Local Gaussian Correlation

Charalampos Kochaimidis

Supervisor: Theodore Simos

Department of Economics

University of Ioannina

Greece

January, 2024

Contents

List of Figures

List of Tables

Acknowledgement

I would like to express my sincere gratitude to my supervisor Theodore Simos. I would, also like to thank my girlfriend and my family for all the continuous support.

Abstract

Many experts in finance and econometrics believe that when the market goes down, the connection between financial assets gets stronger. They think that the correlation almost reaches one during a market crash, which can mess up the advantages of diversification. The purpose of this Thesis is to explore the relationship between markets, assets and commodities by employing a local dependence measure, which provides a precise mathematical description and interpretation of these interactions. We utilize a relatively recent local dependence measure introduced by Dag Tjøstheim in 2013[1] – the local correlation function. This function involves approximating a bivariate density locally through a family of bivariate Gaussian densities using local likelihood. The local correlation is determined at each point by considering the correlation coefficient of the approximating Gaussian distribution. This measure is referred to as Local Gaussian Correlation.

Keywords: Local Dependence Measure, Local Correlation Function, Bivariate Gaussian Densities, Local Gaussian Correlation

Chapter 1

Introduction

Over the last few decades, global financial markets have become increasingly interconnected. A notable focus lies in the spread of crises across these markets, wherein substantial declines in asset values in one country prompt swift declines in other countries. If these falls cannot be accounted for by interdependence or shared macroeconomic factors, it is termed contagion. This phenomenon holds significance for risk management and the performance of international portfolios. A heightened interdependence in financial markets during a crisis suggests that the expected diversification effect may be less pronounced.

The core concept of the Local Gaussian Correlation approach involves estimating a bivariate return distribution through a set of Gaussian bivariate distributions. For each point within the return distribution, a specific Gaussian distribution provides a well-fitting approximation. The correlation of the approximating Gaussian distribution is then considered as the local correlation within that specific neighborhood. This leads to a nonlinear dependence measure that is inherently localized.

In addition the conventional global correlation analysis assumes a Gaussian distribution for financial returns, a local assumption in our approach. Moreover, the local Gaussian correlation avoids the bias problem associated with the conditional correlation which is created when we condition the correlation by focusing on a subset of the sample space and makes difficult the interpretation. Another benefit of employing local Gaussian correlation lies in its ability to identify more intricate, nonlinear changes in the dependence structure that may be obscured by global correlation measures. Consequently, it provides a more comprehensive understanding of the interdependence between markets, particularly in the tails of the distribution and other segments, as opposed to the potential masking effect of global correlation. This contrasts with typical tests for contagion. It's worth noting that while the copula method can uncover nonlinear dependence and tail-dependence, it often involves parameters with indirect interpretations as measures of dependence. In contrast, our procedure has a more intuitive foundation and maintains a correlation interpretation based on local Gaussian approximation.

The objective of this thesis is to investigate the connection among markets, assets and commodities using a local dependence measure. This measure aims to offer a precise mathematical description and interpretation of the interactions between these assets.

Chapter 2

Literature Review

Dag Tjøstheim and Karl Ove Hufthammer (2013) [1] introduced a novel local dependence measure, providing a precise mathematical characterization and interpretation of phenomena like the correlation between financial assets. Their proposed methodology centers around a new local correlation function, which relies on the local likelihood approximation of a bivariate density through a family of bivariate Gaussian densities. The correlation coefficient of the approximating Gaussian distribution at each point serves as the local correlation.This approach establishes the existence, uniqueness, and limit results of the local correlation function. Additionally, they presented various properties associated with the local Gaussian correlation and its estimation. The methodology is illustrated through examples drawn from both simulated and real data.The significance of this research lies in its potential to extend the capability of locally modeling a general density, a task traditionally achieved globally for the Gaussian density. The utilization of local dependence measures contributes to a nuanced understanding of the intricacies within data, particularly in the context of bivariate density estimation.

Georgios Bampinas and Theodore Panagiotidis (2017)[2] conducted an empirical analysis to assess the influence of financial shocks on the interconnections between oil prices and stock markets during four major crises. Employing the local Gaussian correlation method, their study uncovered a noteworthy observation: a regionalization of the two markets prevailed for a substantial duration throughout the 1990s and the early 2000s.The central emphasis of their research was on elucidating the nuanced, nonlinear dependence dynamics between stock and oil markets, specifically within the distinct contexts of the Mexican "Tequila" crisis, the Asian "flu" crisis, the dot-com bubble, and the 2007–2009 financial crisis.

Quynh Nga Nguyen, Sofiane Aboura, Julien Chevallier, Zhang Lyuyuan, and Bangzhu Zhu (2020)[3] conducted a comprehensive examination of the escalating correlations within commodity and U.S. financial markets, as well as among various commodity markets, spanning the years 1992 to 2017. Employing a non-linear framework, their investigation aimed to assess the enduring nature of the financialization phenomenon, taking into account significant events that shaped the landscape of the 2000s.Utilizing a measure of asymmetric dependence, specifically the local Gaussian correlation, the researchers sought to quantify the relationships under consideration. The findings of their study empirically support the existence of the financialization phenomenon, with a particular emphasis on heightened correlations between stock markets and commodity markets, notably following the break date in August 2008.

Håkon Otneim (2021)[4] provided an in-depth exposition on the lg package, delineating its fundamental principles and practical utility. Tailored for the R programming language, the lg package stands as an instrumental resource, incorporating recent methodological advancements in the realm of local Gaussian correlation applications. This encompasses the accurate estimation of the local Gaussian correlation itself, multivariate density estimation, conditional density estimation, a diverse array of tests for independence and conditional independence, and a graphical module specifically crafted for the creation of dependence maps. In essence, the lg package offers a comprehensive toolkit within the R environment, facilitating a broad spectrum of analyses pertaining to local Gaussian correlation.

Geir Drage Berentsen, Tore Selland Kleppe and Dag Tjøstheim (2014)[5] introduced in their paper the R package "localgauss," a robust tool designed to estimate and visually represent a significant measure of localized dependence known as local

Gaussian correlation. The package encompasses a comprehensive set of functionalities, including a dedicated function for precise estimation, a specialized function for conducting local independence tests, and additional functions tailored for visualization purposes. These features are thoughtfully demonstrated through a series of illustrative examples, showcasing the versatility and effectiveness of the package in various scenarios.

Dag Tjøstheim and Karl Ove Hufthammer (2014) [6] researched the phenomenon of financial contagion, specifically investigating whether the interconnections across different financial markets intensify following a shock to a particular country. To explore the contagion effect, they employed the novel metric of local dependence, as introduced by Tjøstheim and Hufthammer (2013). The core concept of this innovative approach revolves around approximating any arbitrary bivariate return distribution through a family of Gaussian bivariate distributions. At each point within the return distribution, a Gaussian distribution is utilized to provide an accurate approximation for that specific point. The local correlation within a given neighborhood is determined by the correlation of the approximating Gaussian distribution. Through a meticulous analysis of the local Gaussian correlation before the shock and after the shock, they employed a bootstrap testing procedure to assess whether contagion has occurred. To illustrate the efficacy of our methodology, they re-evaluate notable events such as the Mexican crisis of 1994, the Asian crisis of 1997-1998, and the financial crisis of 2007-2009. Their findings, derived from this new analytical approach, reveal compelling evidence of contagion and offer insights into the nonlinear dependence structure characterizing these crises.

Chapter 3

Model

This article utilizes the novel approach of the local Gaussian correlation dependence technique, as introduced in the work by Tjøstheim and Hufthammer in 2013. [1] The core concept behind this novel approach is to model any arbitrary bivariate return distribution by using a set of Gaussian bivariate distributions. Specifically, at every location within the return distribution, a Gaussian distribution is employed to approximate that particular point. This approach focuses on local density approximation rather than directly modeling the correlation. The correlation of the approximating Gaussian distribution is determined by the local correlation within that particular neighborhood.

The general bivariate density f for the variables (Xt , Yt), as it is proposed by Quynh Nga Nguyen, et al.[3], can be approximated locally in the neighborhood of each point z = (x, y) by a Gaussian bivariate density defined by:

$$\phi(u, v, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}(\frac{u-\mu_1}{\sigma_1})^2+(\frac{v-\mu_2}{\sigma_2})^2-2\rho(\frac{u-\mu_1}{\sigma_1})(\frac{v-\mu_2}{\sigma_2})\}$$

where $v = (v_1, v_2)^T$ is the running variable in the Gaussian distribution, $\mu_i(z)$, $i = 1, 2$, are the local means, $\sigma_i(z)$, $i = 1, 2$, are the local standard deviations, and $\rho(z)$ is the local correlation at the point $z = (x, y)$.

The population values of the local parameters $\theta_b(z) = \theta(z) = (\mu_1(z), \mu_2(z), \sigma_1(z), \sigma_2(z), \rho(z))$ are obtained by minimizing the local penalty function

$$q = \int_{K_b}(v - z)[\phi(v, \theta(z)) - \log \phi(v, \theta(z))f(v)\, dv.$$

where $K_b(v - z) = (b_1 b_2)^{-1} K_{b_1}^{-1}(v_1 - z_1) K_{b_2}^{-1}(v_2 - z_2)$ is a product kernel with bandwidth $b = (b_1, b_2)$. The local Gaussian correlation $\rho_b(z) = \rho(z)$ is defined as the last element of the vector $\theta(z)$ that minimizes q.

3.1 Methods

3.1.1 Function localgauss

In the R package localgauss, the localgauss() function employs a modified Newton's method with line-search to maximize the local likelihood function for various values of x. The result is an S3 object classified as localgauss. Given that the local likelihood function is not universally concave, eigenvalue modification is incorporated to ensure the positive definiteness of the scaling matrix (Nocedal and Wright 1999 [7]). Both the optimizer and objective function are implemented in Fortran 90 (Metcalf and Reid 1999 [8]), and the source code for the gradient and Hessian of the local likelihood function is generated using the automatic differentiation tool TAPENADE (Hascoët and Pascual 2004 [9]).

The user can manually specify the number M of points $x = (x_1, x_2)$ for estimating the local Gaussian correlation using the argument xy.mat, which is an M times 2 matrix. Alternatively, the selection of these points can be conducted through the methodology introduced by Jones and Koch (2003) [10]. This approach involves placing a regular N $\times$ N grid across the area of interest, which is then screened by selecting grid points $x_1, ..., x_M$ satisfying $\hat{f}(x_j) \geq C$, for some constant C, and a density estimator $\hat{f}$. The screening process is efficiently implemented using the R package MASS (Venables and Ripley 2002 [11]). In localgauss(), the values of N and C are set by the arguments gsize and hthresh, respectively. If xy.mat is not explicitly specified, it will be internally selected through the described method. As the local likelihood function necessitates optimization for each estimation point, the computational time scales proportionally with the size of xy.mat.

3.1.2 Graphics

Facilitating a thorough understanding and polished presentation of the outcomes derived from the function localgauss() as elucidated earlier necessitates the strategic integration of graphics. In this context, the ensuing plotting routines draw inspiration and leverage the capabilities of the R package ggplot2 2 (Wickham 2009 [12]). This robust graphical tool enhances the visual representation of data, providing an enriched and insightful dimension to the interpretation and communication of the analytical results.

In order to plot the local gaussian diagonal we used a function we found on GitHub from the user LarsIndus. This function estimates local Gaussian correlation along the main diagonal of a two-dimensional plane and plots the estimated values in a line plot. Default values are appropriate for standard normal marginal distributions. We modified the function to fit in our data in order to get the wanted results.

Parameter dat represents bivariate input data, accepting either a matrix or a data frame. The parameter diaglow is a numeric value serving as the lower bound for points to be estimated or plotted, while diaghigh is its numeric counterpart as the upper bound. Essentially, the plots are generated within the range from (diaglow, diaglow) to (diaghigh, diaghigh). The parameter stepsize, a numeric value, dictates the distance between points for which the local Gaussian correlation is both estimated and plotted. Two additional parameters, b_1 and b_2, both numeric, play pivotal roles as the first and second bandwidth parameters utilized by the function localgauss. The resulting output is encapsulated in a ggplot object, showcasing the estimates of the local Gaussian correlation along the diagonal.

Chapter 4

Empirical Results

The data that are used in this research are coming from the stock indicators S&P 500 (GSPC) and FTSE 100 (FTSE) , the commodities Crude Oil (CL=F) and Gold (GC=F) and CBOE Interest Rate 10 Year (TNX). The prices of the data come for the period January 05, 2004 until November 03, 2023 (4886 daily observations). The data have been sourced from the Yahoo Finance database.

4.1 Tables and graphics

Table 1 showcases the descriptive statistics for all the financial instruments that are used.

Table 1: Descriptive statistics

Statistics	**S&P 500**	**Bond**	**FTSE 100**	**Gold**	**Crude Oil**
Mean	0.000278	0.000008	0.000102	0.000317	0.000302
Min.	-0.127652	-0.347009	-0.115117	-0.098206	-0.282206
Max.	0.109572	0.404797	0.093842	0.086432	0.319634
St. Deviation	0.012235	0.026641	0.011247	0.011301	0.026670
Skewness	-0.5198888	0.2203605	-0.3900202	-0.328591	0.0983746
Kurtosis	12.8614273	29.8753511	10.1645787	5.3031471	18.0810437

Observing the data, it becomes evident that the Standard Deviation is relatively

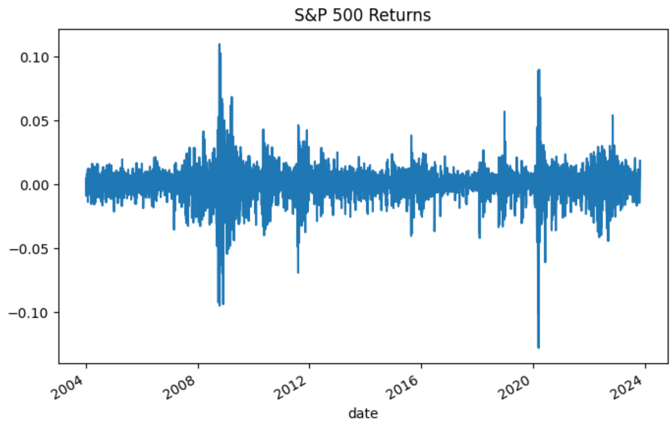


Figure 1: S&P 500 Returns.

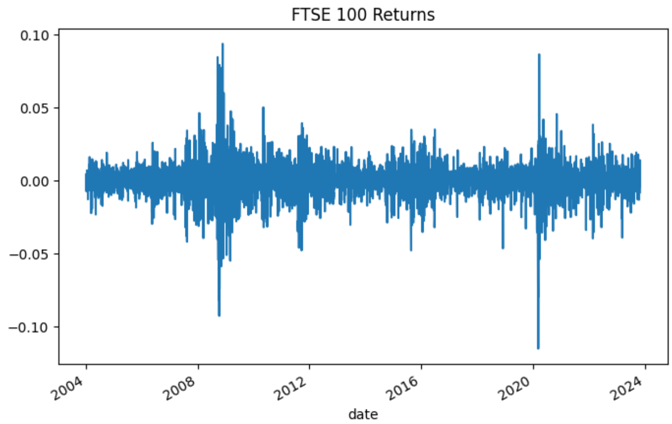


Figure 2: FTSE 100 Returns.

consistent for the two indexes and Gold. However, a slightly elevated Standard Deviation is noticeable for the remaining two instruments, CBOE Interest Rate 10 Year and Crude Oil. This happens due to a higher level of volatility in the years under examination for these particular instruments. It is observed that the skewness values lie within the range of -0.5 to 0.5, signifying a relatively symmetrical distribution of the data. Moreover, the kurtosis values indicate a leptokurtic distribution for all indicators.

Figures 1 - 5 present the returns of the assets that were calculated through Python programming language using the log returns function.

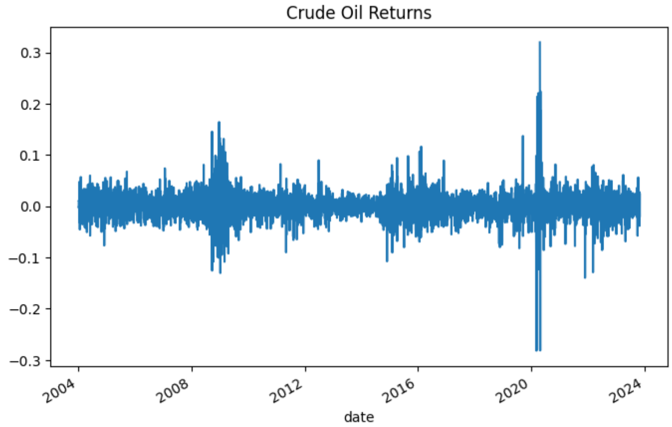


Figure 3: Crude Oil Returns.

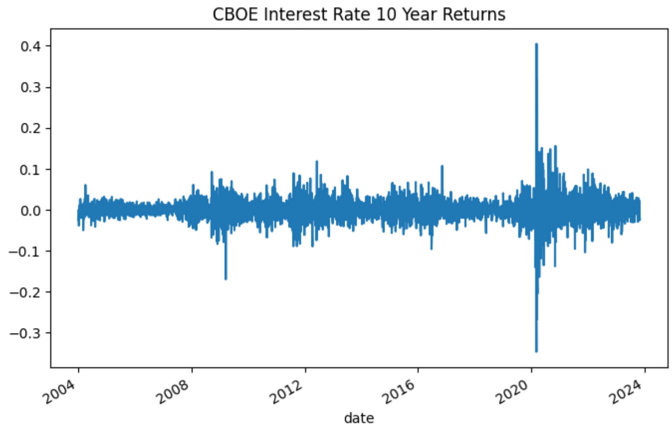


Figure 4: CBOE Interest Rate 10 Year Returns.

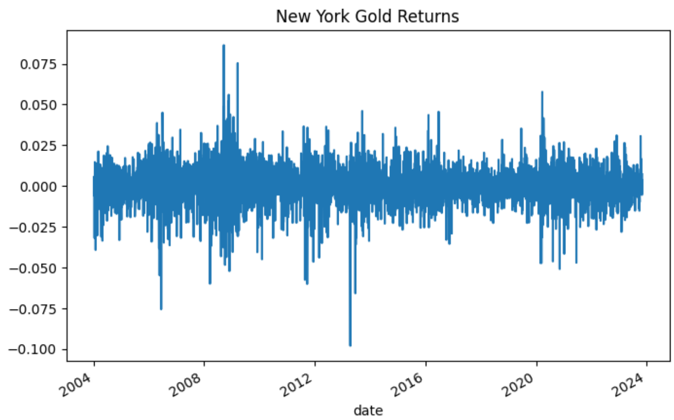


Figure 5: N.Y. Gold Returns.

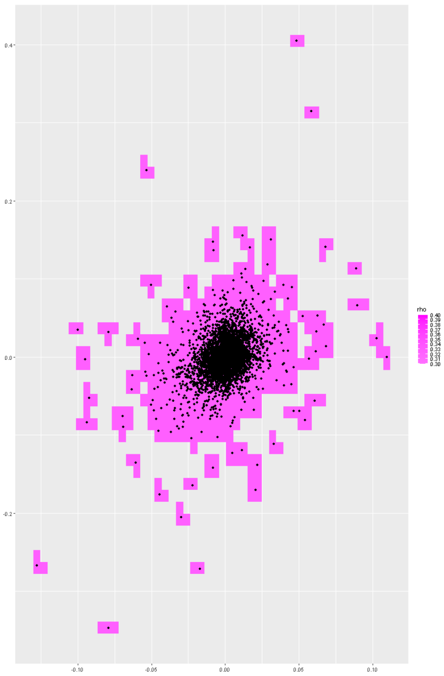
4.2 Local Gaussian Correlation results and graphics

Table 2 displays the arithmetic results for the correlation between two assets.

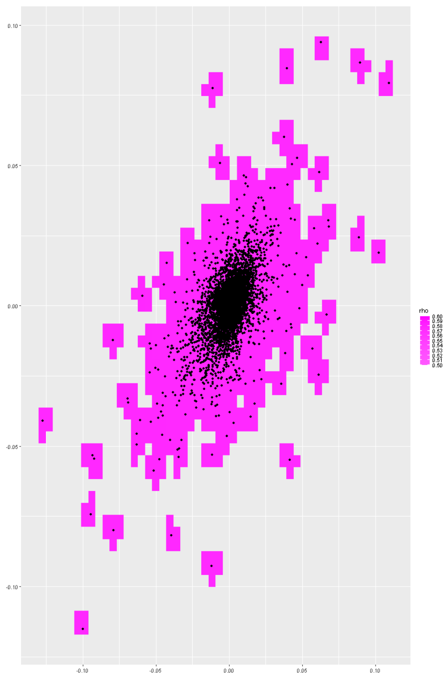
Table 2: Correlation between assets by two.

Assets	**Correlation**
S&P 500 - Bond	**0.3262932**
S&P 500 - FTSE 100	**0.576782**
S&P 500 - Gold	**0.02763491**
S&P 500 - Crude Oil	**0.2724024**
Bond - FTSE 100	**0.3309087**
Bond - Gold	**-0.1617491**
Bond - Crude Oil	**0.2050119**
FTSE 100 - Gold	**0.05883764**
FTSE 100 - Crude Oil	**0.3145499**
Gold - Crude Oil	**0.1998839**

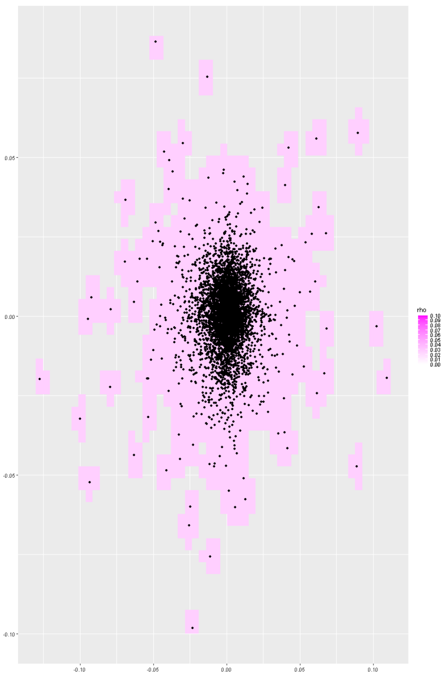
The package lg for the R programming language provided us the following results. The figures that follow depict a dependence scatter map. We can see the Local Gaussian correlation between the assets from the data set, based on 4886 consecutive trading days. In our code the following parameters were used and remained fixed for all the results. These parameters are $b_1 = 1.5, b_2 = 1.5, gsize = 50$ and $hthresh = 0.0025$.



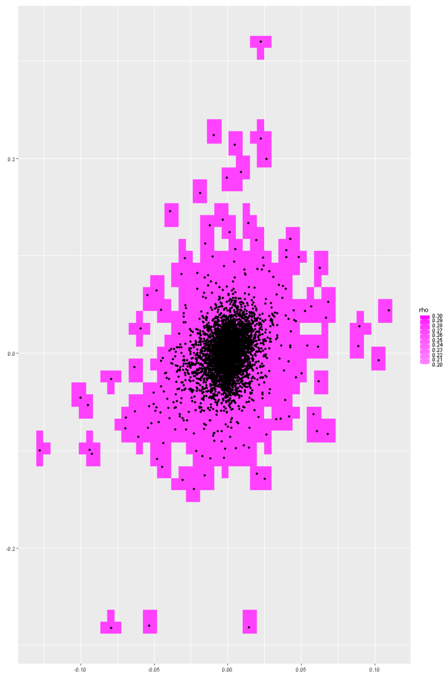
(a) S&P 500 - CBOE Interest Rate 10 Year.



(b) S&P 500 - FTSE 100.

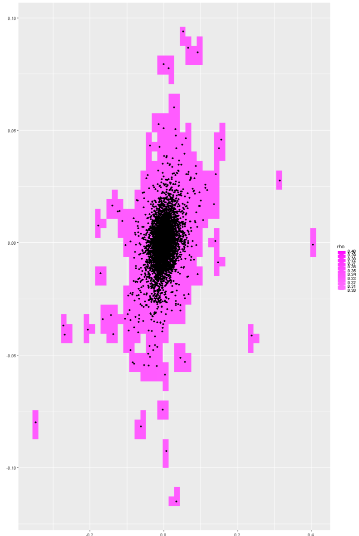


(c) S&P 500 - Gold.

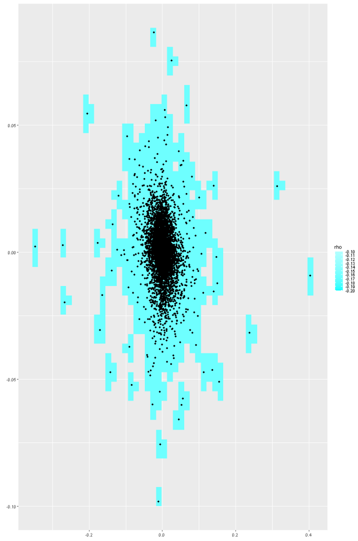


(d) S&P 500 - Crude Oil.

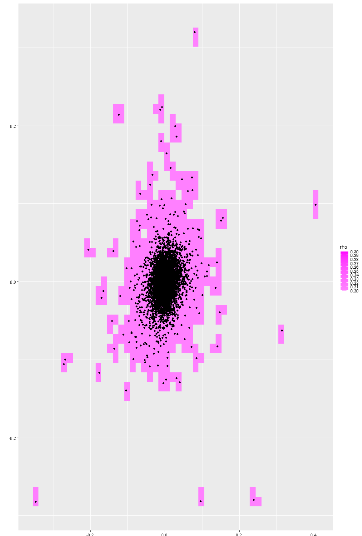
Figure 6: S&P 500 - Assets Local Gaussian correlation.



(a) CBOE Interest Rate 10 Year - FTSE 100.

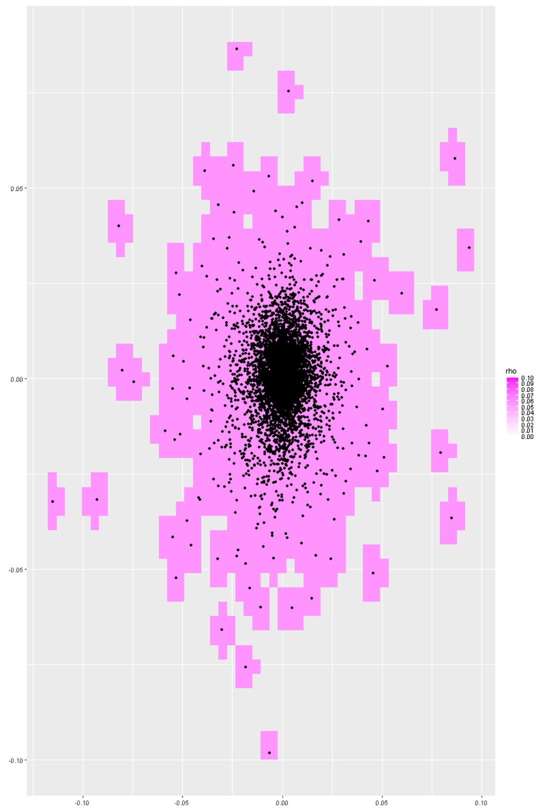


(b) CBOE Interest Rate 10 Year - Gold.

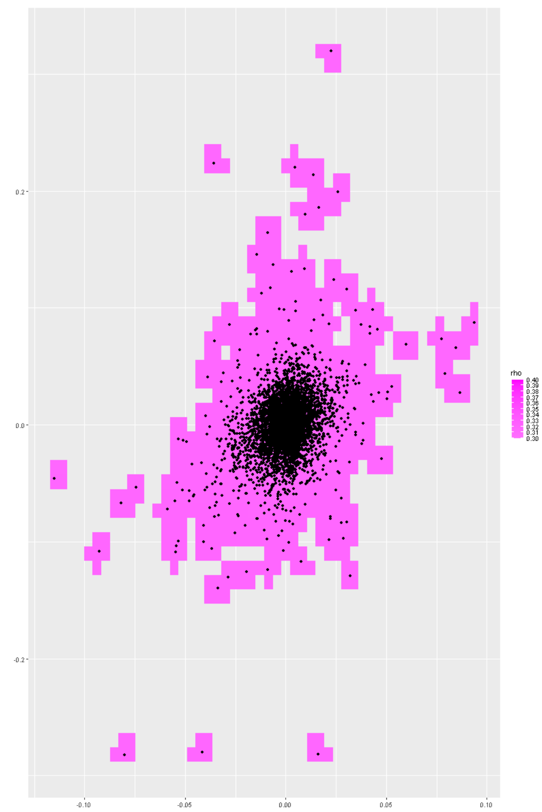


(c) CBOE Interest Rate 10 Year - Crude Oil.

Figure 7: CBOE Interest Rate 10 Year - Assets Local Gaussian correlation.



(a) FTSE 100 - Gold.



(b) FTSE 100 - Crude Oil.

Figure 8: FTSE 100 - Assets Local Gaussian correlation.

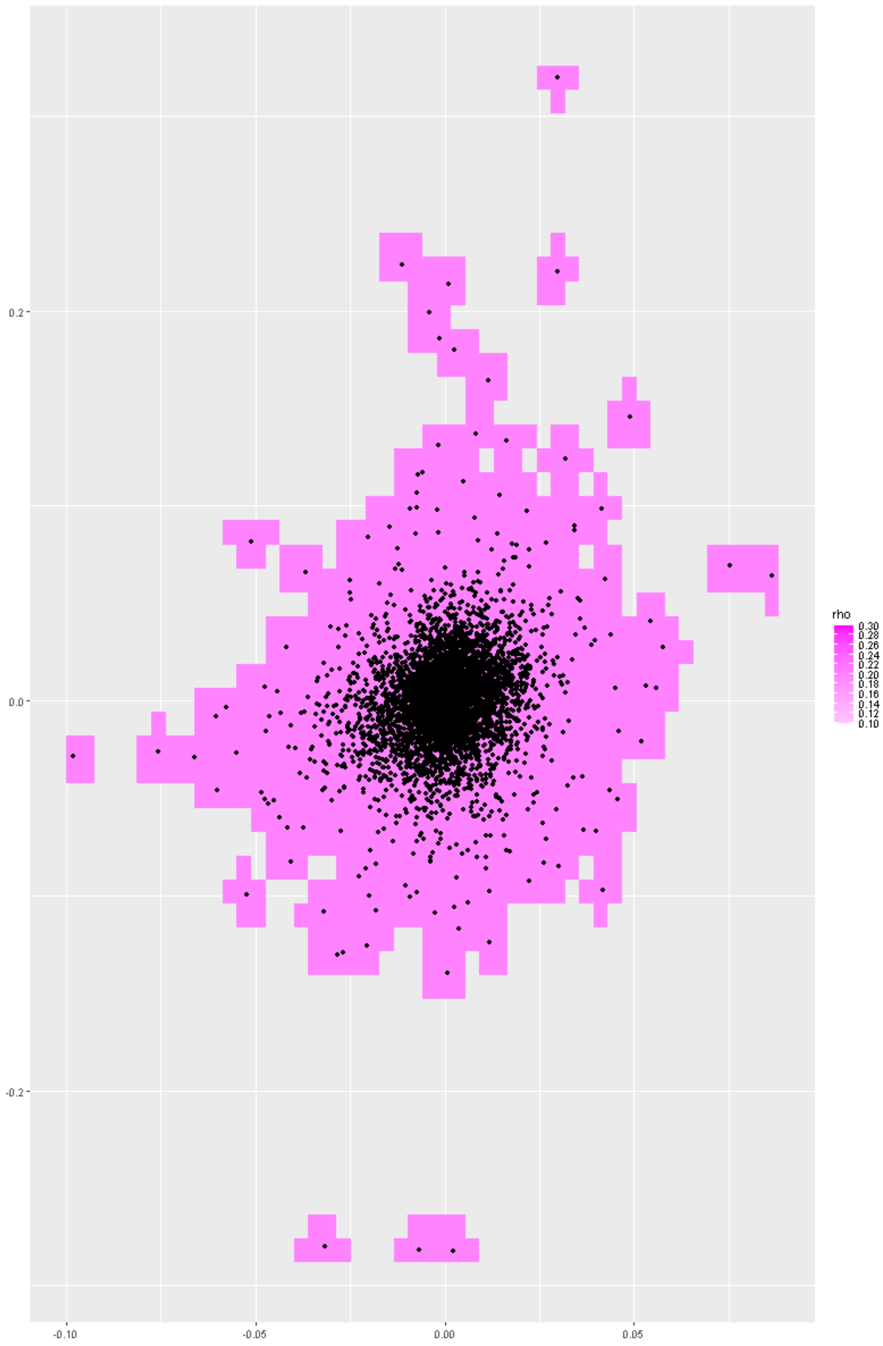
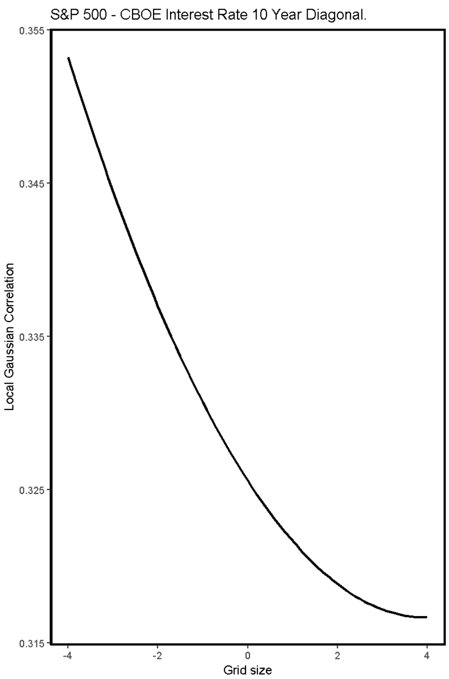


Figure 9: Gold - Crude Oil Local Gaussian correlation.

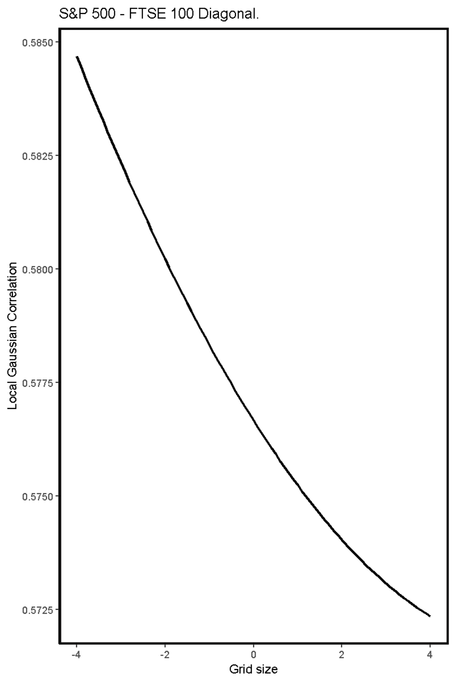
4.3 Local Gaussian correlation estimates along the diagonal

Employing the capabilities of the R programming language, we leverage a dedicated function to estimate local Gaussian correlation, focusing explicitly on the primary diagonal within a two-dimensional plane. The resulting values are thoughtfully visualized through a line plot, with default parameters seamlessly aligning with standard normal marginal distributions. This deliberate choice of plotting along the diagonal frame is strategic, offering an enriched and comprehensive examination of correlation distribution among all possible pairs.

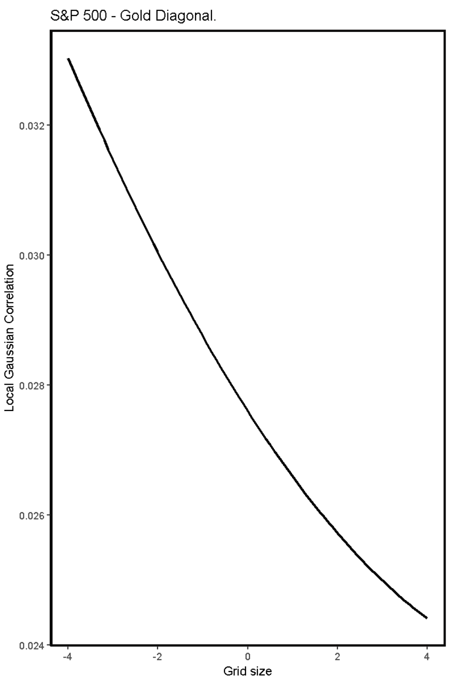
In essence, the local Gaussian diagonal method, implemented in this context, serves as a valuable extension of traditional Gaussian dependence analysis. This particular adaptation facilitates the exploration of non-linear environments, allowing us to discern trends in a more nuanced and comprehensive manner. By adopting this approach, we unveil a deeper understanding of the underlying patterns, providing a more detailed and insightful perspective on the intricacies of the correlation landscape.



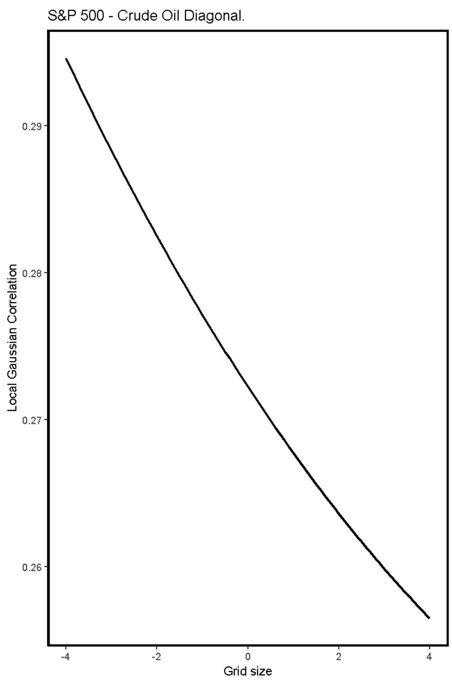
(a) S&P 500 - CBOE Interest Rate 10 Year.



(b) S&P 500 - FTSE 100.

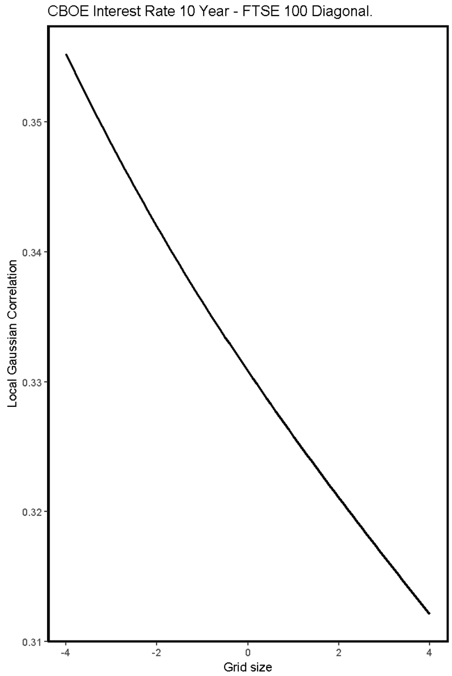


(c) S&P 500 - Gold.

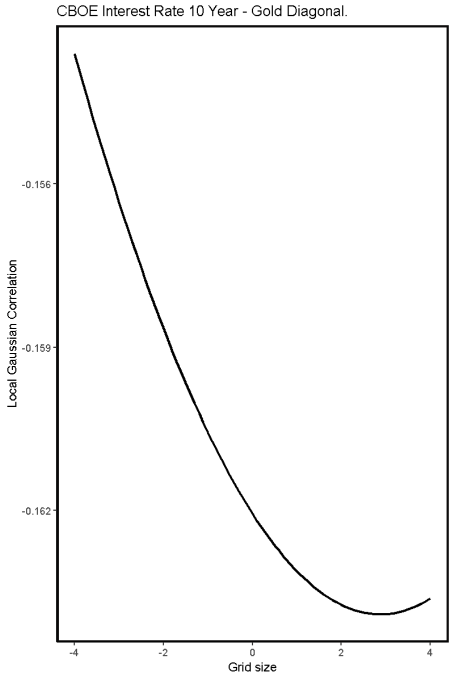


(d) S&P 500 - Crude Oil.

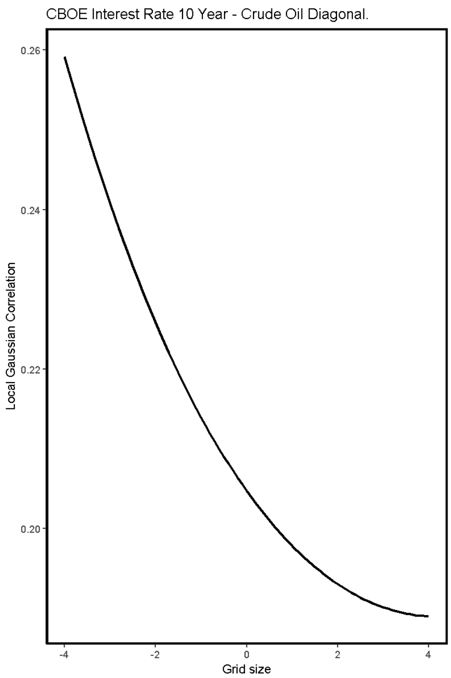
Figure 10: S&P 500 - Assets Local Gaussian correlation diagonal.



(a) CBOE Interest Rate 10 Year - FTSE 100.

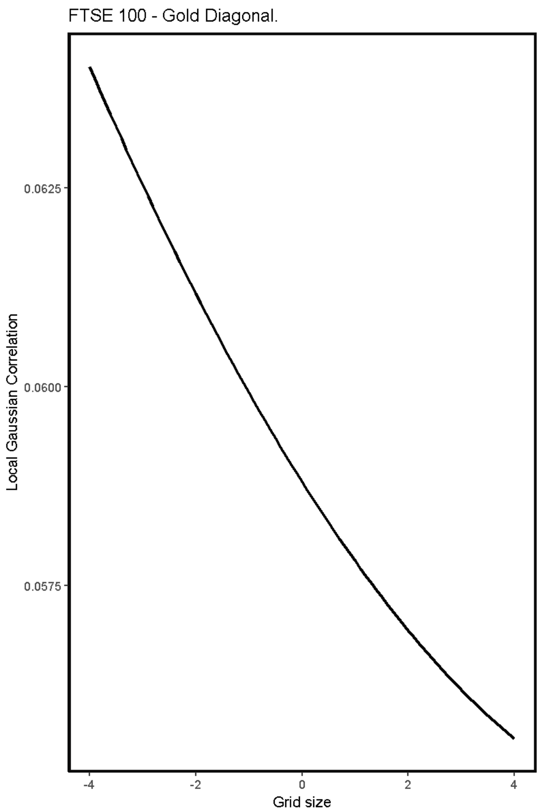


(b) CBOE Interest Rate 10 Year - Gold.

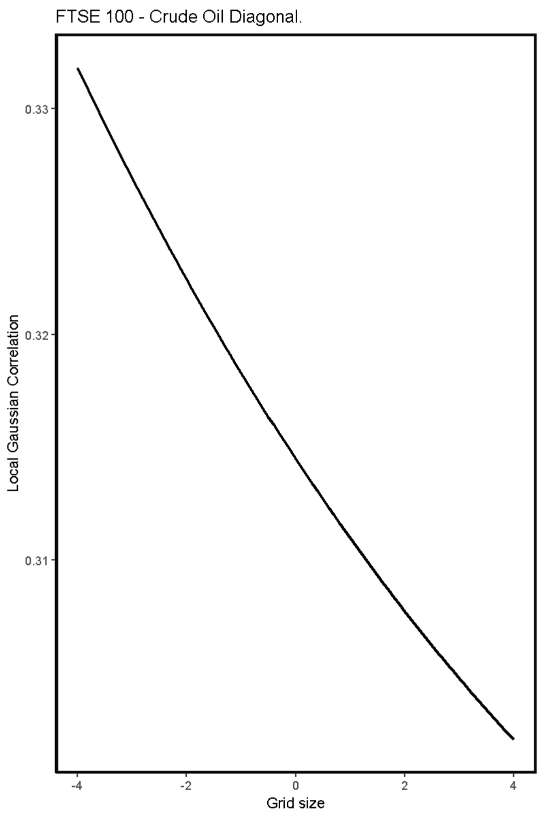


(c) CBOE Interest Rate 10 Year - Crude Oil.

Figure 11: CBOE Interest Rate 10 Year - Assets Local Gaussian correlation diagonal.



(a) FTSE 100 - Gold.



(b) FTSE 100 - Crude Oil.

Figure 12: FTSE 100 - Assets Local Gaussian correlation diagonal.

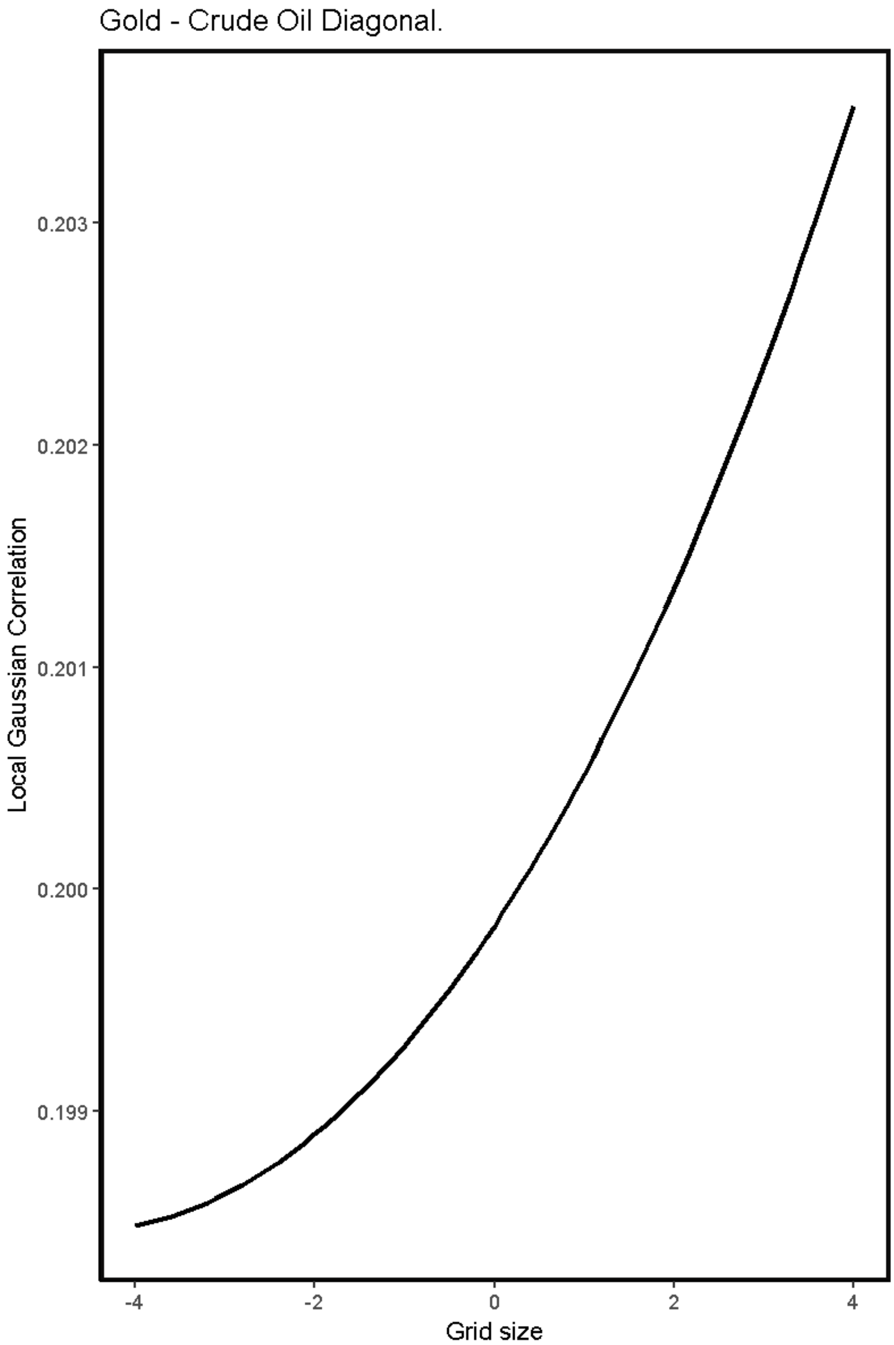


Figure 13: Gold - Crude Oil Local Gaussian correlation diagonal.

Figures 10 - 13 present the computed local Gaussian correlations among the standardized returns derived from the U.S. stock market, the London Stock Exchange index, and various commodities. In each case, the grid size remains constant, ranging from -4 to 4. By segmenting the cases, one can discern and analyze the diagrams in a more granular and detailed manner.

Figure 10 provides a visual representation of the calculated correlation between the U.S. stock market and each asset included in our study. A pattern emerges, yielding consistent outcomes across all scenarios. The correlation coefficients uniformly assume positive values, forming a curve characterized by consistent direction and slope in each diagram, except for the S&P 500 and CBOE Interest Rate 10 Year. In these specific cases, a subtle deviation in the curve becomes apparent towards the conclusion of the diagrams. This deviation signifies a gradual convergence towards a stable correlation value. The same trend is evident in the scatter plot discussed in the preceding section. A close examination of the scatter plot reveals a noticeable convergence of data points toward the center, distinguishing it from other plotted data. A comparative analysis with the S&P 500 and Gold scenario highlights a more dispersed scatter plot, encompassing a broader range of values.

Figure 11 illustrates the estimated correlation utilizing the CBOE Interest Rate 10 Year as the primary paired asset. In this instance, three distinct diagrams come to attention. The initial diagram reveals an almost linear descent in correlation values, a phenomenon anticipated due to the absence of dispersion in the corresponding scatter plot. Rho values are situated aproximately to the center and tend to converge toward a specific value. Transitioning to diagram (b), initial negative values are apparent, exhibiting a descending trend. However, towards the culmination of the curve, there is an observable reversal in direction. It is noteworthy that correlation values change gradually, displaying minimal disparities between them. Conversely, the third diagram illustrates a heightened pace in the progression of values, forming a curve that converges toward a stable numerical outcome.

In Figure 12, the correlation between the FTSE 100 and the commodities market is depicted. Notably, the diagrams exhibit a similar descending trend with positive yet distinct values. A noteworthy observation is the minimal correlation observed between the London Stock Exchange market and the Gold market as a commodity. This pattern is also observed in the case of the S&P 500, where a negligible correlation is evident, largely attributed to the extensive temporal scope of the study. Even when analyze the diagrams derived from the closing prices of these assets over the specified period, a consistent behavioral resemblance becomes apparent, indicating nearly identical regression patterns. However, it is crucial to acknowledge the impact of two significant crises during the studied years: the 2008-2010 Global Financial Crisis (GFC) and the 2020-2022 Covid-19 recession, known as the Great Lockdown, wherein obvious divergences among these assets are clearly observed.

Figure 13 constitutes the final case under consideration, illustrating the correlation between the two commodities: Gold and Crude Oil. Noteworthy is the observation that these two assets not only exhibit a positive correlation but also display an ascending behavior. The depicted correlation values maintain a gradual and stable ascending progression, suggesting a consistent and unchanging correlation between Gold and Crude Oil.

Chapter 5

Conclusions

Within the context of this thesis, a comprehensive exploration unfolded, unraveling the relationships that bind markets, assets, and commodities. This intricate analysis harnessed the power of a local dependence measure, strategically employed as a lens to scrutinize the interplay among these dynamic elements. The overarching aim of this measurement was not merely numerical precision but to intricately portray and comprehend the multifaceted connections embedded within the financial landscape.

Moreover, the revelations derived from this investigation found expression through dual modalities. The pivotal contribution of this thesis manifests in the meticulous representation of data, seamlessly transposed into a continuous diagonal graph. This visualization not only captures the essence of the relationships but also establishes a continuous spectrum for comparative analysis, thereby augmenting the depth and clarity of the findings.

Bibliography

[1] Tjøstheim, D. & Hufthammer, K. O. Local gaussian correlation: A new measure of dependence. *Journal of Econometrics* **172**, 33–48 (2013).

[2] Bampinas, G. & Panagiotidis, T. Oil and stock markets before and after financial crises: A local gaussian correlation approach. *Journal of Futures Markets* **37**, 1179–1204 (2017).

[3] Nguyen, Q. N., Aboura, S., Chevallier, J., Zhang, L. & Zhu, B. Local gaussian correlations in financial and commodity markets. *European journal of operational research* **285**, 306–323 (2020).

[4] Otneim, H. lg: An r package for local gaussian approximations. *R J.* **13**, 15 (2021).

[5] Berentsen, G. D., Kleppe, T. S. & Tjøstheim, D. B. Introducing localgauss, an r package for estimating and visualizing local gaussian correlation. *Journal of Statistical Software* **56**, 1–18 (2014).

[6] Støve, B., Tjøstheim, D. & Hufthammer, K. O. Using local gaussian correlation in a nonlinear re-examination of financial contagion. *Journal of Empirical Finance* **25**, 62–82 (2014).

[7] Nocedal, J. & Wright, S. J. *Numerical optimization* (Springer, 1999).

[8] Metcalf, M. & Reid, J. K. *Fortran 90/95 explained* (Oxford University Press, Inc., 1999).

[9] Pascual, L. H. Tapenade 2.1 user's guide (2004).

[10] Jones, M. C. & Koch, I. Dependence maps: Local dependence in practice. *Statistics and Computing* **13**, 241–255 (2003).

[11] Venables, W., Ripley, B. & Venables, W. Modern applied statistics with s 4 edition springer. *New York* 435–46 (2002).

[12] Wickham, H. ggplot2: elegant graphics for data analysis new york. *NY: Springer* (2009).

Appendix A

Appendix

```
plot_localgauss_diagonal <- function(dat, diag_low = -4, diag_high = 4,
    step_size = 0.1,
                                     b1 = 1, b2 = 1) {

  # convert to matrix if necessary
  if(!is.matrix(dat)) {
    dat <- as.matrix(dat)
  }

  # values for which to calculate LGC
  diag_matrix <- matrix(c(seq(diag_low, diag_high, step_size),
                          seq(diag_low, diag_high, step_size)),
                        ncol = 2)

  # estimate the LGC
  lg.out <- localgauss::localgauss(x = dat[, 1], y = dat[, 2], xy.mat =
      diag_matrix,
                                   b1 = b1, b2 = b2)

  # plotting
```

```
  data.frame(diag = seq(diag_low, diag_high, step_size), rho =
      lg.out$par.est[, "rho"]) %>%
    ggplot(aes(x = diag, y = rho)) +
      geom_line(size=1.5) +
      ylab("Local Gaussian Correlation")+
      xlab('Grid size') +
      ggtitle('')
}
```

```
#example with Gold
import pandas as pd
import yfinance as yf
import numpy as np

#get the data
ticker = ''
gold = yf.download(ticker, start = '2004-01-05', end = '2023-11-05')

#form the data into a data frame
gold.to_excel('gold.xlsx')
pd.read_excel('gold.xlsx', index_col='Date', parse_dates=['Date'])

data = pd.read_excel(".xlsx")
data = data[["Date", "Adj Close"]]

#clean the df
data.dropna()

#get return values
data['log_returns'] = np.log(data['Adj Close'] / data['Adj
    Close'].shift(1))
```

```
data

#form the returns graph
data.set_index("Date", inplace=True)
data['log_returns'].plot(figsize=(8,5), title='')
```