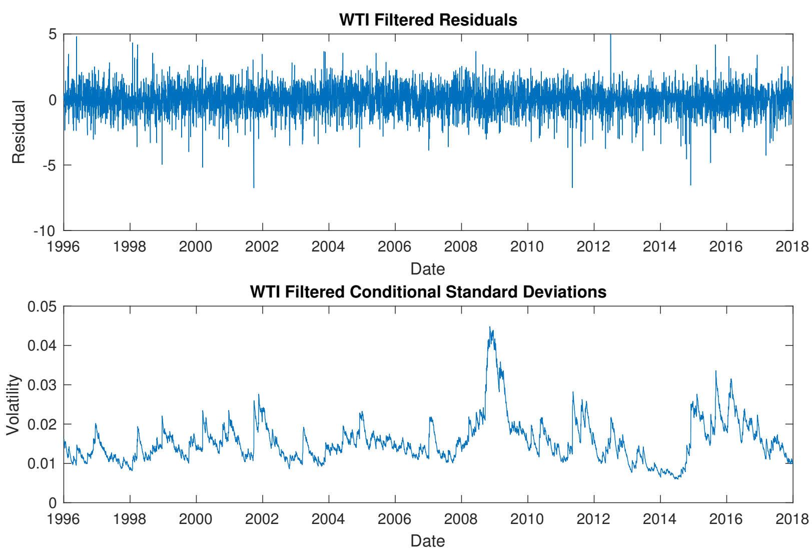

Figure 5.1: Filtered residuals and filtered conditional standard deviation for WTI. Corresponding graphs for the other commodities are in Appendix A6 and A7.

To be able to apply EVT to the tails we need to standardise the filtered residuals from each return series. The standardised residuals are calculated by dividing the filtered residuals with the conditional variance $z_t = \frac{\epsilon}{\sigma_t}$ to obtain mean zero and unit variance. The standardised residuals are plotted in Figure 5.2. We can now see graphically that the residuals are i.i.d. for WTI, and the other commodities show similar results (see Appendix A8 and A9). The residuals are now applicable to be modelled by EVT.

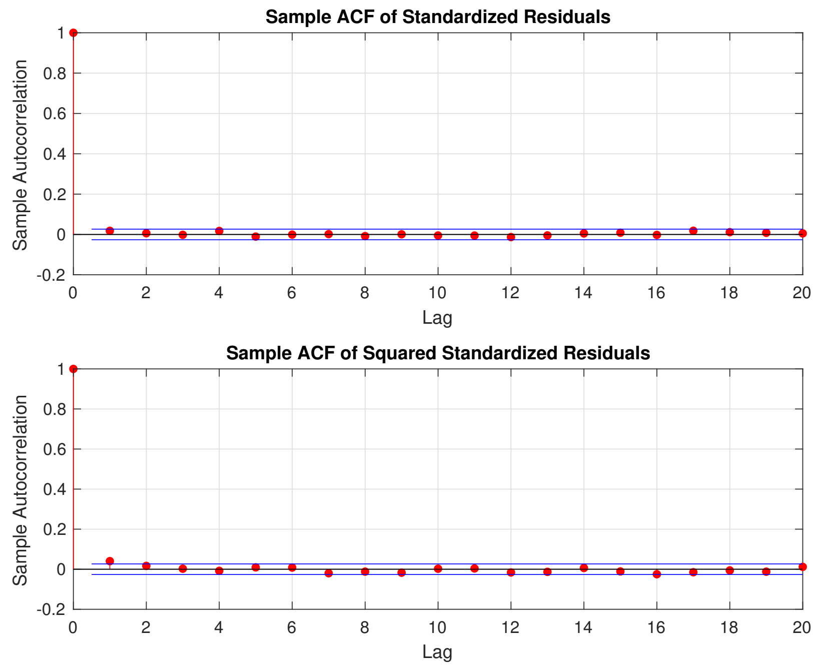

Figure 5.2: Sample autocorrelation plot of standardised and squared standardised WTI residuals, showing that the residuals are now i.i.d. Corresponding figures for the other commodities can be found in Appendix A8 and A9.

5.2 Application of the EVT

We will show WTI as an example in this subsection. Corresponding graphs for the other commodities can be found in the Appendix.

5.2.1 Estimation of the semi-parametric cumulative distribution functions

The next step is to fit the generalised Pareto distribution to the exceedances over threshold by using maximum likelihood. This is according to the Peak over threshold method, also called distribution of exceedances. By optimising the log-likelihood function we estimate the tail indexes ξ and scale parameters β.

We have fitted our data to the GPD with different thresholds between 5% and 15%. This allows us to find a threshold where the tail indexes stabilise. In Table 5.3 we display the 7%, 10%, 11% and 12% thresholds upper tail index. The rest of the parameters from the different thresholds are located in Appendix A1. Notice that the tail index naturally becomes smaller as the threshold allows for more data in the maxima.

	Upper tail (ξ), Threshold = u			
Commodity	u=7%	u=10%	u=11%	u=12%
Wheat	0.0822	0.0223	0.0082	0.0063
Corn	0.1885	0.1775	0.1554	0.1320
Soybeans	0.0816	0.0504	0.0430	0.0356
Live Cattle	0.2781	0.2854	0.2702	0.2714
Copper	-0.0069	-0.0290	-0.0572	-0.0555
Gold	0.2011	0.1450	0.1485	0.1413
Aluminium	-0.0368	0.0022	0.0117	-0.0278
WTI	0.0646	0.0101	-0.0049	-0.0158
Brent	-0.0021	0.0086	-0.0101	0.0006
Natural Gas	0.1583	0.0728	0.0688	0.0871

Table 5.3: Comparison of upper tail parameters (ξ) for different thresholds. The rest of the parameters from the different thresholds are found in Appendix A1.

We will use 10% upper and lower threshold, which is well in the range 5% - 13% suggested by Nyström and Skoglund (2002*b*). The reason is twofold, firstly the 10% threshold seems to be stable, and give the most accurate picture of the maximas. Secondly, previous studies such as Aepli (2011) and Mudry (2013) have chosen a 10% threshold. We see the advantage of choosing a standard threshold as it gives us the opportunity to quality control our results.

The estimated parameters for our risk factors are listed in Table 5.4. We have Fréchet tails for nine of the ten risk factors, only copper shows a negative tail index, which suggests a moderate upper tail. The lower tails are positive for all commodities. This suggests both fat upper and lower tails, and hence tail asymmetry. The findings are consistent with the theory and empirical results for financial time series (Nyström and Skoglund, 2002*b*, Embrechts, Mikosch and Klüppelberg, 1997).

Commodity	ξ Upper tail	ξ Lower Tail	β Upper tail	β Lower Tail
Wheat	0.0223	0.0822	0.6163	0.4942
Corn	0.1775	0.1561	0.5396	0.5081
Soybeans	0.0504	0.1014	0.5595	0.5924
Live Cattle	0.2854	0.2509	0.4534	0.5704
Copper	-0.0290	0.1178	0.5818	0.5654
Gold	0.1450	0.1003	0.5203	0.6121
Aluminium	0.0022	0.1003	0.5551	0.5197
WTI	0.0101	0.0926	0.5166	0.5805
Brent	0.0086	0.0664	0.5152	0.5965
Natural Gas	0.0728	0.0714	0.6126	0.5191

Table 5.4: Maximum likelihood estimators for the generalized Pareto distribution parameters. Threshold (u) = 10%.

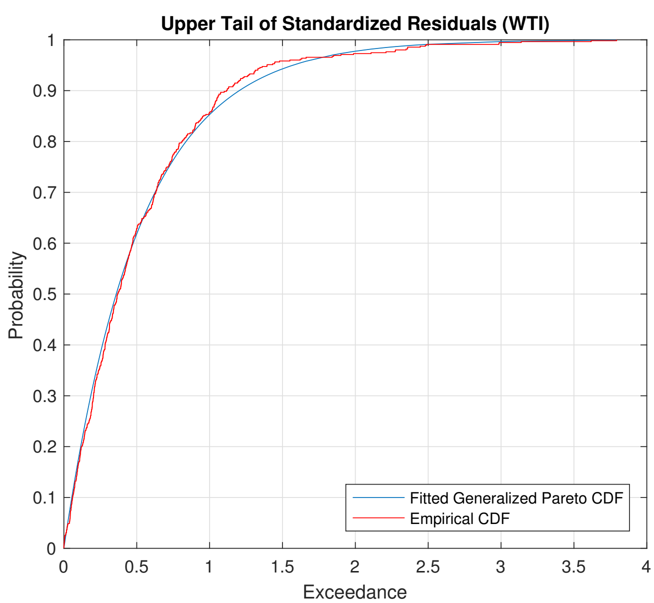

Figure 5.3: Generalized Pareto upper tail of the standardised residuals fitted vs. empirical. Corresponding figures for the other commodities can be found in Appendix A11.

In Figure 5.3 we display the empirical cumulative distribution of the upper tail of the standardised residual exceedances for WTI. The fitted distribution follows the empirical exceedances closely, and so the chosen distribution is well suited to estimate the tails for the commodities.

The last step is to combine the parametric generalized Pareto tails for each commodity with the corresponding Kernel smoothed interior to obtain the entire semi-parametric cumulative distribution function. Figure 5.4 displays the semi-parametric empirical cumulative distribution function of WTI standardized residuals. The piecewise distribution object allows interpolation within the interior of the CDF, displayed in black, and extrapolation in each tail, displayed in red and blue for the lower and upper tail, respectively. The extrapolation allows for estimation of quantiles outside the historical record, and is therefore important for the stress testing exercise.

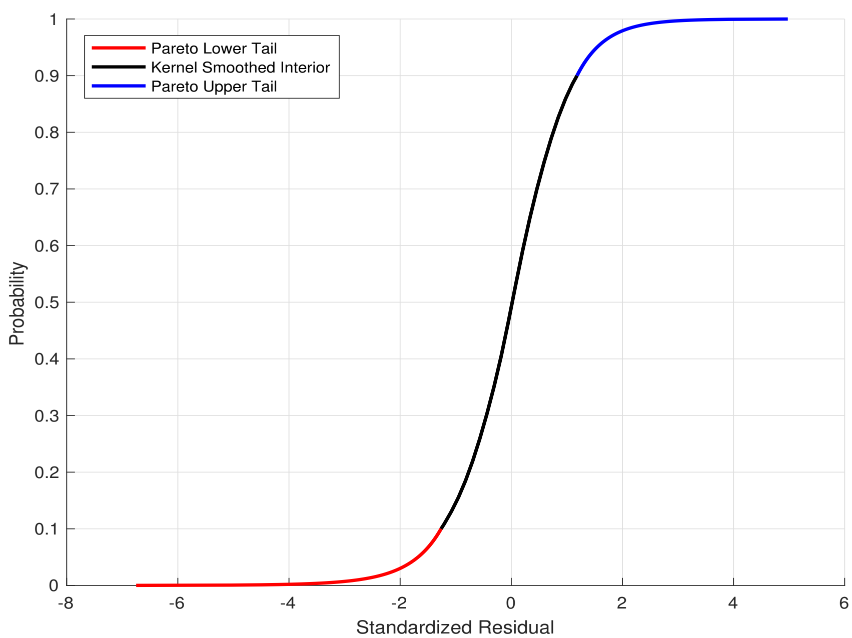

Figure 5.4: Semi-parametric empirical cumulative distribution function of WTI. Corresponding figures for the other commodities can be found in Appendix A10.

5.3 Application of the t copula and simulation steps

Given the parameters of the t copula (the correlation matrix R and the degrees of freedom parameter) we simulate jointly dependent portfolio returns. This is done by first simulating the corresponding dependent standardised residuals. This step is where we transform the uniform variables to standardised residuals through the inversion of the semi-parametric marginal cumulative distribution function of each risk factor. To do this we extrapolate into the generalized Pareto tails, and interpolate into the smoothed interior. This gives simulated standardized residuals consistent with those obtained from the GARCH-GJR(1,1) filtering process described in Section 5.1 . These residuals show no autocorrelation and are i.i.d. Each column of the simulated standardized residuals array represents an i.i.d. univariate stochastic process when viewed in isolation, whereas each row shares the rank correlation induced by the copula. Next on, we reintroduce the autocorrelation and heteroskedasticity observed in the empirical risk factor returns. This is done by using the simulated standardized residuals as the i.i.d. input noise process.

5.3.1 Risk metrics

To compare risk properties the use of Value at risk (VaR) and conditional value at risk (CVaR) is common. These risk metrics provide an indication of the quantile losses. VaR is the amount of maximum potential loss at a given percentage. Critics of the VaR point to the inability to account for the properties beyond the VaR (Alexander, 2008*c*). CVaR corrects for the limitations of VaR, often serving as a more thorough tool for the risk manager than VaR as it measures the average loss in the tail beyond VaR (Alexander, 2008*c*). For our analysis we include both risk metrics at various quantiles to mitigate the potential shortcomings of the individual risk metrics. This is in line with European Banking Authority (2017, p. 28): ”The institutions should stress the identified risk factors using different degrees of severity as an important step in their analysis to reveal nonlinearities, threshold effects, i.e. critical values of risk factors beyond which stress responses accelerate.”.

6 Stress testing and simulation

In this chapter we will perform stress tests on our commodity portfolio. We will simulate the portfolio profit and loss distribution for each stress test case, to illustrate the role of stress testing exercises for the portfolio risk management. For the simulations we calibrate the models on the historical time period from 1996 - 2017 which makes the Baseline scenario. For the historical stress period we use the years 2007 and 2008 to observe the impact of the financial crisis. This time period is known for high market stress with high volatility and captures the simultaneous pricedrop during the financial crisis (see Section 3.2 for discussion). The residuals and parameters of the GARCH process and the t copula are here re-calibrated on the stress horizon, following the same procedure as in Chapter 5. The re-estimated tails from the generalized Pareto distribution parameters can be found in Appendix A2. For all scenarios we run 20 000 simulations over a 22 days horizon. Note that the portfolio weights are held fixed over the risk horizon and that the simulation ignores any transaction costs required to re-balance the portfolio (the daily re-balancing process is assumed to be self-financing). Since we run simulations, we will compare risk metrics across scenarios to see tendencies and form expectations, rather than an exact estimation of the tail losses that would occur.

6.1 Scenario composition

We limit the study mostly to hybrid scenarios, where we calibrate the model for risk factors or the copula to the financial crisis data and shock them simultaneously or one at a time. Even more hypothetical scenarios could have been implemented. Examples could be a scenario with recession in China, which would decrease the demand for aluminium, oil, copper, soybeans and natural gas, and look at the change in dependence structure and volatility. Other scenarios could be natural disasters that affect crops or diseases that affect grains or livestocks. This is however out of the scope of this thesis. We refer the reader to Aepli (2011) for stress testing with hypothetical scenarios.

Our analysis consists of seven different scenarios. Underneath follows a brief description of the scenarios before the analysis is conducted.

Baseline scenario:

The baseline scenarios is a default scenario simulation with the t copula and GARCH-GJR process calibrated on the entire data set. Non of the parameters are stressed and the procedure follows Chapter 5. The baseline scenario is constructed to be a reference for normal times to assess the effect of stressing parameters compared to the steady state.

Historical scenario:

The historical scenario is a scenario without simulations. It is the actual returns over our time period, so we refer to the empirical profit and loss distribution as observed. Unfortunately, due to the limited number of observations when resuming ourselves to observed returns extreme quantiles are hard to estimate.

Hybrid scenarios:

Due to the limitations of the historical empirical scenario, as discussed in Section 2.1, we construct five hybrid scenarios. Hybrid scenarios allow extrapolation beyond realized returns, and are therefore appropriate to estimate extreme quantiles and events that have not yet occurred. The focus in our hybrid scenario construction is to examine which of the estimated parameters challenge mostly the test portfolios' profit and loss distribution. The parameters changed between the different scenarios are the dependencies between risk factors, measured in Degrees of Freedom and correlations, and the individual risk factor distributions (GARCH-GJR parameters). To isolate the effect of various parameters, we construct the scenarios by mixing parameters from the baseline with those during the period of financial distress. The hybrid scenarios are described in Table 6.1.

Scenario	Marginal distribution	Correlations	Degrees of Freedom
Risk factor stress	stress	baseline	baseline
Dependency stress	baseline	stress	stress
Full stress	stress	stress	stress
DoF shock	baseline	baseline	stress
Risk factor stress without EVT	stress	baseline	baseline

Table 6.1: Description of input parameters for simulation in hybrid scenarios. Baseline means the parameters from entire dataset 1996-2017 are used as input. Stress means the parameters are re-calibrated on our chosen time of financial distress, years 2007-2008.

1. **Risk factor stress scenario** aims to show the impact of stressing the individual risk factors marginal distributions on the portfolios' profit and loss distribution, without a change in dependencies between the factors.

2. **Dependency stress scenario** isolates how the dependencies between the returns of portfolio components affect the profit and loss distribution, without changing the parameters for the individual factors model (GARCH-GJR).

3. **Full stress scenario** aims to simulate the effects of a recurring financial crisis on the portfolio.

4. In the **degrees of freedom shock** we shock only the degrees of freedom of the copula, leaving all other parameters unchanged.

5. Risk factor stress **without EVT** highlights how the application of Extreme Value Theory to model the tails of portfolio components returns affect the profit and loss distribution of the portfolio. The risk factor distributions are here not modelled with EVT, but with a Student t distribution (see Section 4.2.1).

6.2 Comparative analysis of simulated profit and loss distributions

6.2.1 Baseline scenario vs. Historical scenario

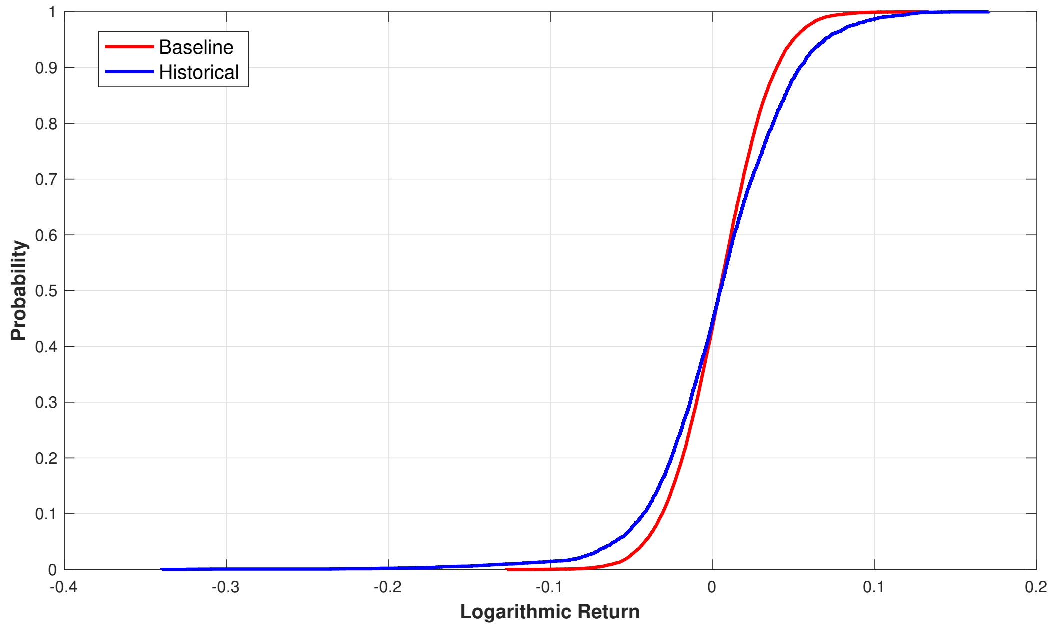

Figure 6.1: Portfolio returns simulation, Baseline vs Historical scenario.

Figure 6.1 displays the profit and loss distribution for the returns in baseline scenario simulated with the t copula and the cdf of the historical returns of the test portfolio. The simulated returns align with the historically observed returns in the centre, but deviate in both the upper and lower tails. This can be further viewed in Table 6.2, where the maximum simulated loss is significantly higher/lower for the historical returns than the simulated baseline, respectively. These results might be linked to the symmetry of the t copula. It can be that the dependencies between the risk factors are skewed. Furthermore, this difference steams from the different time spans of our simulation. The baseline represents normal market conditions, and the simulated extreme losses that were observed in the financial crisis challenge the profit and loss distribution under the hypothesis that a similar financial crisis episode would reoccur. This result highlights the importance of implementing forward-looking scenarios, both to simulate extreme returns in comparison to the baseline and to simulate beyond the historical empirical observed returns.

Metric	Baseline	Historical scenario
Degrees of Freedom	15.28	N/A
Max. Simulated loss	-12.72%	-34.01%
Max. Simulated gain	13.38%	17.08%
Simulated 90% VaR	-3.09%	-4.26%
Simulated 95% VaR	-4.17%	-6.01%
Simulated 99% VaR	-6.13%	-12.47%
Simulated 90% CVaR	-4.48%	-7.38%
Simulated 95% CVaR	-5.39%	-9.75%
Simulated 99% CVaR	-7.30%	-17.82%
Simulated 99.9% CVaR	-10.10%	N/A
Simulated 99.99% CVaR	-12.55%	N/A

Table 6.2: Simulation metrics for Baseline scenario and Historical scenario CDF.

The previous statement is further substantiated when we look at very high confidence levels displayed in Table 6.2. The historical scenario is limited to already experienced events so there are not enough observations in the data set to calculate the expected shortfall at very high confidence levels. This emphasizes the discussion about the scenarios in Section 2.1 and the drawback of using historical scenarios highlighted in Basel Committee on Banking Supervision (2009). In addition the historical scenario neglects the dependence structure between the risk factors, which is an important factor in stress testing. European Banking Authority (2017, p. 24) state that stress tests should take into account changes in correlations between risk types and risk factors and that correlations tend to increase during times of economic or financial distress. This statement and its implications for stress testing exercises will be further investigated in the next subsection where we analyse the hybrid scenarios.

6.2.2 Hybrid scenarios

In the following the results from the hybrid scenarios are analyzed. Table 6.3 shows the risk metrics from the scenarios. The tail dependence for the simulated returns is measured in the degrees of freedom parameter from the t copula. From the entire data set the DoF are 15.28, while during the stressed period they shift to 13.78. Our decrease in DoF signals that the tail dependence in the commodity portfolio is increasing during times of stress. Lower degrees of freedom indicate a higher tendency of extreme events to occur jointly across risk factors (Mudry, 2013, p. 58), which is in line with our simulation result.

	1	2	3	4	5
Metric	Risk factor stress	Dependency stress	Full stress	DoF shock	without EVT
Degrees of Freedom	15.28	13.78	13.78	13.78	15.28
Max. Simulated loss	-36.40%	-18.67%	-45.90%	-16.53%	-34.24%
Max. Simulated gain	31.93%	16.32%	42.89%	13.82%	32.38%
Simulated 90% VaR	-8.84%	-3.75%	-10.85%	-3.08%	-6.91%
Simulated 95% VaR	-11.91%	-5.05%	-14.53%	-4.12%	-9.78%
Simulated 99% VaR	-17.87%	-7.46%	-21.92%	-6.13%	-15.79%
Simulated 90% CVaR	-12.98%	-5.47%	-15.89%	-4.49%	-10.91%
Simulated 95% CVaR	-15.71%	-6.61%	-19.31%	-5.42%	-13.61%
Simulated 99% CVaR	-21.55%	-9.08%	-26.74%	-7.38%	-19.57%
Simulated 99.9% CVaR	-30.35%	-12.99%	-37.17%	-10.63%	-28.04%
Simulated 99.99% CVaR	-35.79%	-17.40%	-44.77%	-15.00%	-33.72%

Table 6.3: Risk Metrics for hybrid scenarios.

Risk factor stress vs. Dependency stress

Figure 6.2 shows the baseline scenario, the scenario where we stress the dependencies between the risk factors, the full stress scenario and the scenario where the individual risk factors are stressed. Starting from the baseline we can see that by only stressing the dependencies, the simulation displays more severe losses (green vs. red). The correlation matrix and the decrease in DoF show that the dependencies between the risk factors increase in times of stress (see Table 6.4), which leads to larger simulated losses for the portfolio overall. However, by stressing only the GARCH-EVT parameters for the individual risk factors the effect on the portfolio is even stronger (red vs. light blue). This result indicates that stressing the individual risk factors has a larger impact on the profit and loss distribution than shifts in the dependencies between the risk factors. This shows that considering the individual portfolio components might be of higher relevance than the shifts in dependencies between them.

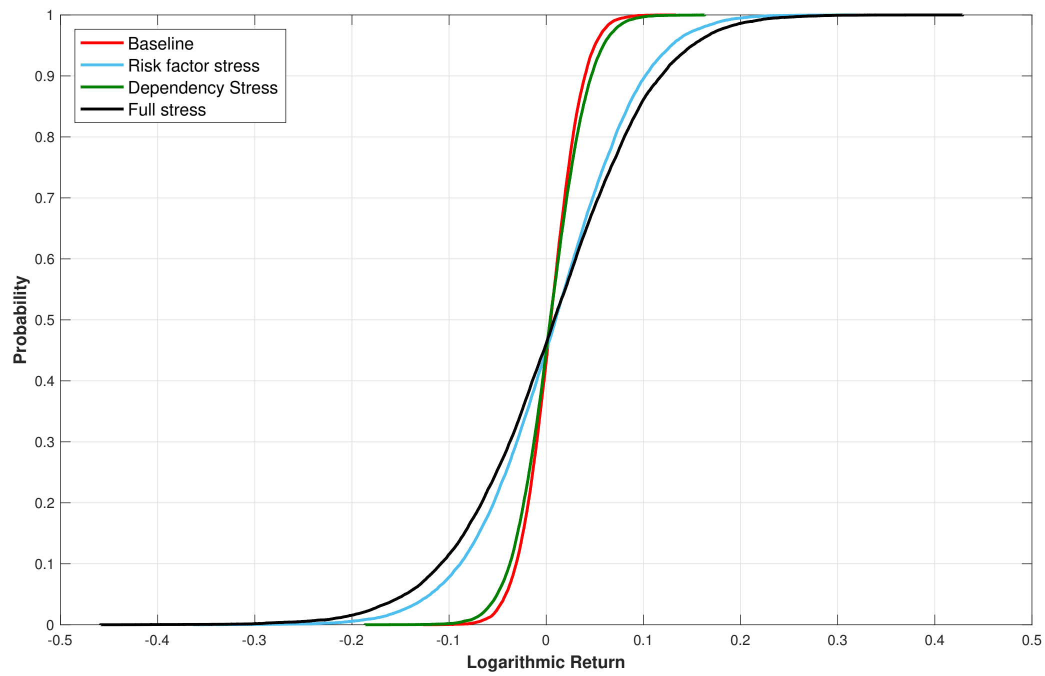

Figure 6.2: Simulated one-month portfolio returns CDF for Baseline vs Hybrid scenarios: Risk factor stress, Dependency stress and Full stress.

Commodity	Wheat	Corn	Soybeans	Live Cattle	Copper	Gold	Aluminium	WTI	Brent	Natural Gas
Wheat	-	-0.0342	0.0085	0.1184	0.1634	0.1518	0.0993	0.1622	0.1688	0.0579
Corn	-0.0342	-	0.0562	0.1542	0.1561	0.1564	0.1569	0.2181	0.2156	0.1259
Soybeans	0.0085	0.0562	-	0.1797	0.1634	0.1851	0.1237	0.2668	0.2614	0.1246
Live Cattle	0.1184	0.1542	0.1797	-	0.1342	0.1327	0.1160	0.1546	0.1552	0.1409
Copper	0.1634	0.1561	0.1634	0.1342	-	0.2088	0.0583	0.1615	0.1631	0.0925
Gold	0.1518	0.1564	0.1851	0.1327	0.2088	-	0.1596	0.2597	0.2582	0.1304
Aluminium	0.0993	0.1569	0.1237	0.1160	0.0583	0.1596	-	0.1573	0.1671	0.1023
WTI	0.1622	0.2181	0.2668	0.1546	0.1615	0.2597	0.1573	-	0.0750	0.1907
Brent	0.1688	0.2156	0.2614	0.1552	0.1631	0.2582	0.1671	0.0750	-	0.1959
Natural Gas	0.0579	0.1259	0.1246	0.1409	0.0925	0.1304	0.1023	0.1907	0.1959	-

Table 6.4: Correlation increases between baseline and stress scenario. The correlation matrices from the Baseline and financial stress period are located in Appendix A3 and A4. (Light blue is correlation increase larger than 0.15. Dark blue is correlation increase larger than 0.20).

Further on, we can compare the mentioned scenarios with the full stress scenario. Naturally this stress scenario simulates the largest tail losses since both the dependencies and the individual parameters are stressed (black line). Comparing the risk metrics in Table 6.3 we see that the risk factor stress scenario simulates the second largest losses, after the full stress scenario, which substantiates the previous result.

Comparing with Mudry (2013) the full stress scenario in his study gives more severe losses overall as well as in the tails. Our study replicated the methodology for the marginal distributions, dependence structure and simulations from this study, making comparison ideal. The difference between the results might be explained by: i) The difference in weights of the test portfolio where our study uses weights from 2017 while Mudry (2013) use the weights from 2013. ii) Our extended data sample. We include the years 1996-1998, and 2011 - 2017 beyond his data set. iii) Differences might be due partially to the randomness in the scenario generation.

In Mudry (2013) natural gas makes 15.11% of the portfolio, while for our portfolio it is 9.6%. From the descriptive statistics natural gas is by far the most volatile commodity, and with the most extreme losses natural gas has performed poorly over the last decades compared to most of the other commodities. Several structural breaks in natural gas prices are also included in our data set, examples being the supply shortfall in Libya 2011 and the Russian export stop in 2012 (Nick and Thoenes, 2014). Including more natural gas the portfolio might therefore be one of the main reasons of the more severe simulated loss in his study.

Soybean is the second commodity with the most deviating weight from Mudry (2013). In our portfolio soybean make 15.66% of the total weight, in comparison to his 6.89%. Over our time period soybean returns showed low volatility. We expect that the increased allocation in soybean in our portfolio provides the same consequences as the down-scaling of natural gas.

DoF shock vs. Dependency stress

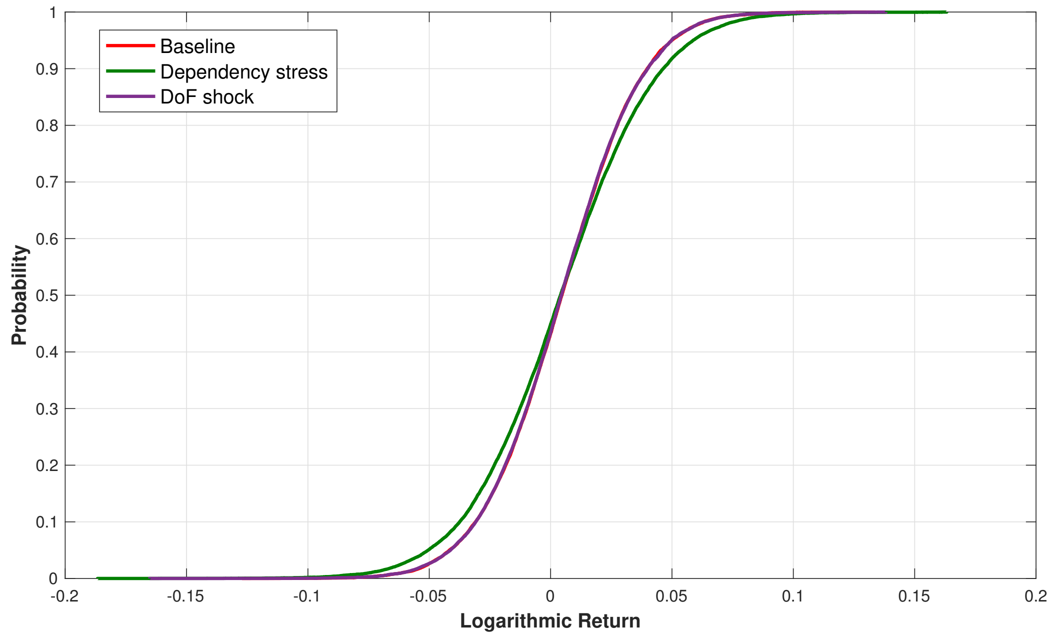

Figure 6.3: Simulated one-month portfolio returns CDF for Baseline vs Hybrid scenarios: Dependency stress and DoF shock.

In Figure 6.3 we compare the scenario where both correlation and DoF are stressed (green line), with the scenario where only the DoF are shocked from 15.28 for the Baseline to 13.78 for the financial crisis (purple line). For both scenarios the individual portfolio component parameters are calibrated on the entire data set. By doing so we can discuss the impact of correlations as a driver of losses in isolation.

From Figure 6.3 we observe that a small shock to degrees of freedom does not provide a significant stress scenario. The baseline scenario and the DoF shock scenario do not deviate much from each other (red vs. purple), although the DoF shock scenario simulates larger extreme losses in the lower quantiles (see Table 6.3). Furthermore we see that the scenario where the correlations between the risk factors as well as the DoF are shocked displays the largest simulated loss of the three scenarios. This indicates that shocking the DoF in isolation is limited in a stress testing exercise.

For more forward-looking hypothetical scenarios the implementation of more severe shocks to

DoF might be of interest. We therefore tested by including a set of hypothetical scenarios where several more substantial downward changes to DoF are included. The results can be found in Table 6.5. For our data we found that extreme shocks to DoF yield no substantial increase in simulated tail losses. Furthermore, the analysis shows that the impact from DoF shocks are very varying. When shifting the parameter some intervals yield higher losses, as expected, while others yield the opposite result. For example, when shifting from 13.78 to 10 DoF the simulated losses decrease, but when we continue the downward shift from 10 to 7 DoF the simulated losses increase. We therefore underline the importance of a more detailed analysis to avoid shortsightedness when implementing hypothetical scenarios.

Metric \ DoF	5	7	10	13.78	15.28	17
Max. Simulated loss	-16.72%	-17.25%	-13.85%	-16.53%	-12.72%	-15.88%
Max. Simulated gain	15.28%	15.47%	14.05%	13.82%	13.38%	16.77%
Simulated 90% VaR	-3.08%	-3.11%	-3.12%	-3.08%	-3.09%	-3.11%
Simulated 95% VaR	-4.12%	-4.18%	-4.20%	-4.12%	-4.17%	-4.19%
Simulated 99% VaR	-6.18%	-6.32%	-6.31%	-6.13%	-6.13%	-6.15%
Simulated 90% CVaR	-4.50%	-4.55%	-4.55%	-4.49%	-4.48%	-4.53%
Simulated 95% CVaR	-5.46%	-5.52%	-5.50%	-5.42%	-5.39%	-5.47%
Simulated 99% CVaR	-7.57%	-7.64%	-7.48%	-7.38%	-7.30%	-7.47%
Simulated 99.9% CVaR	-11.14%	-11.28%	-10.36%	-10.63%	-10.10%	-10.54%
Simulated 99.99% CVaR	-14.44%	-17.13%	-12.81%	-15.00%	-12.55%	-14.20%

Table 6.5: Risk metrics from different DoF shock scenarios. DoF 15.28 is from the Baseline (1996-2017), and DoF 13.78 is from the time of financial stress (2007-2008). The other DoF are hypothetical shocks.

Impact of EVT

In Figure 6.4 we display two hybrid scenarios to highlight the importance of implementing EVT for modelling extremely large return changes of portfolio components before running the actual stress testing. For both scenarios the correlation matrix and the DoF parameter are calibrated on the entire data set, so the difference between them is how the individual risk factors are modelled. In the risk factor stress scenario the tail distributions are modelled with EVT where the tail indexes are calibrated on the financial crisis data, and the other scenario with a Student t distribution. One can see that the scenario where EVT is implemented estimates more severe losses, where simulated

99.99% CVaR is -35.79% in comparison to -33.72% for the scenario without EVT. Overall, the profit and loss distribution in the stress test excluding EVT is shifted to the right. So applying EVT strengthens the accuracy and understanding of the most extreme, potential losses. In light of this we can say that the risk might be underestimated when the individual risk factor distributions are not modelled with consideration to the impact of extreme events (Embrechts, Mikosch and Klüppelberg, 1997).

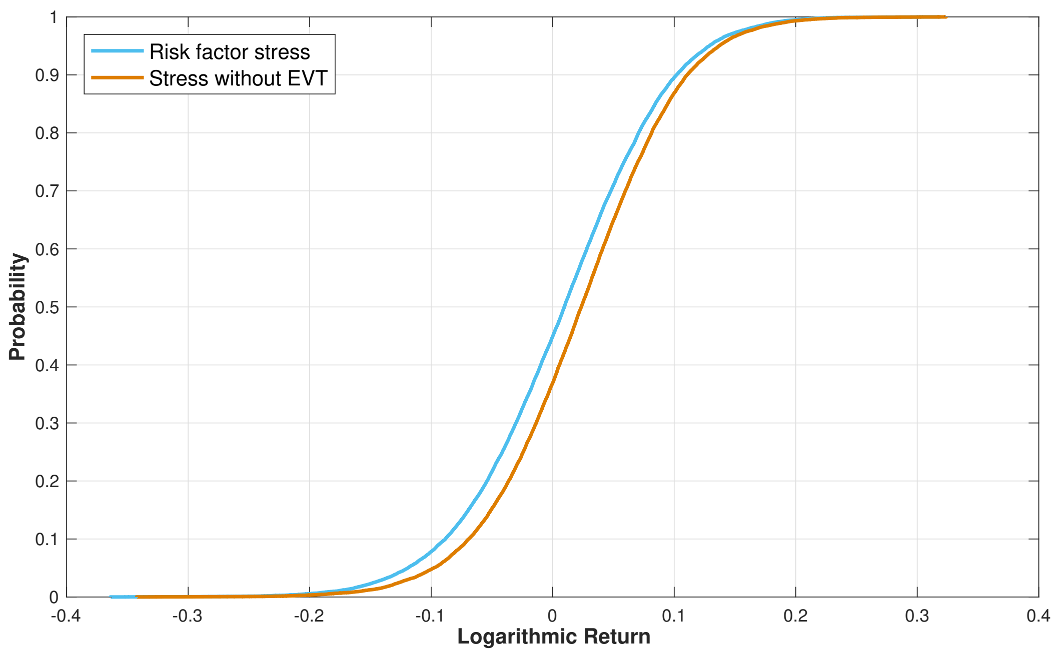

Figure 6.4: Simulated one-month portfolio returns CDF for Baseline and Hybrid scenarios: Risk factor stress and scenario without modelling with EVT.

Stress testing is a part of risk management for extreme events, so called black swans. Our analysis shows that one should not underestimate the impact from the extreme returns observed during the times off financial distress in forward-looking stress testing exercises. Further on we emphasise that it is useful to construct different hybrid scenarios to get a more comprehensive risk picture, in line with the requirements from Basel Committee on Banking Supervision (2009).

7 Conclusion

In this study we update the analysis in Paraschiv, Mudry and Andries (2015) with a more extensive data set, and a more detailed focus on stress testing. In particular, hybrid scenarios are explored. Our stress testing exercise reveals that we can rearrange arbitrarily shocks linked to a specific event or time to reveal the importance of correlations, tail correlations or extreme movements of portfolio components on the profit and loss distribution. This is the first study in the literature that clearly illustrates the marginal impact of the model for the individual portfolio components versus the marginal role of tail dependency on the portfolio risk profile.

We mimic the DJCI, by forming a portfolio of ten commodities. We use a GARCH-GJR approach to model stylized facts observed in commodity return data, and implement Extreme Value Theory to model the tails accurately. To account for the dependence structure we use a t copula. We then stress tested the portfolio with different scenarios, examining the drivers of the profit and loss distribution.

Our study revealed three main results. First, we bring empirical evidence showing the importance of hybrid (forward-looking) scenarios for comprehensive stress testing. In addition, we show the value added of forward looking over historical scenarios and show numerically the drawbacks of the latter. We confirm the stress testing requirements from Basel III accordingly to which different stress testing approaches cannot be used in isolation, but combined, for a comprehensive picture. Our second finding is that before implementing a stress test a special attention should be given to an accurate model identification for the evolution of returns of portfolio components, where a special attention should be given to time-varibility of correlations and tail dependency to make the stress testing outcome more accurate. In addition, our third finding enhanced the previous findings in Mudry (2013) by disentangling the effects of stressing at one time model parameters for the individual portfolio components versus correlation and their tail dependency. We found clear evidence that the first accounts more than the latter while stress testing the portfolio profit and loss profile. At the same time, our analysis represents an integration of the ”model risk” concept into stress testing exercises, highly relevant for portfolio managers. Special attention should be given to extreme tails, time-varying dependencies, in line with the regulatory frame on stress testing.

Our analysis is bounded by the number of scenarios, and simulations as a generator of random numbers. Our analysis is limited to display tendencies, and the numbers generated can not be transferred directly to risk management. On the other hand, the simulations can form expectations and contribute to a overall understanding of stress testing for capital requirements, liquidity risks management etc. For further analysis we recommend to update our analysis with asymmetric copulas to capture the dependence structure. In addition a more extensive use of hypothetical shocks to a commodity portfolio would be of value.

References

Adams, Z. and Glück, T. (2015) Financialization in commodity markets: A passing trend or the new normal?, *Journal of Banking & Finance*, 60, pp. 93–111. doi: 10.1016/j.jbankfin.2015.07.008.

Aepli, M. (2011) On the design of stress tests, Master's thesis, University of St. Gallen.

Aepli, M. D., Füss, R., Henriksen, T. E. S. and Paraschiv, F. (2017) Modeling the multivariate dynamic dependence structure of commodity futures portfolios, *Journal of Commodity Markets*, 6, pp. 66–87. doi:10.1016/j.jcomm.2017.05.002.

Alexander, C. (2008*a*) *Market Risk Analysis: Quantitative methods in finance*, Vol. 1, England: John Wiley & Sons.

Alexander, C. (2008*b*) *Market Risk Analysis: Practical financial econometrics*, Vol. 2, England: John Wiley & Sons.

Alexander, C. (2008*c*) *Market Risk Analysis: Value at Risk Models*, Vol. 4, England: John Wiley & Sons.

Baffes, J., Kose, M. A., Ohnsorge, F. and Stocker, M. (2015), The great plunge in oil prices: Causes, consequences and policy responses, (policy research note), World bank group, Available at: https://openknowledge.worldbank.org/discover?query=the%20great%20plunge (accessed: 26 April 2018).

Basak, S. and Pavlova, A. (2016) A model of financialization of commodities, *The Journal of Finance*, 71(4), pp. 1511–1556. doi: 10.1111/jofi.12408.

Basel Committee on Banking Supervision (2006), International convergence of capital measurement and capital standards, Technical report, Bank for International Settlements. Available at: https://www.bis.org/publ/bcbs128.htm (accessed: 11 April 2018).

Basel Committee on Banking Supervision (2009), Principles for sound stress testing practices and supervision, Technical report, Bank for International Settlements. Available at: https://www.bis.org/publ/bcbs155.htm (accessed: 11 April 2018).

Bhardwaj, G., Gorton, G. and Rouwenhorst, G. K. (2015), Facts and fantasies about commodity futures ten years later, (working paper), National Bureau of Economic Research, Cambridge. doi: 10.3386/w21243.

Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000) Copulas for finance-a reading guide and some applications, *SSRN Electronic Journal*, pp. 1–65. doi: 10.2139/ssrn.1032533.

Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015) *Time series analysis: forecasting and control*, John Wiley & Sons.

Cheng, I.-H. and Xiong, W. (2014) Financialization of commodity markets, *Annual Review of Financial Economics*, 6(1), pp. 419–441. doi: 10.1146/annurev–financial–110613–034432.

Cheung, C. S. and Miu, P. (2010) Diversification benefits of commodity futures, *Journal of International Financial Markets, Institutions and Money*, 20(5), pp. 451–474. doi: 10.1016/j.intfin.2010.06.003.

Commodity Futures Trading Commission (2008), Staff report on commodity swap dealers & index traders wit commission recommendations, Cftf press rel. #5542-08. available at: https://www.cftc.gov/ (accessed: 05 march 2018).

Daskalaki, C. and Skiadopoulos, G. (2011) Should investors include commodities in their portfolios after all? new evidence, *Journal of Banking & Finance*, 35(10), pp. 2606–2626. doi: 10.1016/j.jbankfin.2011.02.022.

Delle Chiaie, S., Ferrara, L. and Giannone, D. (2017), Common factors of commodity prices, (working paper), European Central Bank, Frankfurt am Main. doi: 10.2866/116302.

Embrechts, P., Lindskog, F. and J. Mcneil, E. (2001) *Handbook of Heavy Tailed Distributions in Finance*, Amsterdam: Elsevier Science B.V.

Embrechts, P., Mcneil, E. and Straumann, D. (1999) Correlation: Pitfalls and alternatives, *RISK*, pp. 69–71. doi: 10.1.1.640.6850.

Embrechts, P., Mikosch, T. and Klüppelberg, C. (1997) *Modelling Extremal Events: For Insurance and Finance*, London: Springer-Verlag.

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, *Econometrica*, 50(4), pp. 987–1007. doi:10.2307/1912773.

Erdős, P. (2012) Have oil and gas prices got separated?, *Energies*, 49, pp. 707–718. doi: 10.1016/j.enpol.2012.07.022.

European Banking Authority (2017), Draft guidelines on institution's stress testing, (consultation paper), European Banking Authority, London. Available at: http://www.eba.europa.eu/regulation-and-policy/supervisory-review-and-evaluation-srep-and-pillar-2/guidelines-on-stress-testing-and-supervisory-stress-testing (accessed: 26 April 2018).

Ghorbel, A. and Souilmi, S. (2014) Risk measurement in commodities markets using conditional extreme value theory, *International Journal of Econometrics and Financial Management*, 2(4), pp. 188–205. doi: 10.12691/ijefm–2–5–4.

Gorton, G. and Rouwenhorst, K. G. (2006) Facts and fantasies about commodity futures, *Financial Analysts Journal*, 62(2), pp. 47–68. doi: 10.2469/faj.v62.n2.4083.

He, Y., Wang, S. and Lai, K. K. (2010) Global economic activity and crude oil prices: A cointegration analysis, *Energy Economics*, 32(4), pp. 868–876. doi: 10.1016/j.eneco.2009.12.005.

Kilian, L. and Zhou, X. (2018) Modeling fluctuations in the global demand for commodities, , pp. 0 – 60. Munich: Center for Economic Studies and Ifo Institute (CESifo). Available at: http://hdl.handle.net/10419/173025 (accessed: 11 February 2018).

Koliai, L. (2016) Extreme risk modeling: An evt–pair-copulas approach for financial stress tests, *Journal of Banking & Finance*, 70, pp. 1 – 22. doi: 10.1016/j.jbankfin.2016.02.004.

Liu, J. (2011) Extreme value theory and copula theory: A risk management application with energy futures, Master's thesis, University of Victoria.

Lopez, J. A. (2005) Stress tests: useful complements to financial risk models, *FRBSF Economic Letter*, pp. 1 – 4. Available at: https://www.frbsf.org/economic–research/files/el2005–14.pdf (accessed: 20 March 2018).

McNeil, A. J., Frey, R. and Embrechts, P. (2015) *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton: Princeton University Press.

Mudry, P.-A. (2013) Stress-testing technics for commodities futures portfolios, using extreme-value theory and copulas, Master's thesis, University of St. Gallen.

Nick, S. and Thoenes, S. (2014) What drives natural gas prices?—a structural var approach, *Energy Economics*, 45, pp. 517–527. doi: 10.1016/j.eneco.2014.08.010.

Nyström, K. and Skoglund, J. (2002*a*) A framework for scenario-based risk management, , pp. 1–56. Swedbank, Group Financial Risk Control. Available at: http://gloria–mundi.com/UploadFile/2010–2/knjs.pdf (accessed: 01 March 2018).

Nyström, K. and Skoglund, J. (2002*b*) Univariate Extreme Value Theory, GARCH and Measures of Risk, , pp. 1 – 27. Swedbank, Group Financial Risk Control. Available at: http://gloria–mundi.com/UploadFile/2010–2/knjsug.pdf (accessed: 14 February 2018).

Öztek, M. F. and Öcal, N. (2017) Financial crises and the nature of correlation between commodity and stock markets, *International Review of Economics & Finance*, 48, pp. 56–68. doi: 10.1016/j.iref.2016.11.008.

Paraschiv, F., Mudry, P.-A. and Andries, A. M. (2015) Stress-testing for portfolios of commodity futures, *Economic Modelling*, 50, pp. 9–18. doi: 10.1016/j.econmod.2015.06.005.

Pindyck, R. S. (2004) Volatility and commodity price dynamics, *Journal of Futures Markets*, 24(11), pp. 1029–1047. doi: 10.1002/fut.20120.

Silvennoinen, A. and Thorp, S. (2013) Financialization, crisis and commodity correlation dynamics, *Journal of International Financial Markets, Institutions and Money*, 24, pp. 42–65. doi: 10.1016/j.intfin.2012.11.007.

Sklar, M. (1959) *Fonctions de Répartition À N Dimensions Et Leurs Marges*, Paris: Université Paris 8.

S&P Dow Jones Indices (2016), S&P Dow Jones Indices announces 2017 weights for the Dow Jones Commodity Index, Press release, S&P Global. Available at: https://us.spindices.com/search/?query=2017+weights&Search=Go&Search=Go (accessed: 20 January 2018).

S&P Dow Jones Indices (2017), Dow Jones Commodity Index Methodology, Index methodology, S&P Global. Available at: https://us.spindices.com/documents/methodologies/methodology-dj-commodity-index.pdf (accessed: 12 february 2018).

S&P Dow Jones Indices (2018), S&P GSCI Methodology, Index methodology, S&P Global. Available at: https://us.spindices.com/indices/commodities/sp-gsci (accessed: 12 february 2018).

Stoll, H. R. and Whaley, R. E. (2015) Commodity index investing and commodity futures prices, *Journal of applied finance*, 20(1), pp. 7 – 46. Available at: https://ssrn.com/abstract=2693084 (accessed: 30 March 2018).

Tang, K. and Xiong, W. (2010) Index investing and financialization of commodities, , pp. 1 – 50. doi: 10.3386/w16385.

Tang, K. and Xiong, W. (2012) Index investment and the financialization of commodities, *Financial Analysts Journal*, 68(6), pp. 54–74. doi: 10.2469/faj.v68.n6.5.

Thomson Reuters (2012), Thomson reuters indices: Continuous commodity total return indices, Methodology report, Thomson Reuters. Available at: https://www.thomsonreuters.com (accessed: 25 January 2018).

Tyner, W. E. (2010) The integration of energy and agricultural markets, *Agricultural Economics*, 41(1), pp. 193–201. doi: 10.1111/j.1574–0862.2010.00500.x.

Wang, Z.-R., Chen, X.-H., Jin, Y.-B. and Zhou, Y.-J. (2010) Estimating risk of foreign exchange portfolio: Using var and cvar based on garch–evt-copula model, *Physica A: Statistical Mechanics and its Applications*, 389(21), pp. 4918–4928. doi: 10.1016/j.physa.2010.07.012.

Appendices

Table A1: Comparison of ML estimators for the GPD parameters for different thresholds (u) for years 1996-2017.

u = 7%	ξ		β	
Commodity	Upper tail	Lower Tail	Upper tail	Lower Tail
Wheat	0.082	0.114	0.566	0.477
Corn	0.188	0.233	0.564	0.456
Soybeans	0.082	0.071	0.540	0.645
Live cattle	0.278	0.313	0.516	0.560
Copper	-0.007	0.116	0.551	0.591
Gold	0.201	0.124	0.488	0.609
Aluminum	-0.037	0.148	0.592	0.494
WTI	0.065	0.075	0.466	0.624
Brent	-0.002	0.109	0.527	0.562
Natural gas	0.158	0.123	0.531	0.473

u = 10%	ξ		β	
Commodity	Upper tail	Lower Tail	Upper tail	Lower Tail
Wheat	0.022	0.082	0.616	0.494
Corn	0.178	0.156	0.540	0.508
Soybeans	0.050	0.101	0.559	0.592
Live cattle	0.285	0.251	0.453	0.570
Copper	-0.029	0.118	0.582	0.565
Gold	0.145	0.100	0.520	0.612
Aluminum	0.002	0.100	0.555	0.520
WTI	0.010	0.093	0.517	0.581
Brent	0.009	0.066	0.515	0.597
Natural gas	0.073	0.071	0.613	0.519

u = 11%	ξ		β	
Commodity	Upper tail	Lower Tail	Upper tail	Lower Tail
Wheat	0.008	0.092	0.632	0.480
Corn	0.155	0.152	0.557	0.504
Soybeans	0.043	0.098	0.564	0.591
Live cattle	0.270	0.253	0.454	0.553
Copper	-0.057	0.102	0.620	0.577
Gold	0.148	0.103	0.507	0.602
Aluminum	0.012	0.089	0.546	0.528
WTI	-0.005	0.098	0.534	0.568
Brent	-0.010	0.063	0.535	0.597
Natural gas	0.069	0.062	0.614	0.529

u = 12%	ξ		β	
Commodity	Upper tail	Lower Tail	Upper tail	Lower Tail
Wheat	0.006	0.075	0.634	0.493
Corn	0.132	0.138	0.580	0.514
Soybeans	0.036	0.096	0.571	0.588
Live cattle	0.271	0.266	0.440	0.525
Copper	-0.056	0.097	0.621	0.578
Gold	0.141	0.073	0.508	0.636
Aluminum	-0.028	0.070	0.587	0.549
WTI	-0.016	0.077	0.547	0.588
Brent	0.001	0.044	0.523	0.621
Natural gas	0.087	0.060	0.586	0.528

	ξ		β	
Commodity	Upper tail	Lower Tail	Upper tail	Lower Tail
Wheat	0.0452	0.0021	0.4820	0.6991
Corn	-0.1141	-0.0896	0.4715	0.5985
Soybeans	-0.0309	-0.1642	0.5149	0.7738
Live Cattle	0.2708	0.2140	0.4478	0.5416
Copper	-0.5240	-0.0681	0.9359	0.5994
Gold	0.1864	-0.0017	0.3939	0.7062
Aluminium	0.1976	-0.1715	0.4346	0.6775
WTI	-0.1542	-0.1232	0.5696	0.5557
Brent	-0.1876	-0.1805	0.5822	0.5868
Natural Gas	0.1348	0.0107	0.5633	0.5841

Table A2: Recalibrated Maximum Likelihood estimators for the generalized Pareto distribution parameters for the time of financial distress, years 2007-2008. Threshold: 10%.

Commodity	Wheat	Corn	Soybeans	Live Cattle	Copper	Gold	Aluminium	WTI	Brent	Natural Gas
Wheat	1.0000	0.6217	0.4768	0.1259	0.1251	0.1221	0.1048	0.1468	0.1393	0.0793
Corn	0.6217	1.0000	0.6215	0.1508	0.1556	0.1393	0.1316	0.1689	0.1634	0.1071
Soybeans	0.4768	0.6215	1.0000	0.1462	0.2035	0.1633	0.1676	0.1995	0.1971	0.1144
Live Cattle	0.1259	0.1508	0.1462	1.0000	0.0983	0.0450	0.0907	0.1095	0.1035	0.0493
Copper	0.1251	0.1556	0.2035	0.0983	1.0000	0.2637	0.5829	0.2641	0.2496	0.0668
Gold	0.1221	0.1393	0.1633	0.0450	0.2637	1.0000	0.2280	0.2057	0.1979	0.0735
Aluminium	0.1048	0.1316	0.1676	0.0907	0.5829	0.2280	1.0000	0.2156	0.2023	0.0714
WTI	0.1468	0.1689	0.1995	0.1095	0.2641	0.2057	0.2156	1.0000	0.9083	0.2566
Brent	0.1393	0.1634	0.1971	0.1035	0.2496	0.1979	0.2023	0.9083	1.0000	0.2368
Natural Gas	0.0793	0.1071	0.1144	0.0493	0.0668	0.0735	0.0714	0.2566	0.2368	1.0000

Table A3: Correlation matrix for the baseline, years 1996-2017. Light blue is correlation between 0.15 and 0.19. Dark blue is correlation larger than 0.20.

Commodity	Wheat	Corn	Soybeans	Live Cattle	Copper	Gold	Aluminium	WTI	Brent	Natural Gas
Wheat	1.0000	0.5875	0.4854	0.2443	0.2885	0.2738	0.2041	0.3090	0.3080	0.1372
Corn	0.5875	1.0000	0.6777	0.3050	0.3117	0.2956	0.2885	0.3870	0.3790	0.2330
Soybeans	0.4854	0.6777	1.0000	0.3259	0.3670	0.3484	0.2913	0.4663	0.4586	0.2391
Live Cattle	0.2443	0.3050	0.3259	1.0000	0.2326	0.1777	0.2067	0.2641	0.2587	0.1902
Copper	0.2885	0.3117	0.3670	0.2326	1.0000	0.4725	0.6412	0.4256	0.4127	0.1594
Gold	0.2738	0.2956	0.3484	0.1777	0.4725	1.0000	0.3876	0.4654	0.4561	0.2038
Aluminium	0.2041	0.2885	0.2913	0.2067	0.6412	0.3876	1.0000	0.3729	0.3694	0.1737
WTI	0.3090	0.3870	0.4663	0.2641	0.4256	0.4654	0.3729	1.0000	0.9833	0.4473
Brent	0.3080	0.3790	0.4586	0.2587	0.4127	0.4561	0.3694	0.9833	1.0000	0.4326
Natural Gas	0.1372	0.2330	0.2391	0.1902	0.1594	0.2038	0.1737	0.4473	0.4326	1.0000

Table A4: Correlation matrix for scenarios re-calibrated on time of financial distress, years 2007-2008. Light blue is correlation between 0.15 and 0.19. Dark blue is correlation larger than 0.20.

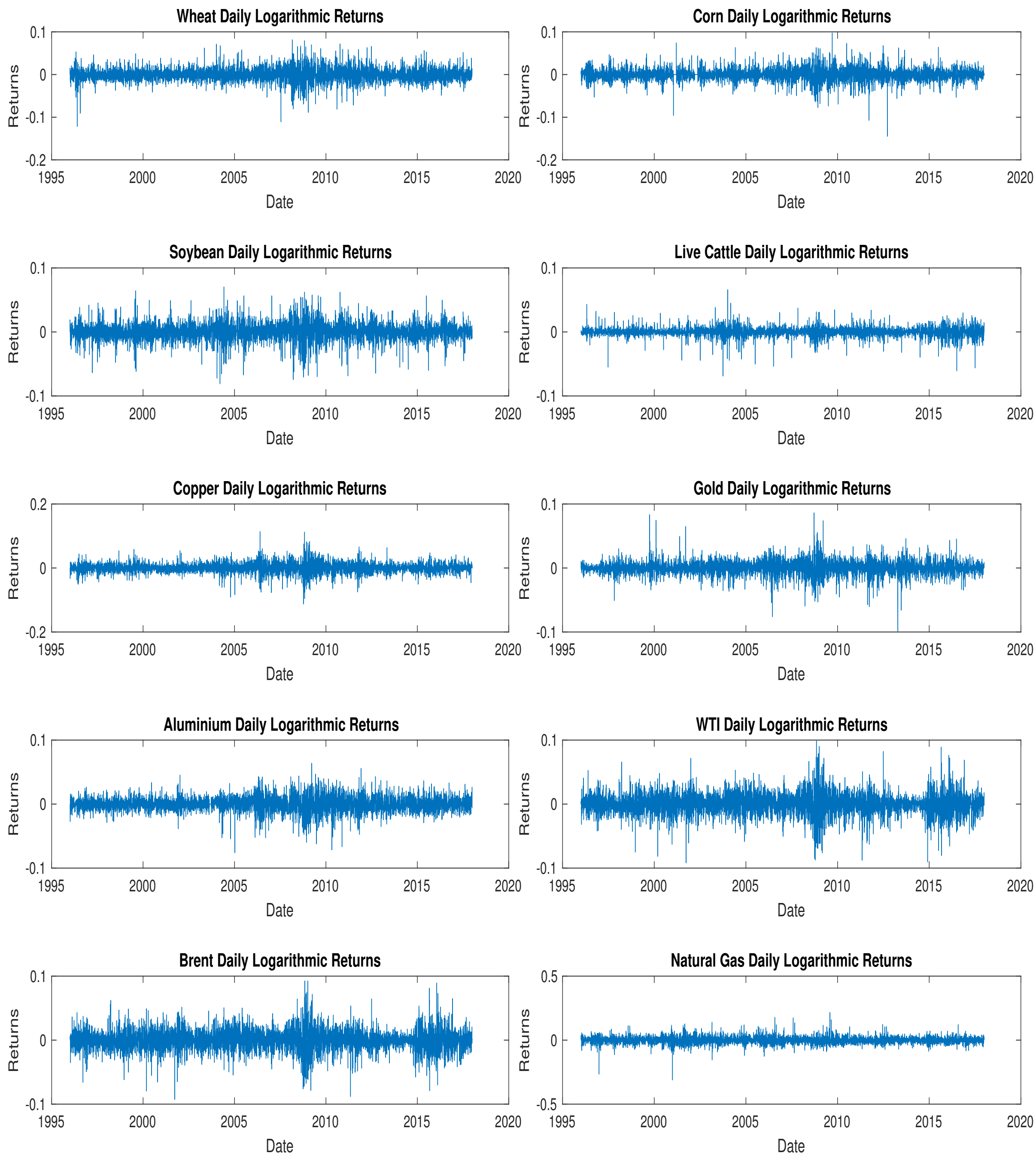

Figure A1: Daily Logarithmic Returns.

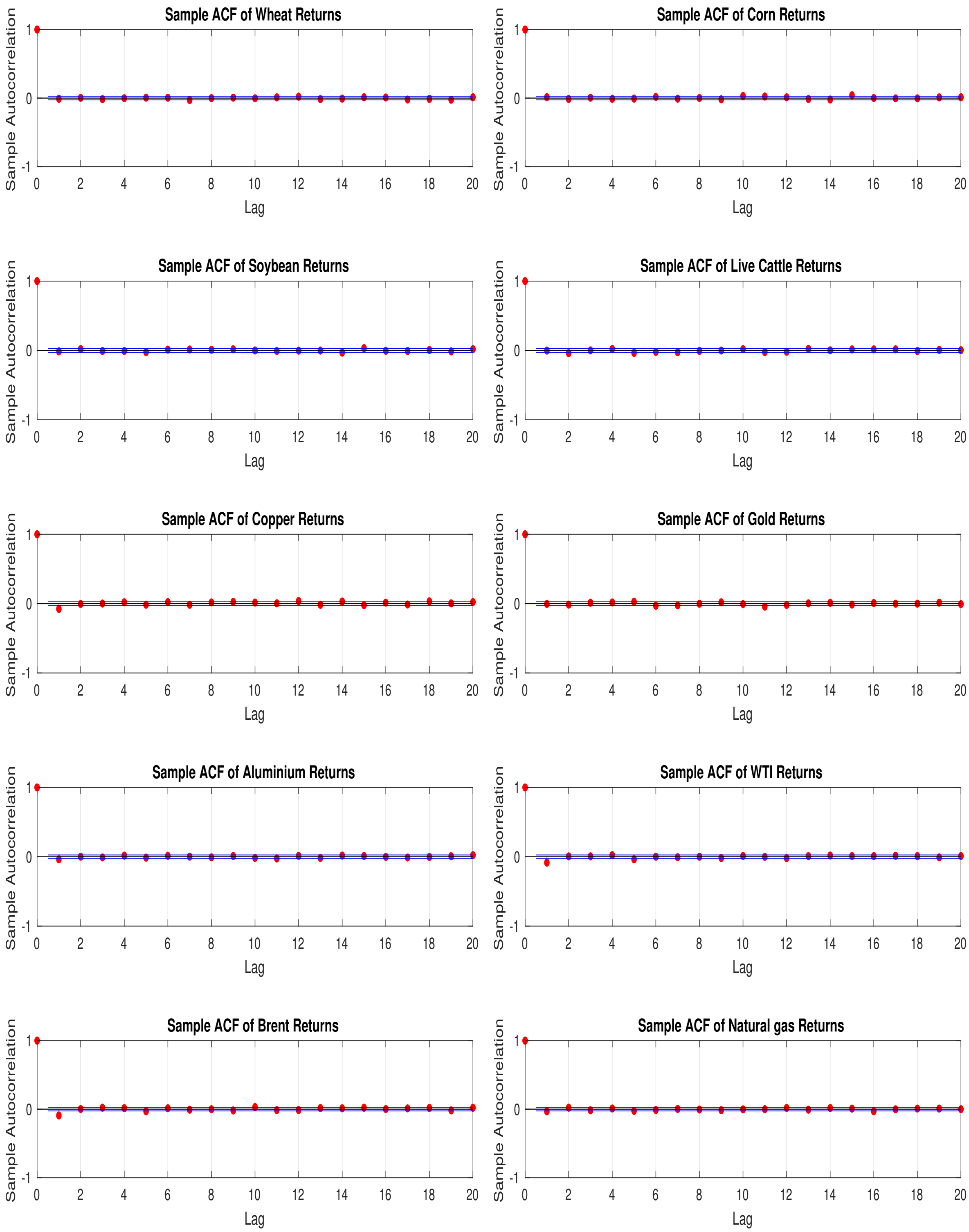

Figure A2: Sample ACF of Returns.

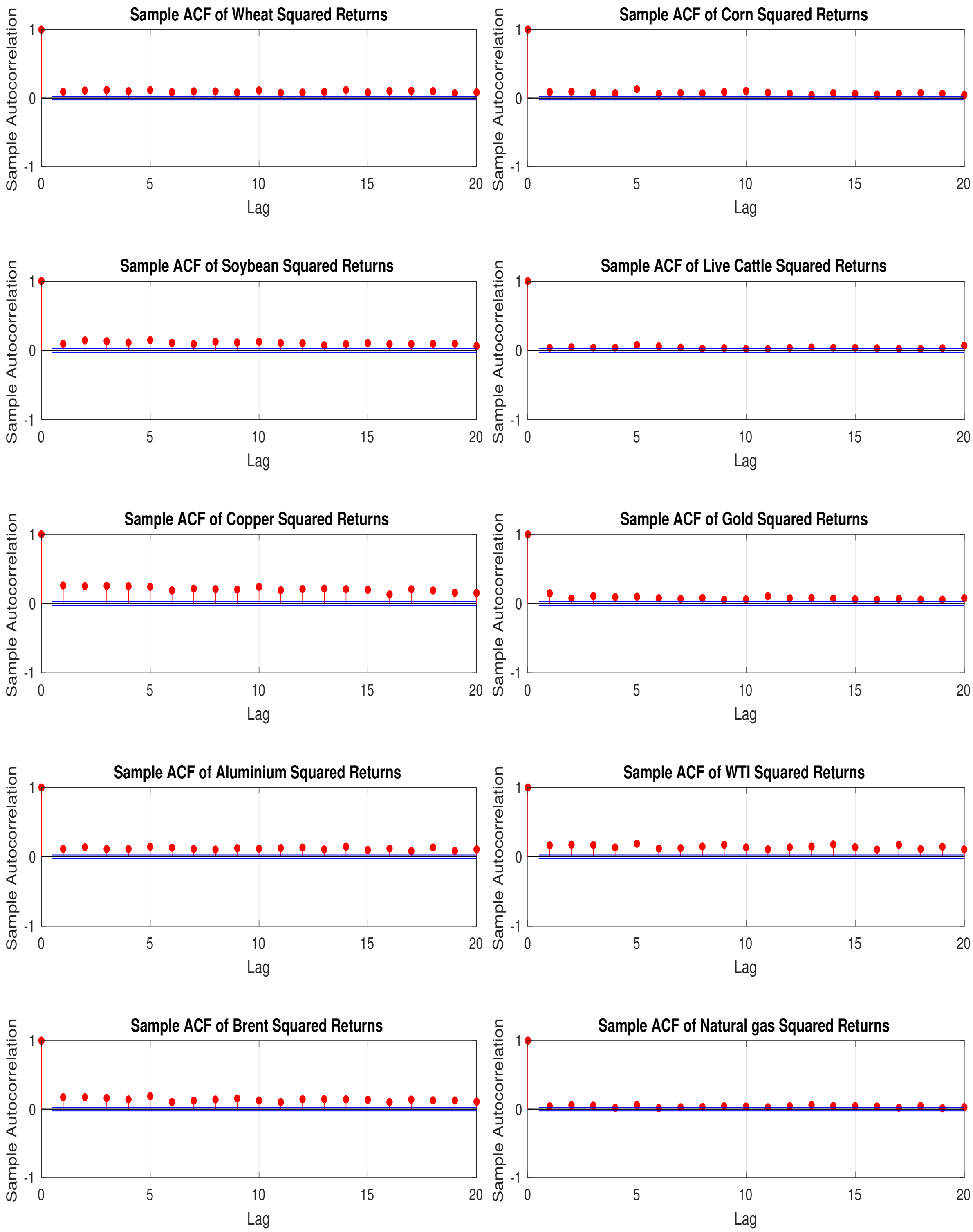

Figure A3: Sample ACF of Squared Returns.

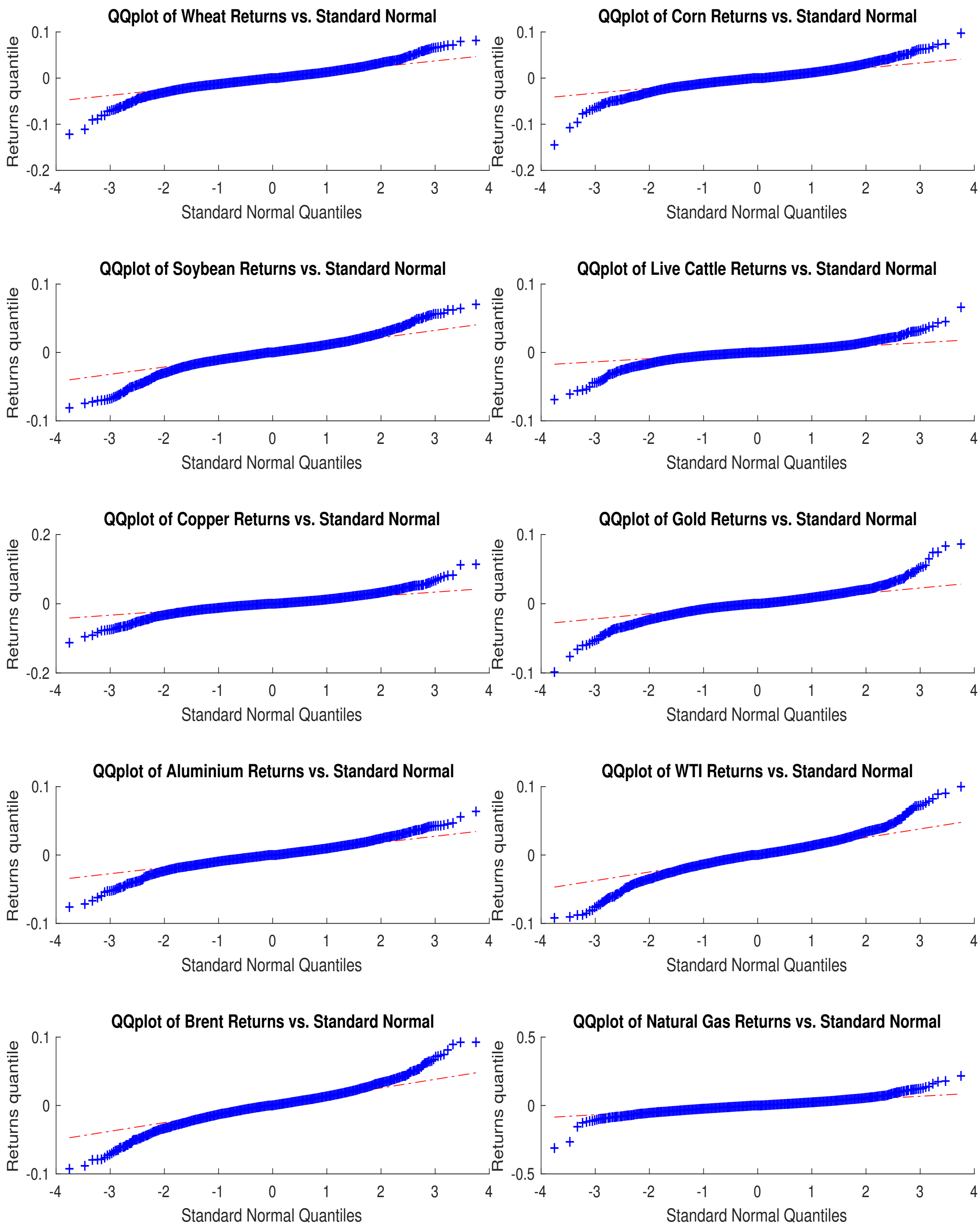

Figure A4: QQ Plot of Returns vs Standard Normal Distribution.

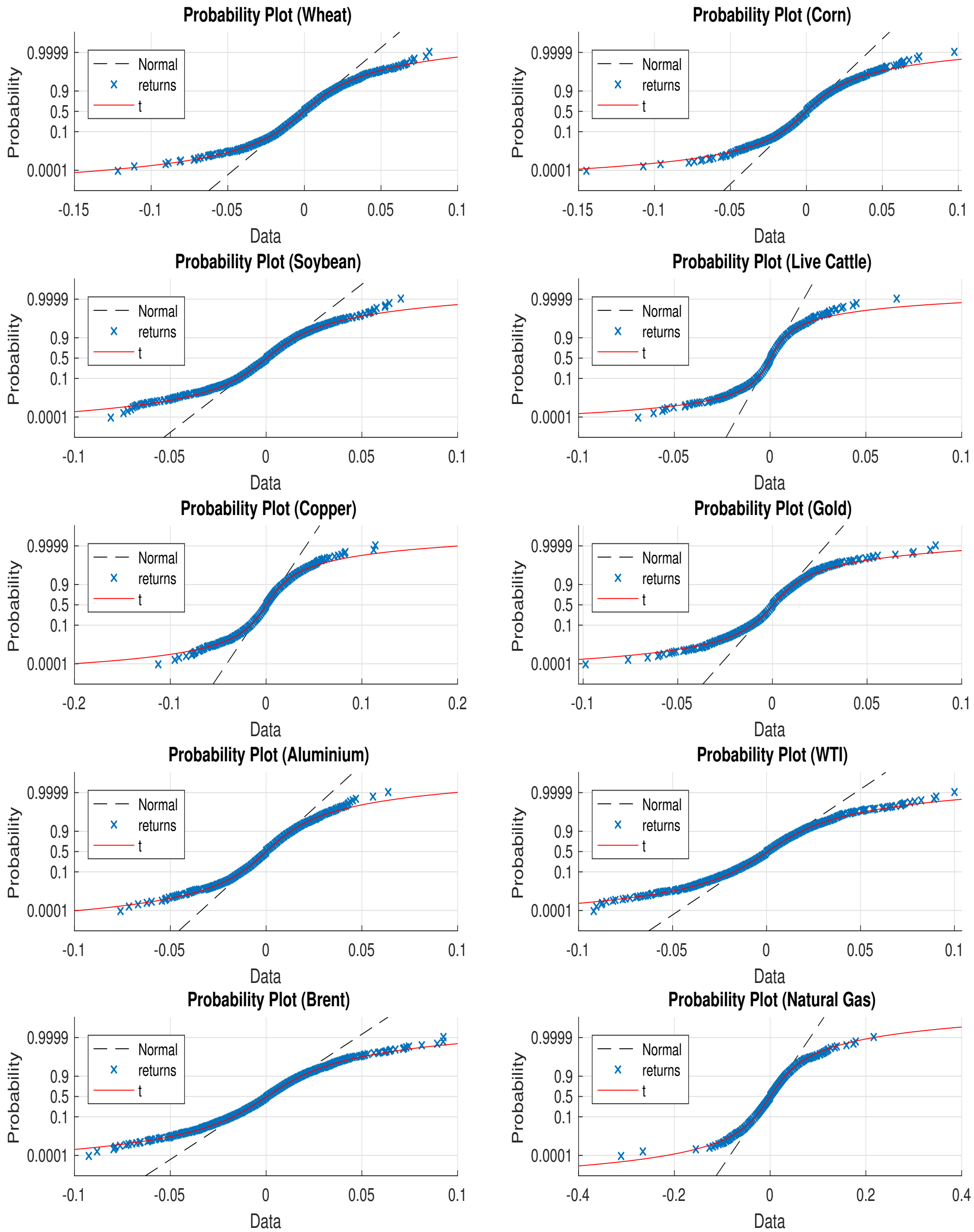

Figure A5: Probability Plots of Returns vs Standard Normal vs t distributions.

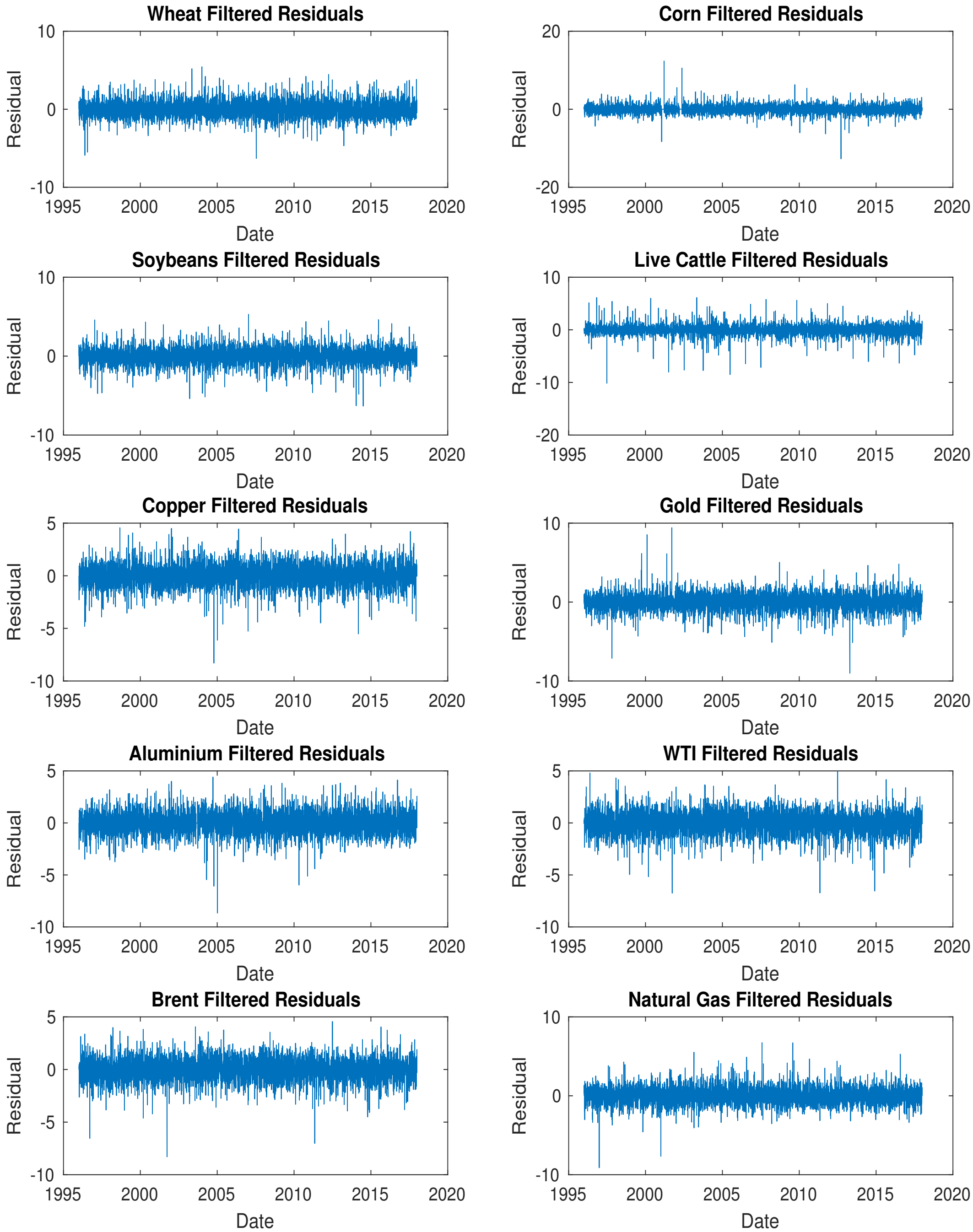

Figure A6: Filtered Residuals.

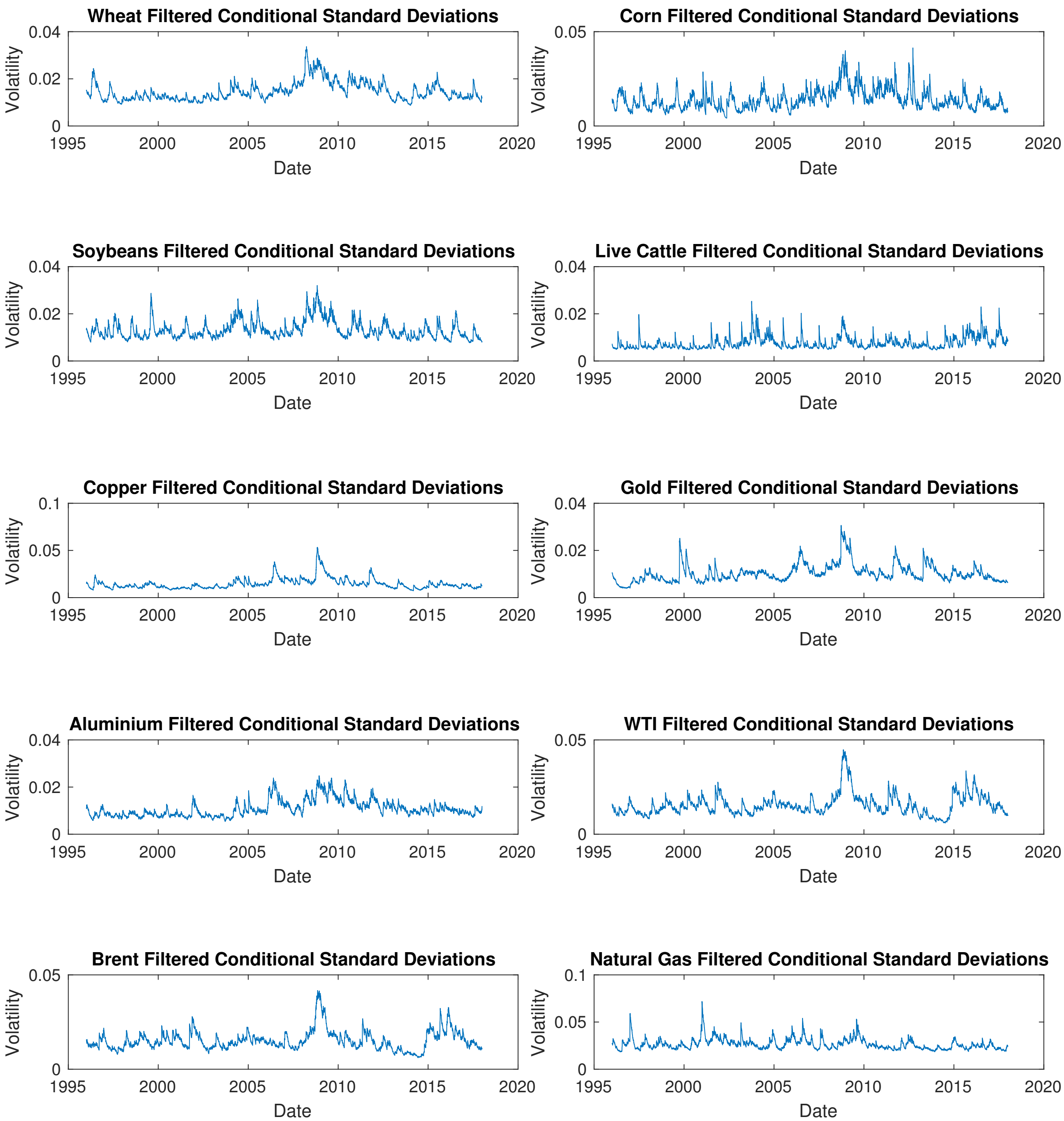

Figure A7: Filtered Conditional Standard Deviations.

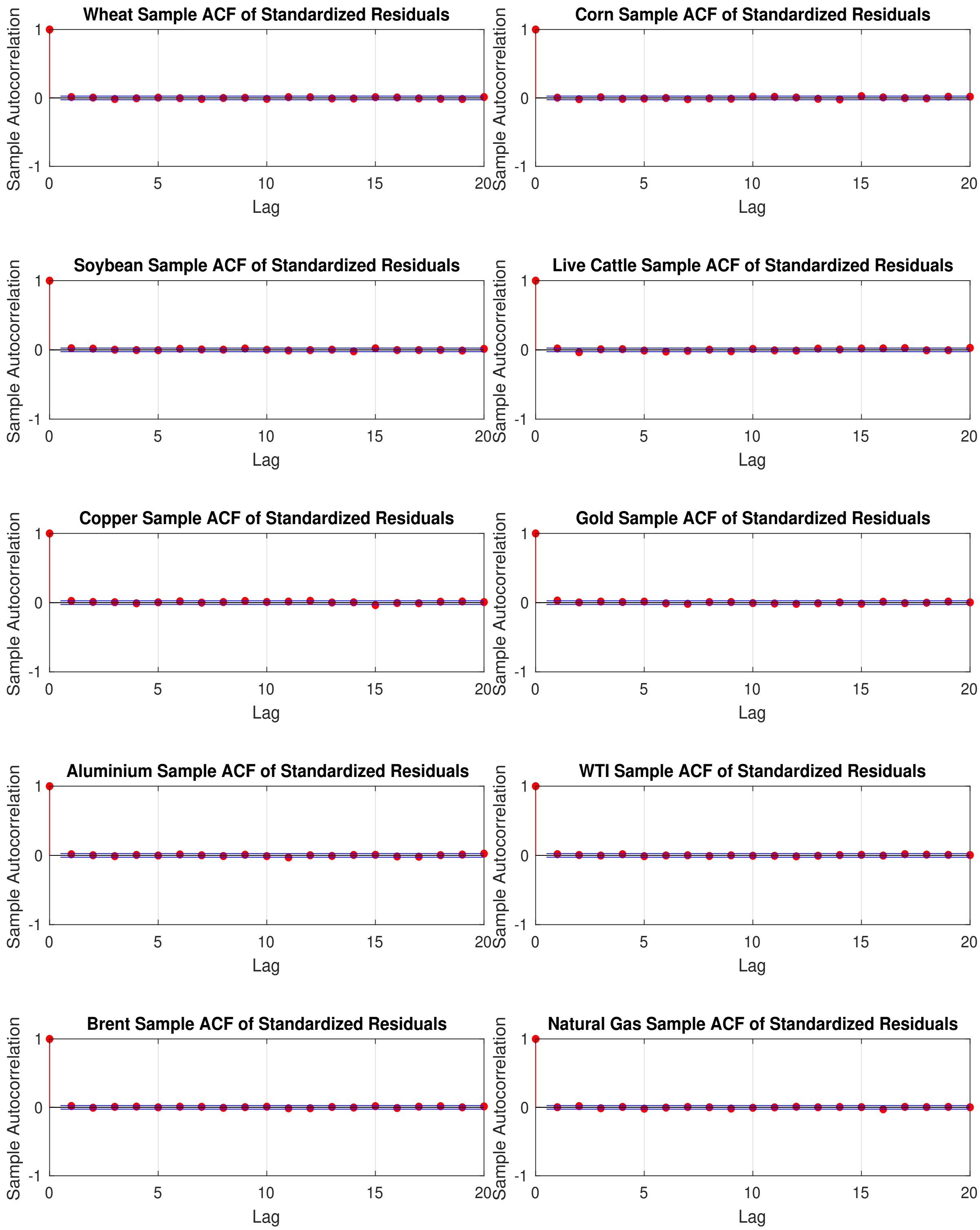

Figure A8: Sample ACF of Standardized Residuals.

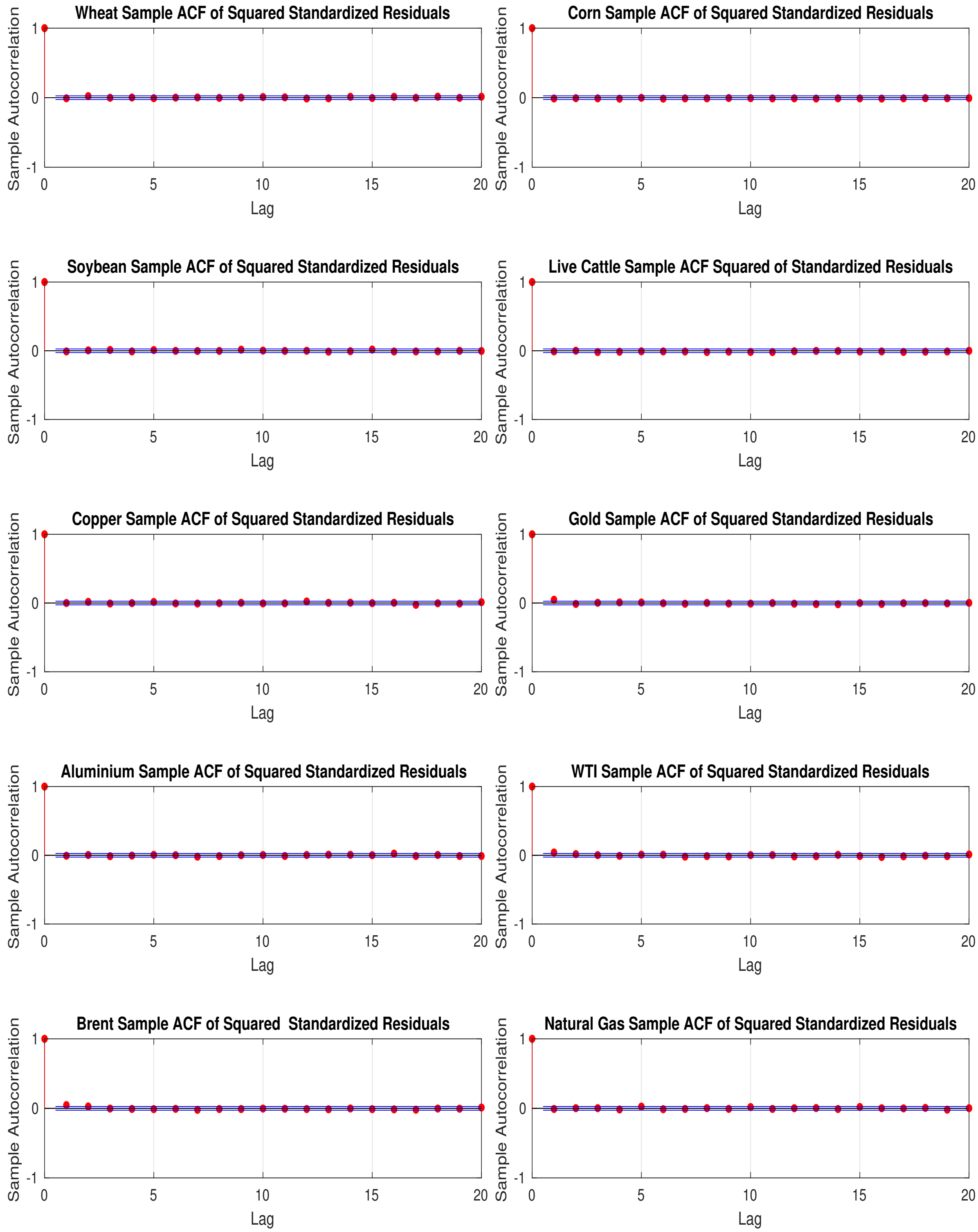

Figure A9: Sample ACF of Squared Standardized Residuals.

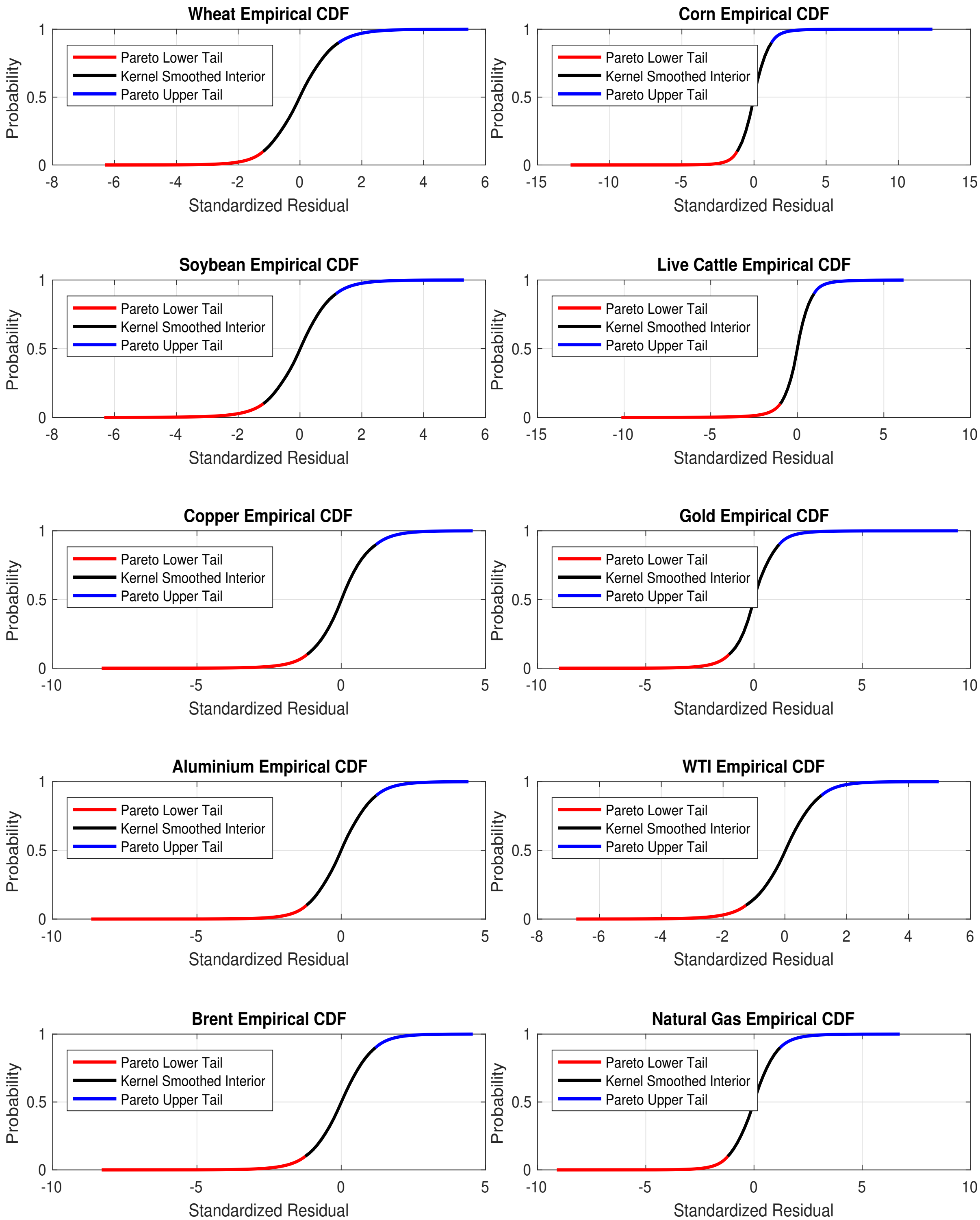

Figure A10: Semi-parametric Empirical CDFs.

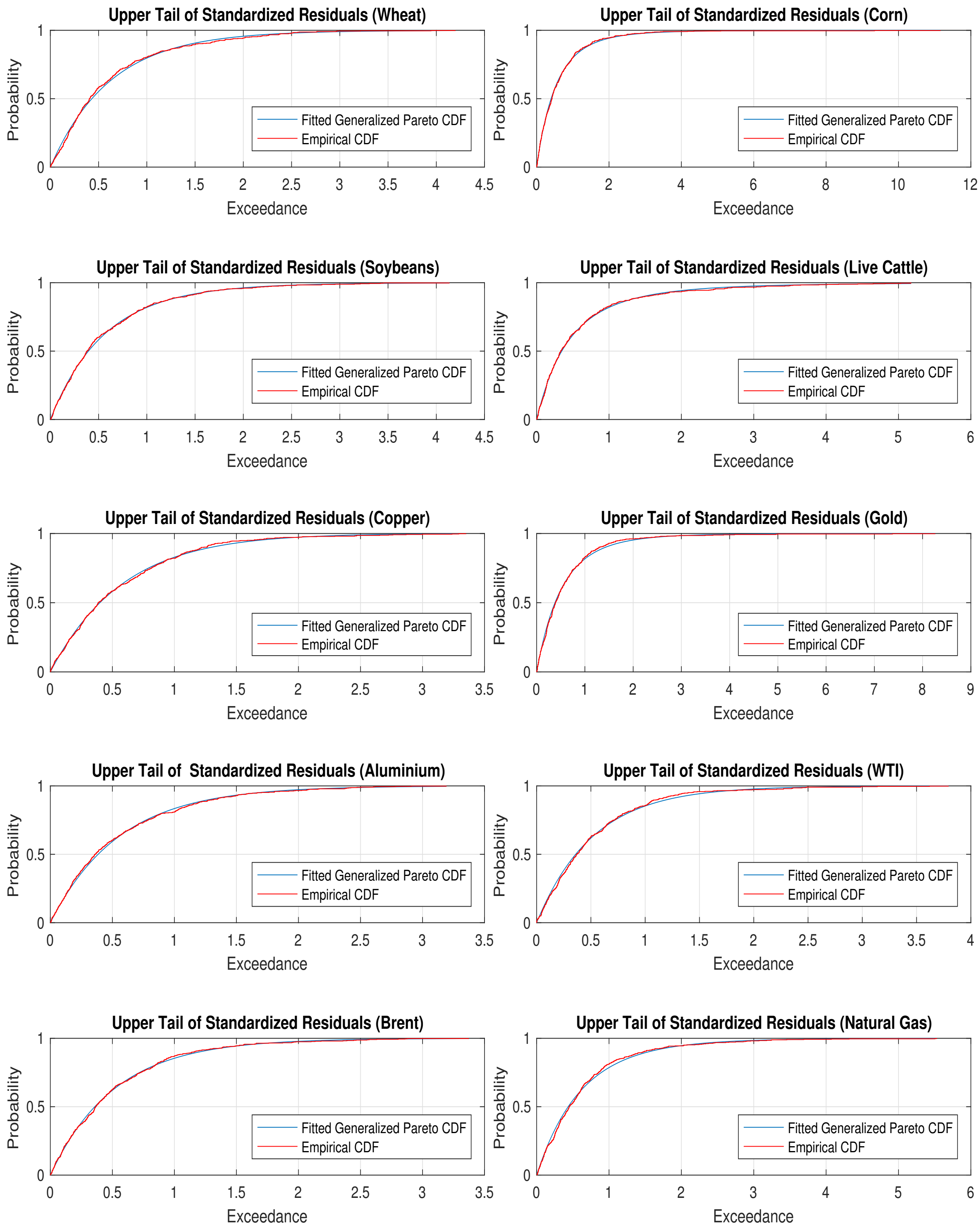

Figure A11: Fitted vs empirical upper tails of the standardized residuals.