Time-Varying Stock-Bond Correlations: Advanced DCC(1,1) GARCH Modeling, Change Point Detection Methods and Macroeconomic Influences

By
B.A.Stokebrand
SNR: 2037956

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Quantitative Finance and Actuarial Science

Tilburg School of Economics and Management
Tilburg University

Supervised by:
L.Coumans MSc
Second reader:
Prof. Dr. B. Melenberg

Transtrend:
A.Bakx MSc
M.M.A.van Enkhuizen MSc

June 28, 2024

Contents

Acknowledgement

First and foremost, I sincerely thank my university supervisor, Lieske Coumans. Her constant support, helpful feedback and willingness to revisit questions have been incredibly valuable throughout the process of writing this thesis.

I am also very grateful to my supervisors at Transtrend, Anne Bakx and Michael van Enkhuizen. They provided essential data and helped me find additional sources. Their guidance during our regular meetings was very helpful, allowing me to address all my questions.

On a personal note, I want to thank my parents, Ronald and Saskia Stokebrand, for their emotional support during my entire academic career and for carefully reading through my thesis multiple times. Their encouragement has been a great source of strength.

Lastly, I am deeply thankful to my partner, Jasper Delahaije, who has been a sounding board for my ideas and solutions related to my thesis and coding challenges. His patience and support have been invaluable.

Abstract

This thesis addresses two primary questions: (1) Are stock-bond correlations time-varying according to the models used? (2) Are structural breaks in these correlations driven by specific macroeconomic factors?

We investigate the first question using two different models: the Rolling Window Correlation (RWC) and the Dynamic Conditional Correlation (DCC) GARCH model. The findings reveal that while both models show that the correlation is time-varying and show a similar trend in correlations over time, the RWC with a 30-day window exhibits greater volatility compared to the DCC(1,1) GARCH model. The stock-bond correlation being time-varying aligns with prior research indicating that stock-bond correlations are indeed time-varying, fluctuating between positive and negative values.

Applying a change point detection method to both correlations gives a notable similarity in breakpoints identified by both models suggesting robust breakpoints. While some breakpoints are consistent across both methods, the DCC(1,1) GARCH correlation identifies one additional breakpoint, and some of the detected breakpoints occur at different times, either earlier or later. Due to its reduced noise, the DCC(1,1) GARCH model was used for further analysis to identify macroeconomic factors influencing the stock-bond correlation. The study applies change point detection methods to macroeconomic variables such as the Consumer Price Index (CPI) growth rate, the federal funds effective rate, the 10-year expected inflation, and the years of service of the Federal Reserve chairs. Among these factors, the 10-year expected inflation had the highest F1 score of 0.400. Since this is not a very high score, all of these factors do not cause a structural break in the stock-bond correlation, according to this thesis. Which answers the second research question.

The DCC(1,1) GARCH model, in particular, provides a robust tool for capturing the actual dynamics of the stock-bond relationships, aiding in more accurate financial analysis, risk management, and policy decision-making.

1 Introduction

In recent years, Western economies have experienced periods of relatively low inflation, resulting in a challenging financial environment characterized by low interest rates and economic stress. A recent surge in inflation has further disrupted this situation. In response, central banks worldwide have increased interest rates, a monetary policy adjustment with significant implications for various financial instruments, including stocks and bonds. Therefore, studying the drivers affecting the correlations between stocks and bonds is crucial, particularly in the context of time-varying relationships and structural breaks influenced by macroeconomic factors.

Understanding the dynamic correlations between stocks and bonds is essential for several reasons. First, the stock-bond correlation has profound implications for portfolio diversification and risk management. Second, understanding what factors change or influence these correlations can provide insight into certain market dynamics such as, for example, the influence of the expected inflation of the stock-bond correlation.

This thesis aims to extend the literature in several ways through combining multiple papers. First, it builds on the work of Andersson et al. (2008), who presented two methods for calculating time-varying correlations between bonds and stocks. Second, it utilizes the framework developed by Truong et al. (2020) for identifying breakpoints in correlation sequences. Third, it draws from studies by the European Central Bank (2022a), Li (2002), Campbell and Ammer (1993), Czasonis et al. (2020), Pericoli (2018); Ilmanen (2003), which have examined various factors influencing the stock-bond correlation over different periods.

Despite substantial research on stock-bond correlations, significant gaps remain. Previous studies have yet to comprehensively address how varying economic scenarios impact the dynamic correlations between stocks and bonds. Most methods for determining these macroeconomic factors are linear regressions with the macroeconomic factors as independent variables. However, these methods look at all the data, and are not very interesting if, for example, the federal funds effective rate is stable and the correlation is also stable. Our research extends existing studies by employing advanced change point detection methods to identify structural breaks in the correlation between stocks and bonds over the past 60 years. We analyze these breaks to pinpoint the macroeconomic factors driving them. Therefore, our primary contribution is a detailed examination of how economic conditions influence the stock-bond relationship by identifying breaks in their correlation. Our secondary contribution involves analyzing these breaks to uncover the macroeconomic factors driving these structural shifts, providing valuable insights for policymakers and investors.

We use data from the last 60 years to address these research questions and apply the methods outlined by Andersson et al. (2008) and Truong et al. (2020). Specifically, we compare two models to evaluate their effectiveness in capturing time-varying correlations and identifying structural breaks. We find the correlations between bonds and stocks to be time-varying, with multiple changes between positive and negative correlations. When finding the structural breaks in the data, they were not driven by the federal funds effective rate, the Consumer Price Index (CPI) growth rate, the periods in which a chair was active at the Federal Reserve, or the 10-year expected inflation rate. However, the latter had the most significant influence on all four factors based on its F1 score of 0.400.

In what follows, we first review the existing literature on stock-bond correlations. Subsequently, Section 2 provides a detailed overview of the data and the criteria for the chosen models. Section 3 explains the models and methods used to calculate time-varying correlations, to detect breaks in these correlation sequences and to calculate the F1 score as metric for the influence of macroeconomic factors. Section 4 presents the results, followed by a conclusion in Section 5 and a discussion in Section 6. Finally, we offer recommendations for future research.

1.1 Literature Review

1.1.1 Introduction to Dynamic Correlation

The dynamic correlation between stock and bond returns is an essential aspect of financial market analysis, requiring advanced econometric models to capture these time-varying relationships. This literature review addresses two main research questions: (1) Are the correlations time-varying between stocks and bonds based on two different models? (2) Are there any structural breaks in these correlations caused by macroeconomic factors?

The assumption of constant correlations between stocks and bonds has been extensively challenged. Empirical evidence suggests significant time variation in response to changing economic conditions, market volatility, and shifts in monetary policy (Gulko, 2002; Cappiello et al., 2006; Ilmanen, 2003; Connolly et al., 2005; Ang and Bekaert, 2002). These findings highlight the complexity of financial markets and the inadequacy of static models in capturing the nuanced dynamics of asset correlations.

1.1.2 Models for Estimating Time-Varying Correlations

Several models have been developed to estimate time-varying correlations between stocks and bonds, each with its own advantages and disadvantages. Two prominent models are the Rolling Window Correlation (RWC) and the Dynamic Conditional Correlation (DCC) GARCH model.

Rolling Window Correlation (RWC) Model

The RWC model is recognized for its standard application in financial time series analysis, enabling the tracking of evolving correlations over time. It computes correlation coefficients within a moving window of returns, offering insights into the changing dynamics between two different assets (Andersson et al., 2008). In this thesis, the correlation is assessed on a monthly basis, employing a time-based rolling window. This approach ensures consistency across calculations, which is crucial for financial datasets characterized by non-trading periods (weekends and holidays). One advantage of the time-based rolling window is its accommodation of datasets with an unequal number of observations between stocks and bonds; providing flexibility is essential in financial markets where data synchronicity cannot always be assumed. However, the RWC model has limitations in fully understanding cross-return dynamics due to equal weighting of return observations, leading to slow adjustments to new information since new data is weighted $\frac{1}{T}$ of the total information. This results in recent information not getting more weight. Another limitation is the sensitivity to the choice of window since the correlation becomes less volatile if the window size increases. (Andersson et al., 2008; Forbes and Rigobon, 2002).

Dynamic Conditional Correlation (DCC) GARCH Model

GARCH models, such as the DCC(1,1) GARCH, are essential for modeling financial time series characterized by volatility clustering. Standard correlation estimates assume that the relationship between asset returns is static over time. However, in reality, correlations can change depending on market conditions. The DCC(1,1) GARCH model, introduced by Engle (2002), extends the Constant Conditional Correlation (CCC) model by allowing correlations to vary over time while retaining the simplicity of univariate GARCH models in the first stage of estimation. This two-step estimation process involves estimating univariate GARCH models for each asset and then modeling the dynamic correlations. The DCC model is computationally efficient and well-suited for large datasets, making it a popular choice in empirical applications (Engle, 2002).

The DCC(1,1) GARCH model adjusts for heteroscedasticity and estimates the correlation coefficients of standardized residuals, providing unbiased correlation estimates even in volatile markets (Chiang et al., 2007; Celik, 2012; Cho and Parhizgari, 2009). While they offer flexibility, accuracy, and robustness, their limitations include instability in highly volatile data and incorrect risk estimation in the presence of structural breaks (Bauwens et al., 2006). One disadvantage of the DCC(1,1) GARCH is that the parameters α and β are scalars, which is necessary to ensure that the R_t matrix is positive definite for all t. However, because α and β are scalars, all the correlations follow the same dynamics (Bauwens et al., 2006).

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Models

In financial econometrics, various Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models have been developed to model time-varying volatility and correlations. These models are critical in understanding the dynamic behavior of financial time series, particularly for large datasets.

Standard GARCH Model

The standard GARCH model, introduced by Bollerslev (1986), models the conditional variance as a function of past squared returns and past variances. While it is simple and widely used, it is limited to univariate time series and does not capture the potential time-varying correlations between multiple financial assets.

Multivariate GARCH Models: VEC and BEKK

In multivariate settings, the Vector Error Correction (VEC) and BEKK (Baba, Engle, Kraft, and Kroner) models were proposed by Engle and Kroner (1995). The VEC model parameterizes the covariance matrix directly but becomes computationally infeasible for large datasets, making it less usefull if the dimensionality increases. The BEKK model, while ensuring the positive definiteness of the covariance matrix and capturing dynamic interactions between multiple time series, remains computationally challenging for very high-dimensional datasets.

Diagonal Versions of VEC and BEKK Models

Alternative to the VEC and BEKK models are the diagonal versions of the VEC and BEKK models. These variations simplify estimation and enhance computational efficiency by reducing the number of parameters, achieved through the assumption of no cross-asset effects in the conditional variances. In the diagonal VEC model, this means that each variable is influenced only by its own past short-term changes, and not by those of other variables. Consequently, these models rely solely on long-term changes. The conditional correlations between variables may be lower compared to those estimated using a full VEC model, as immediate interactions between variables are not considered. Similarly, in the diagonal BEKK model, the conditional covariance between two variables depends solely on long-term average covariances, thereby ignoring short-term dynamics.

Constant Conditional Correlation (CCC) GARCH Model

The Constant Conditional Correlation (CCC) GARCH model, introduced by Bollerslev (1990), assumes that correlations between assets are constant over time. This model simplifies estimation significantly, but its assumption of constant correlations is often unrealistic in financial markets where correlations can change rapidly due to market conditions.

Extensions: Multivariate GARCH (MGARCH) Models

Other extensions include the Multivariate GARCH (MGARCH) models, which provide a general framework for modeling conditional variances and covariances of multiple time series. The Exponential GARCH (EGARCH) model captures asymmetries in volatility but can be complex to estimate in a multivariate context.

Choosing the DCC(1,1) GARCH Model

Given the trade-offs between model complexity and computational feasibility, we will use the DCC(1,1) GARCH model of Engle (2002) in our analysis. This choice is motivated by its balance of flexibility and efficiency, allowing us to capture time-varying correlations in a computationally manageable framework.

1.1.3 Structural Breaks in Stock-Bond Correlations

Identifying structural breaks in stock-bond correlations is crucial for understanding how macroeconomic factors influence financial markets. Various studies have shown that correlations between stocks and bonds are not stable over time and can be significantly affected by economic conditions such as inflation, economic growth, and interest rates. Previous studies have shed light on the factors influencing the correlation between stocks and bonds. Inflation has been identified as a key determinant, with high inflation periods leading to changes in discount rates that outweigh adjustments in cash flow expectations, thus affecting the stock-bond correlation (Ilmanen, 2003; Andersson et al., 2008). Additionally, market volatility has been shown to potentially induce an upward bias in correlation estimates during periods of market stress (Forbes and Rigobon, 2002).

1.1.4 Methodologies for Detecting Structural Breaks

Several methodologies have been proposed for detecting structural breaks in financial time series, online and offline. Online methods are methods that are able to detect breaks in real-time. We look at offline methods to identify structural breaks in data over the last 60 years. Among these offline methods is the change point detection method by Truong et al. (2020) which offers a robust framework for identifying breaks. This method involves optimizing a cost function to detect multiple breakpoints in a time series, providing a detailed picture of how and when correlations change.

Cost Functions

In selecting appropriate cost functions for change point detection across different datasets, aligning the function with the analysis's underlying data characteristics and specific objectives is crucial. Therefore, we look at non-parametric options since we do not assume a specific distribution for the time-varying correlations. For detecting breaks in these correlations, which often involve complex, non-linear interactions between variables, the Radial Basis Function (RBF) cost function is used, just like Arlot et al. (2012). This choice is motivated by the RBF's flexibility and capability to handle complex data structures, making it well-suited for modeling non-linear relationships (Truong et al., 2020). However, this approach comes with considerations of potential overfitting and increased computational demands, which are mitigated through careful regularization and parameter tuning.

Conversely, for datasets where relationships are expected to be more linear and less volatile, a linear cost function is preferred. The advantages of using a linear cost function include computational efficiency which is particularly beneficial for large datasets. This is essential for econometric analyses where computational speed is a priority (Truong et al., 2020). Additionally, the linear cost function is less prone to overfitting compared to the RBF, ensuring that the models remain robust and generalizable across similar economic datasets. The deliberate choice of cost functions is aligned in these papers with the nature of the data and the analytical goals, ensuring that each model is optimally fitted to provide reliable and interpretable results.

Iterative Methods for Structural Break Detection

To find these structural breaks, we have to iterate over the data in an efficient way. The following methods are based on the analysis of Killick et al. (2012), who studied several methodologies that have been developed to identify shifts in data distributions.

Binary Segmentation Binary Segmentation is one of the earliest and easiest approaches for multiple change point detection. It iteratively applies a single change point detection method to progressively smaller segments of the data. Binary Segmentation is computationally efficient, with a complexity of $O(n \log n)$, making it suitable for large datasets. The algorithm is also relatively simple to implement and understand, which contributes to its widespread use. However, Binary Segmentation is not guaranteed to find the global minimum of the cost function. It often provides a suboptimal solution because it only considers a local subset of possible change points at each step.

Segment Neighborhood Method Alternatively, the Segment Neighborhood method, proposed by Auger and Lawrence (1989), is an exact search method that explores all possible segmentations to find the optimal one. Unlike Binary Segmentation, Segment Neighborhood guarantees finding the global minimum of the cost function, thus providing an optimal segmentation. It can also deal with various penalty functions, allowing for customization based on the specific requirements of the analysis. However, the method has a large computational cost of $O(Qn^2)$, where Q is the upper limit on the number of changepoints. If the number of changepoints increases linearly with the data size, the cost becomes $O(n^3)$, which is the case for this thesis, since we do not impose a restriction on the number of breakpoints.

Optimal Partitioning Alternatively, Optimal Partitioning, as proposed by Jackson et al. (2005), aims to minimize a cost function over all possible segmentations by conditioning on the last point of change. Similar to Segment Neighborhood, Optimal Partitioning provides an exact solution, ensuring that the segmentation is optimal. However, the method also has a high computational cost of approximately $O(n^2)$.

Pruned Exact Linear Time (PELT) Method Alternatively, the PELT method is designed to find multiple changepoints with a computational cost that is linear under certain conditions. It uses pruning to reduce the number of potential change points that need to be considered, significantly improving efficiency. PELT's primary advantage is its linear computational cost, $O(n)$, making it very efficient and scalable to larger datasets. Despite its efficiency, PELT provides an exact solution, ensuring optimal segmentation without sacrificing accuracy. However, the pruning mechanism and the conditions under which it achieves linearity can be complex and difficult to implement correctly.

1.1.5 Conclusion

This literature review highlights the significance of employing time-varying models to capture the dynamic correlations between stocks and bonds. The RWC model is selected for its simplicity and its utility as a benchmark for comparing other models. Additionally, the DCC(1,1) GARCH model is chosen for its computational efficiency and its capability to account for cross-asset relations when determining the covariance. Furthermore, to capture the non-linear relationship in the correlation data, we use the Radial Basis Function (RBF) cost function. For the macroeconomic factors data, we use eather a linear cost function or an RBF cost function based on what fits better with the underlying data. For detecting structural breaks in the time series, we adopt the Pruned Exact Linear Time (PELT) method, which provides an exact solution with a lower computational cost of $O(n)$

This study underscores the importance of using advanced models and methods to better understand the fluctuating nature of stock-bond correlations and the macroeconomic factors influencing these changes. The integration of the RWC and DCC(1,1) GARCH models, along with the RBF and linear cost functions, and the PELT method, provides a robust framework for analyzing these complex relationships.

2 Data

In this section, we present the data used in the analyses, detailing the data itself. We also discuss the assumptions made about the data, the conditions that the data must satisfy to be able to use the DCC(1,1) GARCH model effectively, and we introduce the macroeconomic factors that could potentially influence the stock-bond correlation, such as the 10-year expected inflation, the growth of the CPI, the federal funds effective rate and the moments where the chair of the Federal Reserve changed.

The data used to represent stocks are the prices of the S&P500[1] from 1950 until now. The dataset contains daily opening and closing prices. The descriptive statistics of this data set can be found in Table 2.1. The data used to represent bonds are the 10-year notes[2]. This dataset contains the market yield on U.S. Treasury Securities at 10-year constant maturity, quoted on an investment basis, percent, daily, not seasonally adjusted from 1962 until now. Both datasets have been used in multiple other papers such as Andersson et al. (2008) and Chen et al. (2009). The descriptive statistics for this dataset can also be found in Table 2.1. A graphical representation of both datasets can be found in Appendix A.1. The analysis uses the intersection of the two datasets, focusing exclusively on days for which data is available. We use this approach because it covers the longest period possible with the available data. An extended time-frame is essential due to the substantial evolution in the relationship between stock and bond returns over time.

2.1 Assumptions

The data must satisfy specific properties to use a DCC(1,1) GARCH model effectively. First, the data must be stationary, meaning the statistical properties such as mean and variance should not change over time. Second, we test for the presence of autocorrelation using a Ljung-Box test. Third, we will test for volatility clustering in the residuals in Section 4. Finally, in the latter section, we will test for the normality of the residuals using the Jarque-Bera test.

---

[1] The data belongs to Transtrend and is based on the SPX index ISIN US78378X1072.

[2] From https://fred.stlouisfed.org/series/DGS10

Table 2.1: Descriptive Statistics of Log Returns

Statistic	10-year Note Log Returns	S&P 500 Log Returns
Count	15,455	15,455
Mean	3.63e-8	2.45e-4
Standard Deviation	1.6e-2	1.04e-2
Minimum	-0.315	-0.229
25th Percentile	-5.48e-3	-4.35e-3
Median	0.00	4.24e-4
75th Percentile	5.35e-3	5.19e-3
Maximum	0.342	0.103

Notes: The log returns are calculated from the yields of the 10-year Note and the closing prices of the S&P 500. The descriptive statistics are based on the data from February 2nd, 1962, to January 30th, 2024.

2.1.1 Stationarity

The data needs to be stationary. Stationarity is essential because the error terms in the model capture variability, and non-stationary returns would introduce trends that could compromise the reliability of the DCC(1,1) GARCH results. Typically, financial returns are not stationary (Nazir et al., 2021; Hsu, 1984). We will use the Augmented Dickey-Fuller (ADF) test to assess if the data is stationary. When there is no unit root (indicating stationarity), a series moves around a consistent long-term average, suggesting that the series possesses a finite variance that remains constant over time (Glynn et al., 2007). The hypotheses for the ADF test are as follows:

- H_0: The time series has a unit root (i.e., the time series is non-stationary).
- H_1: The time series does not have a unit root (i.e., the time series is stationary).

The Augmented Dickey-Fuller test examines the stationarity of the time series of the log returns of bonds. The ADF statistic of -17.75 significantly exceeds the critical values at the 1%, 5%, and 10% levels, being smaller than -3.43, -2.86, and -2.57, respectively. This indicates a strong rejection of the null hypothesis, which is evidence against the presence of a unit root in the time series.

The p-value, approximately 3.36e-30, further supports this rejection, suggesting that the probability of observing such a test statistic under the null hypothesis is nearly zero. This leads to the conclusion that the time series is stationary, with no unit root present, at a high confidence level. This also holds for the log returns of the bonds with maturities 2-, 5- and 30-years.

Given these findings, the time series is considered to exhibit consistent statistical properties over time, thus validating the use of models that require stationarity as a fundamental assumption, like the DCC(1,1) GARCH model.

The Augmented Dickey-Fuller (ADF) test is also used to evaluate the stationarity of the time series of the log returns of stocks. The test statistic of -19.93 is below the critical values at all levels of significance (1%, 5%, and 10%), which are −3.43, −2.86, and −2.57, respectively. This substantial deviation from the critical values allows for the robust rejection of the null hypothesis, which is evidence against the presence of a unit root within the series, indicating stationarity. A p-value of 0.0 reinforces the rejection of the null hypothesis, suggesting an extremely high confidence level in the stationarity of the time series under investigation.

2.1.2 Autocorrelation

To test for the presence of autocorrelation, we use a Ljung-Box test. Generally, there is no autocorrelation among stock and bond returns (Engle, 2002), although some studies, such as Andersson et al. (2008), account for autocorrelation. The hypotheses for the Ljung-Box test are as follows:

- H_0: The time series exhibits no autocorrelation at the specified lag.
- H_1: The time series exhibits autocorrelation at the specified lag.

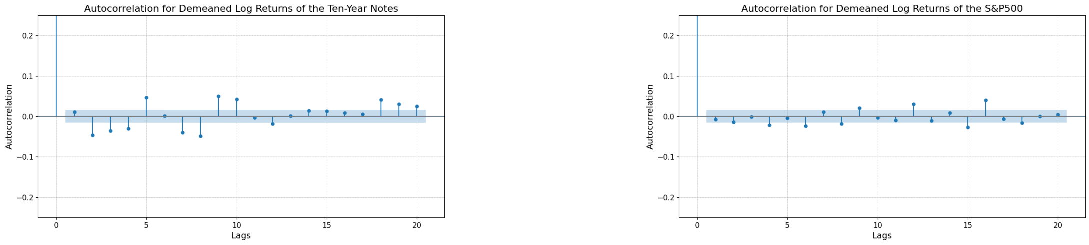
The Ljung-Box test results for the log returns of bonds and stocks are summarized in Table 2.2. Graph 2.1 presents a graphical representation of the autocorrelation.

Table 2.2: Ljung-Box Test Results for Log Returns

Data Type	**Lags**	**Statistic**	**p-Value**
Log Bond Returns	1	1.74	0.187
	2	35.41	2.04e-8***
	3	55.53	5.29e-12***
	4	70.21	2.05e-14***
	5	104.38	6.30e-21***
	10	231.38	4.41e-44***
	20	294.63	1.01e-50***
Log Stock Returns	1	1.87	0.171
	2	2.20	0.333
	3	2.20	0.532
	4	10.41	3.40e-2**
	5	10.65	5.88e-2*
	10	30.66	6.68e-4***
	20	54.40	5.00e-5***

Note: This table summarizes the Ljung-Box test results for autocorrelations in the log bond and stock returns. Using the 10-year notes as bond data. Significance levels: $* : p < 0.10, ** : p < 0.05, *** : p < 0.01$.

The Ljung-Box test statistic and corresponding p-values for log bond returns reveal significant autocorrelations at lags greater than 1. Specifically, the p-values at lags two and



(a) Autocorrelation for Demeaned Log Returns of the 10-year Notes

(b) Autocorrelation for Demeaned Log Returns of the S&P500

Figure 2.1: Autocorrelation Function (ACF) Plots

higher are all extremely low (smaller than 10e-8), firmly rejecting the null hypothesis of no autocorrelation. This indicates that log bond returns exhibit significant autocorrelation, suggesting that past returns substantially impact future returns at these lags. This result is the same for bond returns.

In contrast, the log stock returns show a different pattern. The p-values are above a 5% significance level for the first three lags, suggesting no significant autocorrelation at these short lags. However, at lag 4, the p-value drops to 0.034, indicating a weak rejection of the null hypothesis at a 5% significance level. As the lags increase, the p-values decrease, with lags 10 and 20 showing significant autocorrelation (p-values of 6.68e-4 and 5e-5, respectively). This implies that while log stock returns do not exhibit substantial autocorrelation at very short lags, there is evidence of autocorrelation at longer lags.

To further verify the robustness of these findings, the Ljung-Box test is performed on bond data with maturities of 2, 5, and 30 years. Autocorrelation is even more noticeable in these test results, particularly at lag 1 for bonds with 2-year and 5-year maturities, as shown in Table 2.3. This table also indicates that the results for bonds with 30-year maturities remain consistent with the initial findings.

Overall, these results suggest that while both bond and stock returns exhibit autocorrelation, the pattern and strength of this autocorrelation differ between the two asset classes. Bond returns show significant autocorrelation at nearly all tested lags, while stock returns show significant autocorrelation primarily at longer lags. The presence of autocorrelation aligns with the findings of Baele et al. (2010), who also found some autocorrelation in bonds and stock returns.

Therefore, an AR(1) model is used to find the error terms of the log returns.

Table 2.3: Ljung-Box Test Results for Log Bond Returns

Bond Type	Lags	Statistic	p-Value
2 year maturity	1	213.02	3.009e-48***
	2	221.16	9.48e-49***
	3	221.42	9.92e-48***
	4	222.39	5.73e-47***
	5	225.37	1.05e-46***
	10	238.58	1.36e-45***
	20	268.81	1.80e-45***
5 year maturity	1	10.34	1.30e-3**
	2	35.73	1.75e-08***
	3	54.60	8.36e-12***
	4	54.98	3.28e-11***
	5	63.19	2.65e-12***
	10	109.01	8.43e-19***
	20	162.04	3.04e-24***
30 year maturity	1	0.523	0.469
	2	23.60	7e-6***
	3	39.28	1.51e-08***
	4	59.57	3.57e-12***
	5	76.14	5.37e-15***
	10	158.17	7.72e-29***
	20	207.63	3.47e-33***

Note: This table summarizes the Ljung-Box test results for autocorrelations in log bond returns across various lags for 2YR, 5YR and 30YR bonds. Significance levels: $*$: $p < 0.10, ** : p < 0.05, *** : p < 0.01$.

2.1.3 Volatility Clustering

Volatility clustering is a characteristic that justifies the use of a model that can accommodate changing volatility. Volatility clustering is a fundamental characteristic of financial time series, where periods of high volatility tend to be followed by similarly high volatility, and low volatility periods are followed by low volatility. Typically, financial data shows volatility clustering (Lo and MacKinlay, 2011; Cont, 2001). We test for volatility clustering using Engle's Autoregressive Conditional Heteroskedasticity (ARCH) test in Section 4.

2.1.4 Normality of the Residuals

To use the DCC(1,1) GARCH model effectively, it is necessary to assume a distribution for the residuals. A common assumption is that the residuals follow a normal distribution, an assumption supported by various studies Engle (2002); Andersson et al. (2008); Bautista

(2003); Cappiello et al. (2006); Manera et al. (2006). Consequently, we test for the normality of the residuals using the Jarque-Bera test. The results of this test are presented in Section 4.

2.2 Macroeconomic Factors

Based on factors identified in previous studies, we examine various additional elements that empirically influence the correlation between stocks and bonds or are expected to do so. Czasonis et al. (2020) demonstrated that inflation significantly affects the stock-bond correlation. Consequently, we analyze the monthly Consumer Price Index (CPI) growth rate for all items in the United States as a measure of inflation[3]. The data is illustrated in Figure 2.2.

Several potential breakpoints can be identified from a visual inspection of the CPI data in Figure 2.2. The 1960s to the early 1970s shows relative stability with minor fluctuations. Around the mid-1970s, there is a noticeable increase in volatility. The early 1980s exhibit significant spikes in the rate, reflecting high inflation. Following the early 1980s, the period from 1990 until 2000 seems relatively stable with less volatility. Around 2008-2010, there was a sharp drop followed by increased volatility, likely due to the global financial crisis. Post-2010, the rate exhibits more fluctuations, but the mean remains relatively constant. After 2020, the CPI growth rate starts to increase again.

The expected breakpoints in the CPI growth rate data are around the mid-1970s, early 1980s, late 1990s, late 2000s, and early 2020s. These breakpoints will be identified in Section 4.3.2.

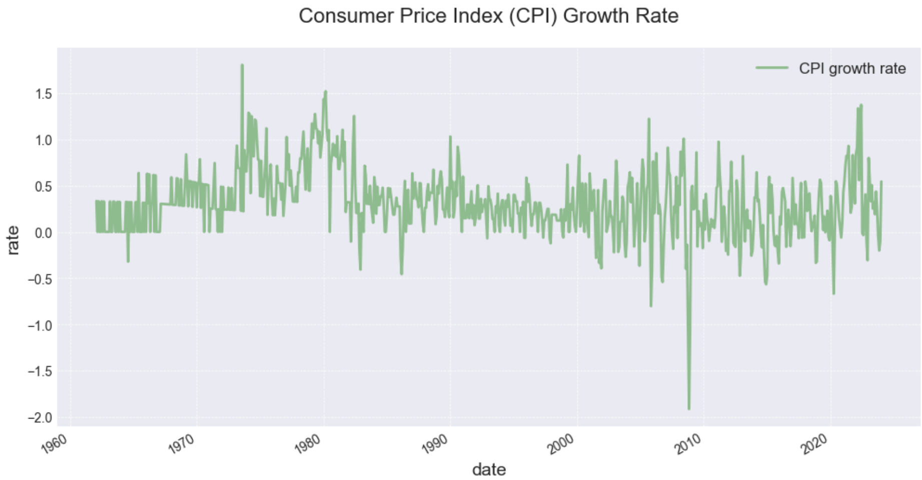


Figure 2.2: Consumer Price Index (CPI) growth rate

Furthermore, David and Veronesi (2001) indicated that fluctuations in inflation uncertainty explain certain variations in the volatilities and covariances of stock and bond returns. Therefore, we include the 10-year expected inflation rate for the United States[4], as shown in

[3] https://fred.stlouisfed.org/series/CPALTT01USM657N
[4] https://fred.stlouisfed.org/series/EXPINF10YR/

Figure 2.3.

Visual inspection of the data in Figure 2.3 identifies several potential breakpoints. The period from 1980 to around 1987 shows a large decline in the rate, followed by a period of relative stability with fluctuations between 1985 and 1995. Around 1995, another significant decline was observed, lasting until approximately 1998.

The rate seems to have moved downwards from 2000 to 2010. After 2010, it appears to stabilize. Finally, there is a noticeable increase in the rate after 2020.

The expected breakpoints in the 10-year expected inflation data are around 1987, around 1995, around 2000, late 2000s, and early 2020s. These breakpoints will be identified in section 4.3.2.

Figure 2.3: 10-Year Expected Inflation

Pericoli (2018) showed that unconventional monetary policies can impact the stock-bond correlation. To investigate this, we consider two different aspects. First, we examine the monthly federal funds effective rates set by the Federal Reserve[5], shown in Figure 2.4. Additionally, we analyze the periods during which each individual served as the Chair of the Federal Reserve. Table 2.5 lists all Federal Reserve Chairs starting from the beginning of our dataset. A graphical representation is given in Figure 2.6. We use the transitions between these chairs as breakpoints to test their potential to explain the stock-bond correlation.

Several potential breakpoints can be identified based on the visual inspection of the federal funds effective rate data in Figure 2.4. There seems to be a sharp increase in the federal funds effective rate around the mid-1970s. The late 1970s to early 1980s also show a sharp increase in the rate, reaching peaks above 17.5%. This is followed by a significant decline in the mid-1980s. The late 1980s exhibited fluctuations with moderate peaks and troughs, reflecting economic adjustments. The early 1990s show a large drop. The early 2000s show a large drop followed by a peak of the same size and an even more significant drop around 2008, reflecting the response to the global financial crisis with near-zero interest rates. Post-2015, the rate gradually increases. Finally, there is a noticeable increase around

[5]https://fred.stlouisfed.org/series/fedfunds

2020.

The expected breakpoints in the federal funds effective rate data are around the mid-1970s, the late 1970s to early 1980s, mid-1980s, late 1980s to early 1990s, 2000s, 2008-2010, post-2015, and 2020 onwards. These breakpoints will be identified in Section 4.3.2.

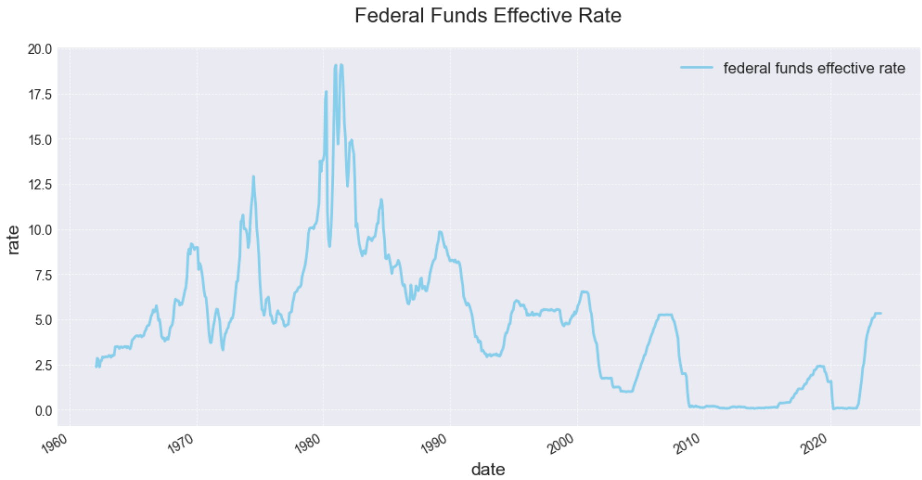


Figure 2.4: Federal Funds Effective Rate

Figure 2.5: Chairs of the Federal Reserve and Their Years of Service

Chair	**Years as Chair**
William McChesney Martin	1951-1970
Arthur F. Burns	1970-1978
G. William Miller	1978-1979
Paul Volcker	1979-1987
Alan Greenspan	1987-2006
Ben Bernanke	2006-2014
Janet Yellen	2014-2018
Jerome Powell	2018-Present

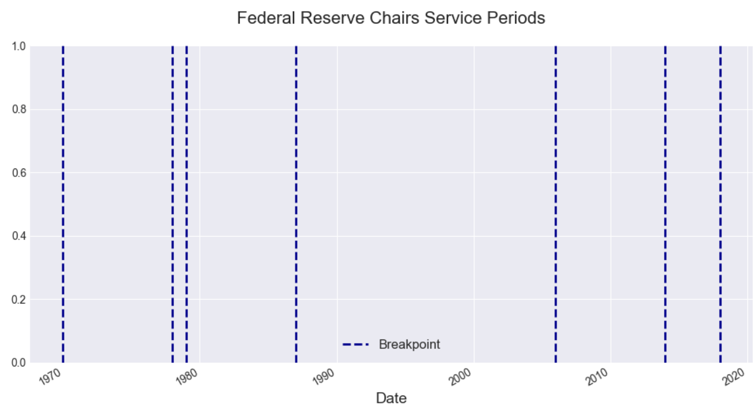


Figure 2.6: Federal Reserve Chairs

3 Methodology

The analysis in this thesis explains two primary models: the Rolling Window Correlation (RWC) model and the Dynamic Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity (DCC(1,1) GARCH) model, as introduced by Engle (2002). The DCC model is a more efficient variant of the multivariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model (Andersson et al., 2008). This section provides a detailed explanation of both models, alongside a change point detection method and the formulas of the F1 score which are used to evaluate the effectiveness of various factors in explaining structural breaks in the correlation data. These methodologies together facilitate a robust analysis of the time-varying relationships between stock and bond returns.

3.1 Rolling Window Correlation

The Rolling Window Correlation (RWC) method is a fundamental approach for analyzing the dynamic relationship between financial assets, such as stocks and bonds. The mathematical formulation of the RWC, as adapted from Andersson et al. (2008) and Nazir et al. (2021), can be found in Equation 3.1, which is presented below:

$$\rho_{t,S,B} = \frac{\sum_{i=1}^{T}(r_{S,t-i} - \bar{r}_{S,t})(r_{B,t-i} - \bar{r}_{B,t})}{\sqrt{\sum_{i=1}^{T}(r_{S,t-i} - \bar{r}_{S,t})^2 \sum_{i=1}^{T}(r_{B,t-i} - \bar{r}_{B,t})^2}} \tag{3.1}$$

Here, $r_{t,S}$ and $r_{t,B}$ denote the returns on stocks and bonds, respectively, while $\bar{r}_S$ and $\bar{r}_B$ represent the average returns over the T day window looking back from time t. It is acknowledged that for certain days within the window, data for $r_{S,t-i}$ or $r_{B,t-i}$ may be unavailable. These instances are consequently excluded from the calculation, acknowledging that some T day windows may contain a variable number of data points. This approach shows the model's adaptability in handling real-world data inconsistencies.

3.2 DCC(1,1) GARCH model

The DCC(1,1) GARCH model, initially proposed by Engle (2002) and adapted by Andersson et al. (2008), offers a sophisticated approach to modeling the conditional correlation between stocks and bonds over time. This model uses a two-step estimation process to capture the dynamics between these financial assets.

The first step deviates from the original model by Engle (2002). It incorporates possible autocorrelation in the returns, based on modifications made by Andersson et al. (2008). This involves fitting a linear regression to both the autoregressive process of returns and the autoregressive process of volatilities. Let r_t denote the demeaned log returns of bonds and stocks, represented as $r_t = \{r_{S,t}, r_{B,t}\}$, where $r_{S,t}$ and $r_{B,t}$ are the returns of the stocks and bonds, respectively.

The log returns follow an AR(1) process, as shows in Equation 3.2.

$$r_t = \gamma + \phi r_{t-1} + \varepsilon_t \tag{3.2}$$

where γ is the constant term accounting for a potential non-zero mean and ϕ captures the autocorrelation in the returns. The error term ε_t is assumed to follow a normal distribution with a mean of zero and a time-varying variance $H_{ii,t}$, conditional on past information $\Im_{t-1}$, $\varepsilon_{i,t}|\Im_{t-1} \sim N(0, H_{ii,t})$. There are multiple papers that also assume normality of the error term such as Engle (2002), Andersson et al. (2008), Bautista (2003), Cappiello et al. (2006) and Manera et al. (2006). The matrix H_t is given by:

$$H_t = D_t R_t D_t \tag{3.3}$$

In the next step, we model the conditional variance D_t. The conditional variance is given by:

$$D_t^2 = \text{diag}\{\omega\} + \text{diag}\{\kappa\} \circ \varepsilon_{t-1}\varepsilon_{t-1}' + \text{diag}\{\lambda\} \circ D_{t-1}^2 \tag{3.4}$$

In this equation, D_t is a diagonal matrix with elements $\sqrt{h_{BB,t}}$ and $\sqrt{h_{SS,t}}$ representing the conditional standard deviations of the returns of the bonds and stocks, respectively. ω represents the baseline variance. Which we expect to be close to zero as the variance is then completely captured by the other variables. κ represents how recent shocks in the log returns impact the volatility. λ shows how much past variances influence the current variance.

Where R_t represents the conditional correlation matrix. Given the assumption of normally distributed error terms, we define the standardized residuals z_t, as:

$$z_t = D_t^{-1}\varepsilon_t \sim N(0, I) \tag{3.5}$$

In the second step we use the standardized residuals, z_t, to model the conditional covariance matrix Q_t as:

$$Q_t = (1 - \alpha - \beta)\bar{Q} + \alpha z_{t-1}z_{t-1}' + \beta Q_{t-1} \tag{3.6}$$

Here, $\bar{Q} = \frac{1}{T}\sum_{t=1}^{T} \hat{z}_t\hat{z}_t{}'$ is the unconditional covariance matrix of z_t[1]. The parameters α and β are very important as they ensure the stationarity of the model, with the condition $\alpha + \beta < 1$.

Finally, the time-varying correlation matrix R_t is derived from the conditional covariance matrix Q_t:

[1] Multiple papers choose different matrices for $\bar{Q}$, for this thesis we will follow the approach of Lee et al. (2006) and Bautista (2003) who use the unconditional covariance matrix.

$$R_t = \{\rho_{ij,t}\} \text{ where } i,j \in S,B \tag{3.7}$$

with the correlation between assets i and j given by:

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}} \tag{3.8}$$

Where $q_{ij,t}$ are entries of Q_t.

To estimate the parameters of the model, we maximize the log-likelihood function, which can be decomposed into the likelihood of the volatility component $L_V(\theta)$ and the likelihood of the correlation component $L_C(\theta,\phi)$.

$$L = -\frac{1}{2}\sum_{t=1}^{T}(\underbrace{n\log(2\pi) + 2\log|D_t| + r_t' D_t^{-1} D_t^{-1} r_t}_{L_V(\theta)} \underbrace{-\varepsilon_t'\varepsilon_t + \log|R_t| + \varepsilon_t' R_t^{-1}\varepsilon_t}_{L_C(\theta,\phi)}) \tag{3.9}$$

By first maximizing $L_V(\theta)$, we obtain the estimates of the volatility and return parameters, $\gamma, \phi, \omega, \kappa$ and λ. These estimates are then filled in Equation 3.2 and 3.4 which are both used to calculate 3.5. After which we then maximize $L_C(\theta,\phi)$ with the optimal values from $L_V(\theta)$ to obtain the estimates of the conditional covariance parameters, α and β. These estimates are then used to calculate the conditional covariance in Equation 3.6 and the matrix Q_t is then used to calculate the correlation like in Equation 3.8. Which then leads to the correlation matrix R_t as given in Equation 3.7.

3.3 Change Point Detection

To identify structural breaks in a data sequence, it is essential to use change point detection methods. These methods find the moments at which the statistical properties of the sequence, such as the mean or variance, change significantly. All the following methods are based on the paper of Truong et al. (2020) unless mentioned otherwise.

Consider a time series signal $y = \{y_t\}_{t=1}^T$ representing our input data. We define a sub-signal of length $b-a$ from the series $\{y_t\}_{t=a+1}^b$ $(1 \le a < b \le T)$ as $y_{a..b}$, thus the complete signal can be expressed as $y = y_{0..T}$. A set of indexes is denoted by $\mathcal{T} = \{t_1, t_2, \ldots\} \subset \{1,\ldots,T\}$, with $|\mathcal{T}|$ denoting its cardinality. For a finite index set $\mathcal{T} = \{t_1,\ldots,t_K\}$, the dummy indices $t_0 := 0$ and $t_{K+1} := T$ are included. Here, K represents the number of change points.

In this context, change point detection assumes that the data process $y = \{y_t\}_{t=1}^T$ is piecewise stationary, implying that some characteristics change abruptly at unknown times $t_1^* \le t_2^* \le \ldots \le t_K^*$. Change point detection aims to estimate these indices t_K. The number of change points K is unknown and must be estimated.

Change point detection is a model selection problem, which involves selecting the optimal segmentation $\mathcal{T}$ based on a quantitative criterion $V(\mathcal{T}, y)$ that must be minimized.

We assume that the criterion function $V(\mathcal{T})$ for a given segmentation is the sum of the costs of all segments defining the segmentation:

$$V(\mathcal{T}, y) := \sum_{k=0}^{K} c(y_{t_k..t_{k+1}}) \tag{3.10}$$

where $c(\cdot)$ is a cost function that evaluates the goodness-of-fit of the sub-signal $y_{t_k..t_{k+1}} = \{y_t\}_{t=t_k+1}^{t_{k+1}}$ to a specific model. The "best segmentation" $\hat{\mathcal{T}}$ is the minimizer of the criterion $V(\mathcal{T})$. Given the unknown number of changes, the change point detection problem involves solving the following discrete optimization problem:

$$\min_{\mathcal{T}} V(\mathcal{T}) + \text{pen}(\mathcal{T}) \tag{3.11}$$

where $\text{pen}(\mathcal{T})$ is a penalization term that influences the number of detected breaks. A higher value results in fewer breaks, while a lower value allows for more breaks.

The cost function $c(\cdot)$, as introduced in Equation 3.10, is a measure of "homogeneity." Its selection dictates the type of changes detectable. Ideally, $c(y_{a..b})$ is low if the sub-signal $y_{a..b}$ is "homogeneous" (i.e., contains no change points) and high if $y_{a..b}$ is "heterogeneous" (i.e., contains one or more change points).

We employ a non-parametric kernel-based detection method to avoid assumptions about the underlying data. This method, proposed by Harchaoui and Cappé (2007), performs change point detection in a non-parametric setting by mapping the original signal y onto a reproducing kernel Hilbert space (RKHS) $\mathcal{H}$ associated with a user-defined kernel function $k(\cdot,\cdot) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. The mapping function $\phi : \mathbb{R}^d \to \mathcal{H}$ is implicitly defined by $\phi(y_t) = k(y_t, \cdot) \in \mathcal{H}$, leading to the following inner product and norm:

$$\langle \phi(y_s) | \phi(y_t) \rangle_{\mathcal{H}} = k(y_s, y_t) \quad \text{and} \quad \|\phi(y_t)\|_{\mathcal{H}}^2 = k(y_t, y_t) \tag{3.12}$$

for any samples $y_s, y_t \in \mathbb{R}^d$. The corresponding cost function, denoted c_{kernel}, is defined as follows:

$$c_{\text{kernel}}(y_{a..b}) := \sum_{t=a+1}^{b} \|\phi(y_t) - \bar{\mu}_{a..b}\|_{\mathcal{H}}^2 \tag{3.13}$$

where $\bar{\mu}_{a..b} \in \mathcal{H}$ is the sample mean of the signal $\{\phi(y_t)\}_{t=a+1}^{b}$ and $\|\cdot\|_{\mathcal{H}}$ is defined in 3.12.

After algebraic manipulation, which can be found in Appendix B.1, $c_{kernel}(y_{a..b})$ can be expressed as:

$$c_{\text{kernel}}(y_{a..b}) = \sum_{t=a+1}^{b} k(y_t, y_t) - \frac{1}{b-a} \sum_{s,t=a+1}^{b} k(y_s, y_t). \tag{3.14}$$

Different kernel functions can be combined with this cost function (Shawe-Taylor and Cristianini, 2004). Examples used in this thesis include the linear kernel and the Gaussian kernel, as shown in Equation 3.15. According to Truong et al. (2020), these are also among the most commonly used kernels.

$$k(x, y) = \begin{cases} \langle x, y \rangle = x^T y \text{ with } x, y \in \mathbb{R}^d & \text{linear kernel} \\ \exp(-\gamma \|x - y\|^2) \text{ with } x, y \in \mathbb{R}^d & \text{gaussian kernel} \end{cases} \tag{3.15}$$

where $\gamma \geq 0$ is the bandwidth parameter. While most papers do not specify a value for γ (Killick et al., 2012; Jackson et al., 2005), Garreau and Arlot (2018) provided a value for the bandwidth parameter, setting γ to 50.

By substituting these kernel functions into the cost functions, we obtain:

$$\begin{aligned} c_{\text{rbf}}(y_{a..b}) &:= (b-a) - \frac{1}{b-a} \sum_{s,t=a+1}^{b} \exp(-\gamma \|y_s - y_t\|^2) \\ c_{\text{linear}}(y_{a..b}) &:= \sum_{t=a+1}^{b} y_t^T y_t - \frac{1}{b-a} \sum_{s,t=a+1}^{b} y_s^T y_t \end{aligned} \tag{3.16}$$

After determining the cost function, it is essential to incorporate a penalty to limit an excessive number of breaks. The choice of penalty is subjective and highly dependent on the characteristics of the underlying data. Therefore, to identify an appropriate penalty value, we must evaluate the detected breaks using various penalties. The goal is to select the penalty that aligns best with the breaks we visually interpret as a break. This iterative process ensures that the chosen penalty value reflects the true underlying structure of the data, balancing the trade-off between overfitting and underfitting.

When the number of changes is unknown, we solve the following penalized optimization problem:

$$\hat{K}, \{\hat{t}_1, \ldots, \hat{t}_{\hat{K}}\} := \arg \min_{K, \{t_1, \ldots, t_K\}} V(t_1, \ldots, t_K) + \beta K \tag{3.17}$$

where $\beta > 0$ is a parameter set to different values for different purposes and $\hat{K}$ is the estimated number of change points. Higher values of β yield fewer $\hat{K}$.

The objective is to solve Equation 3.17. A straightforward approach involves applying the optimization method for different values of K (from 1 to a sufficiently large $K_{\max}$) and then selecting the segmentation that minimizes the penalized criterion. However, this approach is computationally inefficient due to the complexity of the optimization method.

A more efficient algorithm exists for a broad class of penalty functions, specifically linear penalties, expressed as:

$$\text{pen}(\mathcal{T}) = \beta |\mathcal{T}|$$

where $\beta > 0$ is the same parameter. The Pelt algorithm (”Pruned Exact Linear Time”) is introduced to find the exact solution efficiently by using a specific pruning rule to eliminate unlikely change points (Killick et al., 2012). For two indexes t and s (where $t < s < T$), the pruning rule is defined as follows:

$$\text{if } \left[\min_{\mathcal{T}} V(\mathcal{T}, y_{0..t}) + \beta |\mathcal{T}| \right] + c(y_{t..s}) \geq \left[\min_{\mathcal{T}} V(\mathcal{T}, y_{0..s}) + \beta |\mathcal{T}| \right],$$

then, t cannot be the last change point prior to s. This rule significantly speeds up the computation. Assuming that the length of a regime is randomly drawn for a uniform distribution, the complexity of the Pelt algorithm is linear, i.e., $O(T)$. The algorithm's pseudo-code can be found in the paper of Truong et al. (2020).

3.4 F1 Score

After using the change detection method and finding these breaks, we will assume that these detected breaks in the correlation between bonds and stocks, as determined by the DCC(1,1) GARCH model, represent the true breaks. Using the same change point detection methods, we will analyze additional data sequences of different macroeconomic factors to find breaks, applying different cost functions and penalties to accommodate the distinct properties of these datasets.

Subsequently, we will use an F1 score to evaluate the influence of macroeconomic factors on the stock-bond correlation. The F1 score uses precision and recall to calculate a metric that can compare the breaks found in both data sequences. To the best of our knowledge, this approach has not been previously applied in this context. Traditionally, this metric is utilized to assess the performance of change point detection methods by comparing the identified breakpoints against the true breakpoints, as demonstrated by Truong et al. (2020).

Precision measures the ratio of true positive predictions to the total number of positive predictions made by the model, indicating how many predicted breaks in the macroeconomic factor dataset coincide with breaks in the correlation data within an error margin $M > 0$. On the other hand, recall measures the ratio of true positive predictions to the total number of actual positive cases, reflecting how many true breaks in the correlation data are also breaks in the macroeconomic factor dataset.

True positives (TP) are defined where a structural break in a macroeconomic factor is within a margin M of a break in the time-varying correlation. Formally, this can be represented as:

$$\text{TP}(\mathcal{T}^*, \hat{\mathcal{T}}) := \left\{ t^* \in \mathcal{T}^* \mid \exists \hat{t} \in \hat{\mathcal{T}} \text{ s.t. } |\hat{t} - t^*| < M \right\} \tag{3.18}$$

Precision and recall are then given by:

$$\text{PREC}(\mathcal{T}^*, \hat{\mathcal{T}}) := \frac{|\text{TP}(\mathcal{T}^*, \hat{\mathcal{T}})|}{|\hat{K}|} \quad \text{and} \quad \text{REC}(\mathcal{T}^*, \hat{\mathcal{T}}) := \frac{|\text{TP}(\mathcal{T}^*, \hat{\mathcal{T}})|}{|K^*|} \tag{3.19}$$

Where $\hat{K}$ is the number of predicted breakpoints in the macroeconomic factor dataset and K^* is the number of true breakpoints found in the correlation data. Precision and recall are defined between 0 and 1 if the margin, M, is less than the minimum length between two true change point indices t^*_k and t^*_{k+1}. Over-segmentation of a signal results in precision approaching zero and recall approaching one, whereas under-segmentation has the reverse effect. The F1 score is given by:

$$F_1\text{-Score}(\mathcal{T}^*, \hat{\mathcal{T}}) = 2 \times \frac{\text{Prec}(\mathcal{T}^*, \hat{\mathcal{T}}) \times \text{Rec}(\mathcal{T}^*, \hat{\mathcal{T}})}{\text{Prec}(\mathcal{T}^*, \hat{\mathcal{T}}) + \text{Rec}(\mathcal{T}^*, \hat{\mathcal{T}})} \tag{3.20}$$

This approach allows us to assess which macroeconomic factors have the most significant impact on the stock-bond correlation by measuring the alignment of breakpoints across datasets using the F1 score.

4 Results

This section presents the results from the Rolling Window Correlation (RWC), the DCC(1,1) GARCH model correlations, the structural breaks obtained from applying change point detection methods, and the analysis of macroeconomic factors explaining these structural breaks.

4.1 Rolling Window Correlation

Figure 4.1 presents the results of the Rolling Window Correlation (RWC) model using a 30-day window[1]. The results reveal significant volatility in the correlation between stock and bond returns. From 1960 to 1970, the correlation fluctuated around zero, showing more negative values until 1967 and then becoming more positive. This pattern aligns with findings from Ilmanen (2003). Between 1970 and 2000, the correlation was predominantly positive, after which it shifted to being negative until the mid-2000s. Following this period, the correlation turned positive again until around 2008, during the financial crisis, when it once more became negative, remaining so until the early 2020s. This observation is consistent with the European Central Bank's Financial Stability Review (European Central Bank, 2022b), which shows a positive stock-bond correlation in the United States from 1990 to 2000, mirroring the movements observed in the RWC model from 2000 to 2010. Notably, the correlation from 2013 to 2022 is more negative in our results than the ECB report, showing a movement around -0.3. For a robustness check, the market yields on U.S Treasury Securities at a 2-year, 5-year and 30-year[2] constant maturity, quoted on an investment basis, percent, daily, not seasonally adjusted are used. The figures for the Rolling Window Correlations can be found in the Appendix D.1. These results show a similar pattern to the pattern observed in Figure 4.1.

[1] Appendix C.1 provides the outcomes for other window lengths.

[2] From https://fred.stlouisfed.org/series/DGS2, https://fred.stlouisfed.org/series/DGS5 and https://fred.stlouisfed.org/series/DGS30.

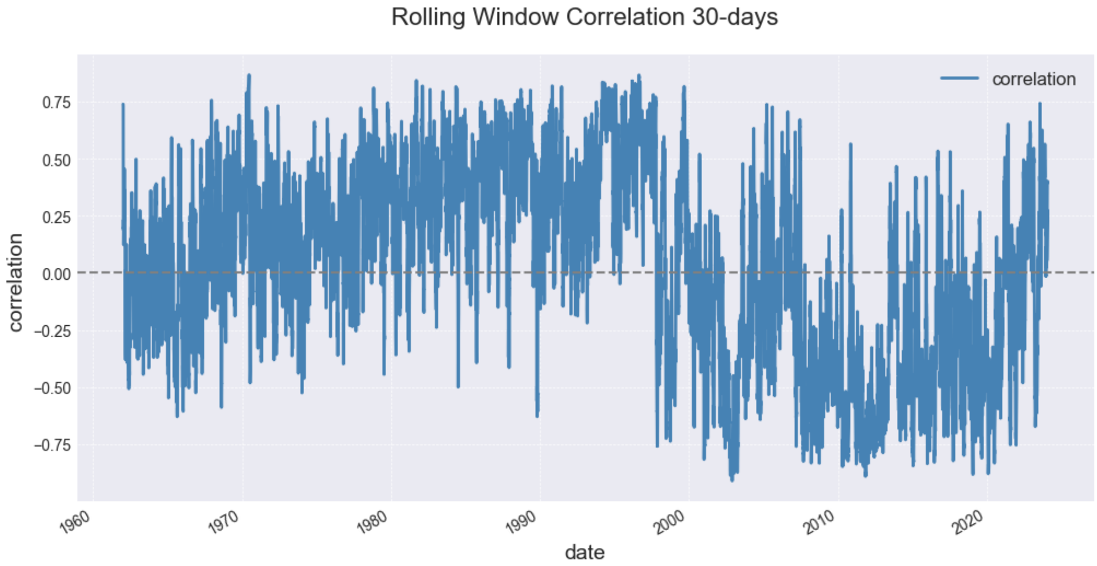


Figure 4.1: Rolling Window Correlation

4.2 DCC(1,1) GARCH

Table 4.1: Parameter Estimates for Log Bonds and Stocks Returns with T-Statistics

	Bond	**T-Statistic**	**Stock**	**T-Statistic**
γ	0.0002	1.342e2***	0.0031	1.621e2***
ϕ	0.1000	1.47*	0.1000	1.01
ω	1.024e-06	0.27	5.13e-07	0.16
κ	0.2167	6.454e3***	0.2165	6.453e3***
λ	0.7997	5.692e2***	0.8002	6.124e2***

Note: This table displays parameter estimates and their corresponding t-statistics, where the t-statistics are calculated using a hessian matrix. Significance: $* : p < 0.1, ** : p < 0.05, *** : p < 0.01$.

The results of the first step of the DCC(1,1) GARCH correlation analysis are shown in Table 4.1. The coefficient γ is very small, indicating minimal additional volatility dynamics not captured by the other parameters. This is consistent with the findings of Andersson et al. (2008), who reported γ values of zero, with γ_B not being statistically significant. However, in this case, the t-statistics for both γ coefficients are remarkably high, reflecting the precision and statistical significance of these estimates.

The parameter ϕ, nearly identical for both bonds and stocks (approximately 0.10), suggests an autoregressive component in both return series. This similarity indicates that returns for both assets are similarly influenced by their past values, implying that past returns have a moderate impact on current returns. The t-statistics for ϕ show that the parameter for bonds is statistically significant at a 90% confidence level, while the parameter for stocks is not. The ϕ value of 0.1000 found in this study is slightly higher than the results of Andersson et al. (2008), who reported a statistically insignificant ϕ_S of -0.007 and a statistically significant ϕ_B of 0.054. This suggests that there is almost no autocorrelation in the stock returns in Andersson et al. (2008)'s study, while there is some in this thesis.

The baseline variance, denoted by ω, measures the baseline level of volatility. This parameter varies between the two assets, indicating differences in their inherent volatility. The t-statistics for ω are very low and not statistically significant. Andersson et al. (2008) reported ω estimates of 0.005 for bonds and 0.003 for stocks, both statistically significant. While these low values align with Andersson et al. (2008)'s results, the values found by our model might be too low.

Both assets exhibit a relatively large κ value (around 0.216), indicating that recent shocks substantially impact current volatility. The equality in the magnitude of this parameter for both asset classes underscores a similar responsiveness to new information or market events in terms of increasing the volatility. The high t-statistics for κ confirm the strong effect of recent shocks on volatility. In contrast, Andersson et al. (2008) reported lower κ values of 0.042 for stocks and 0.034 for bonds, both statistically significant, suggesting that recent shocks influence current volatility less in their case, which seems less logical.

High values of λ (close to 0.8 for both assets) suggest that past conditional variances are highly predictive of current variances, indicating strong volatility clustering. This means that bonds and stocks tend to maintain their previous levels of volatility over time. The t-statistics for λ are also very high, indicating that the persistence in volatility is highly statistically significant. Andersson et al. (2008) reported λ values of 0.954 for stocks and 0.951 for bonds, both statistically significant, showing a higher persistence of variance in their parameters.

The κ and λ values are very close across the two assets, suggesting similar dynamics in terms of volatility clustering and reaction to new information. However, differences in γ and ω imply that the baseline levels and sensitivities to new shocks are different, with stocks generally being more volatile. The high t-statistics across most of the parameters reflect the robustness of these findings and provide strong evidence for the described dynamics in bond and stock returns.

The results of the same analysis, but using different data for the log bond returns as a robustness check, are presented in the Appendix, Table D.1[3]. The findings, estimates, and t-statistics are very similar, indicating that the method is highly robust. The implications that the parameters have on the correlation will be discussed in Section 4.2.3.

4.2.1 Volatility Clustering

We apply Engle's Autoregressive Conditional Heteroskedasticity (ARCH) test to the residuals of log returns for bonds and stocks to detect this phenomenon. The ARCH test specifically

[3]The estimates for the log stock returns also vary since the datasets with different maturities have different time-frames. Therefore, the results are not identical.

checks whether there are significant autocorrelations in the squared residuals of the time series, an indicator of volatility clustering. For Engle's ARCH test, the hypotheses are:

- H_0: No ARCH effects (the residuals are not conditionally heteroskedastic, which implies no significant autocorrelations in the squared residuals)
- H_1: Presence of ARCH effects (the residuals exhibit conditional heteroskedasticity, indicating significant autocorrelations in the squared residuals)

For the residuals of log bond returns, the test statistic is significantly high at 6,816.08 with a p-value of 0.0. Similarly, the log stock returns exhibit a substantial test statistic of 2,143.85 with a p-value of 0.0. Due to these p-values, we reject the null hypothesis of no autoregressive conditional heteroskedastic effects, confirming the presence of volatility clustering. This pattern is also observed in the residuals of log bond returns with maturities of 2, 5, and 30 years[4].

These findings confirm the need to use models capable of accommodating the changing volatility, such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. Since the DCC(1,1) GARCH model addresses volatility clustering by dynamically updating the volatility matrix at each time point, thereby capturing the time-varying nature of volatility.

4.2.2 Normality of the Residuals

The Jarque-Bera test assesses the distributional characteristics of the daily returns data. This test is beneficial for spotting deviations from the normal distribution by capturing both its skewness and kurtosis.
For the residuals of the log bond returns, the Jarque-Bera test statistic is 2.07e6, with a p-value of 0.000. This extremely high test statistic and corresponding p-value indicate strong evidence against the null hypothesis that the residuals are normally distributed. As a robustness check, the Jarque-Bera test was also performed on the log bond returns for maturities of 2, 5, and 30 years[5]. In each case, the results similarly indicate a strong rejection of the normality assumption.

Similarly, the log stock returns' residuals yield a Jarque-Bera test statistic of 3.68e5 with a p-value of 0.000. Like the bond returns, this result strongly rejects the null hypothesis of normality in the distribution of the residuals. These findings suggest that the log returns on both stocks and bonds significantly deviate from a normal distribution when considered over a period of 60 years. This deviation could be attributed to the presence of heavy tails and skewness in the distribution of returns, which is typical in financial markets due to factors like market sentiment, large jumps in prices due to economic events, and other market anomalies (Lo and MacKinlay, 2011).

The error terms are not normally distributed. However, for this thesis, we will assume normality. This approach follows the methodology of Andersson et al. (2008), who analyzed

[4]The results are not included in this thesis since the results show no new insights.
[5]The results are not included in this thesis since the results show no new insights.

the same datasets over a shorter timeframe. We have redone the analysis of Andersson et al. (2008) on the timeframe that they used to test for normality. The results are indicated in Table E.1. These results show that the residuals for the dataset within this timeframe used by Andersson et al. (2008) are also not normally distributed. Additionally, testing different periods within the timeframe used by Andersson et al. (2008) reveals that the residuals remain non-normally distributed in both their results and those derived in this thesis, as shown in Tables E.2 and E.3.

4.2.3 DCC(1,1) GARCH Step 2

The estimated parameters of the second step of the DCC(1,1) GARCH model are as follows:

Table 4.2: Parameter Estimates

Parameter		**Bootstrap 50**	**Bootstrap 10**
	Estimate	T-Statistic	T-Statistic
α	0.0325	8.41***	6.14***
β	0.9619	3.44***	4.08***

Note: This table displays parameter estimates and their corresponding t-statistics, where the t-statistics are calculated based on a bootstrap method. Bootstrap 50 means that the data was resampled 50 times. Significance: $*** : p < 0.01$.

The estimated value of α is 0.0325, indicating the responsiveness of the correlation dynamics to new shocks or innovations. This relatively low value suggests that new shocks have a small impact on changing the correlation between bonds and stocks. This implies that daily market fluctuations or specific news events do not tend to change the correlation structure between these two asset classes drastically. The parameter estimate is statistically significant for 10 and 50 bootstrap samples, as presented in Table 4.2, which supports the findings of Andersson et al. (2008), who reported a value for α of 0.042.

The estimated value of β is 0.9619, indicating high persistence in correlations. This high β value signifies that once established, the correlation between stock and bond returns tends to remain stable over time. The parameter estimate is statistically significant at a 99% confidence interval for 10 bootstrap samples and 50, as shown in Table 4.2. This resembles the results from Andersson et al. (2008) closely, as they reported a value for β of 0.950.

The sum of α and β is approximately 0.9944, slightly below 1, which is needed for model stability. This near-unit sum indicates that the correlations are mean-reverting, ensuring that the model's predictions remain realistic and bounded, thus avoiding extreme over- or underestimations of correlations over time, which was also an assumption of the model.

The results of the same analysis, but using different data for the log bond returns as a robustness check, are presented in Appendix D.2. The findings are very similar, indicating that the method is highly robust.

The Dynamic Conditional Correlations (DCC) graph can be found in Figure 4.2. The DCC model correlations follow a similar dynamic to the Rolling Window correlations, although it exhibits less volatility and consequently less noise. The stock-bond correlation is slightly negative from 1962 until 1967, aligning with the theory proposed by Ilmanen (2003). Subsequently, the correlation becomes positive from 1967 until a slight dip around 1990, after which it returns to positive values until approximately 1999. It then turns negative but reverts to positive a year later, briefly dropping to very negative values, even below -0.6. Around 2005, the correlation became slightly positive again, only to collapse around 2008 and remain negative until the early 2020s. These observations are consistent with the movements described in the European Central Bank's Financial Stability Review (European Central Bank, 2022b).

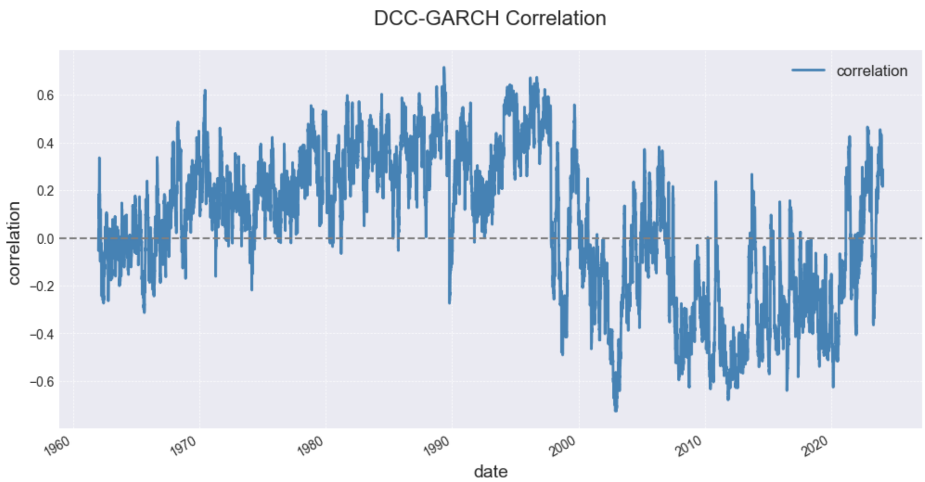


Figure 4.2: Correlation DCC(1,1) GARCH using the 10-year Note

4.3 Break-test

4.3.1 Correlation

When detecting breakpoints in data sequences, this is done through the minimization of a cost function plus a penalty. For these results, the penalties were selected based on visual inspection of the correlation graphs as shown in Figure 4.3 and 4.4. This approach ensured the detection of large changes in the correlation. The Radial Basis Function (RBF) cost function was used to identify these breakpoints. This method was chosen because visual inspection indicated it was the most effective in detecting the expected breaks.

The penalties were calibrated to detect only larger changes, thus avoiding overfitting the model with an excessive number of breakpoints. This calibration involved checking multiple penalties and examining the resulting breakpoints. Very low penalties could result in over 100 breakpoints, which is beyond the scope of this thesis. The objective is to identify macroeconomic factors affecting the bond-stock correlation on a broader basis rather than on a monthly basis. For the rolling window correlations, a penalty of 50 was chosen, resulting in 9 structural breaks over approximately 60 years[6]. This aligns with the findings of A'Hearn

[6]The last break is excluded because the model always shows a break after the final data point.

and Woitek (2001), who noted that a typical business cycle has an average duration of 7-10 years. For the DCC(1,1) GARCH correlations, a penalty of 100 was selected, resulting in 10 structural breaks over the same period[7], which is also consistent with the results of A'Hearn and Woitek (2001). In contrast, the rolling window correlation (RWC) method results in fewer breakpoints, likely due to higher volatility in the correlation estimates, as shown in Figure 4.3, compared to those generated by the Dynamic Conditional Correlation (DCC) GARCH model, as illustrated in Figure 4.4. Consequently, the RWC method failed to detect a noticeable strutural break that the DCC(1,1) GARCH model could identify, which is the break around 2013[8]. However, all other breaks seems to be rather similar.

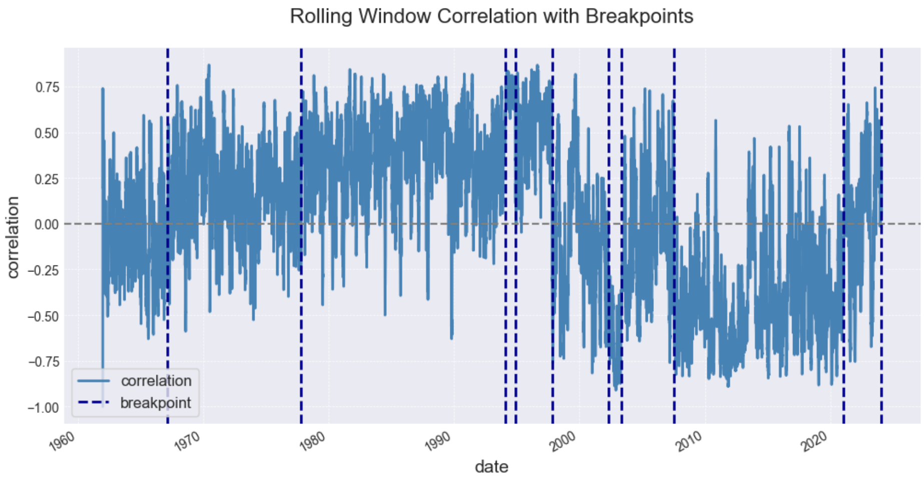


Figure 4.3: Breakpoints Rolling Window Correlation 30 day time window

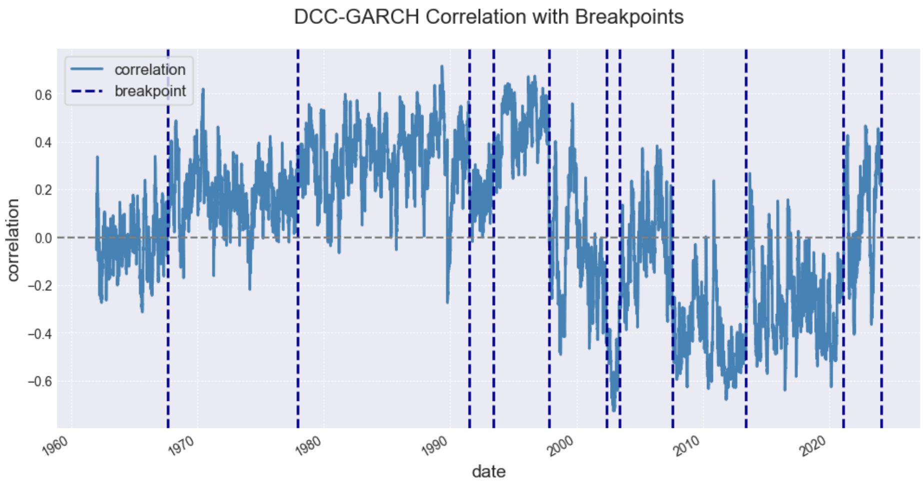


Figure 4.4: Breakpoints DCC(1,1) GARCH Correlation

[7] The same exclusion applies to the breaks of the DCC(1,1) GARCH correlation.
[8] Adjusting the rolling window size might resolve this issue.

4.3.2 Macroeconomic Factors

A different kernel function was used for all additional datasets compared to the one used for detecting breaks in the correlation data. Specifically, a Gaussian kernel was used for the correlation data, whereas a linear kernel was applied to the additional datasets. This decision was based on a visual inspection of applying both kernels to all the datasets and selecting the option that most effectively detected the expected breaks.

The approach for determining the penalties for all additional datasets mirrors the method employed for calibrating the penalties for the correlation data. These penalties are precisely calibrated to identify only larger changes, thereby preventing the model from overfitting with an excessive number of breakpoints. This calibration process includes evaluating multiple penalty values and analyzing the corresponding breakpoints.

The results from the change point detection method applied to the federal funds effective rate are shown in Figure 4.5. A penalty of 100 was chosen for the analysis, and the results proved to be highly robust, as the identified breakpoints remained consistent, with penalties ranging from 81 to 140. As expected, a breakpoint was observed around the mid-1970s. Additionally, an unexpected break was detected before 1970. We predicted a break between the late 1970s and early 1980s, between the late 1980s to early 1990s, around 2000, and around the financial crisis in 2008 which were all confirmed. Additionally, an unanticipated break between 2000 and 2008 appears justified due to the rate's increase followed by a decrease in 2008. Finally, a break in the early 2020s was detected, aligning with our expectations, although this break occurred slightly later than predicted.

Overall, the results align closely with our initial expectations, demonstrating the method's effectiveness in identifying significant changes in the federal funds effective rate.

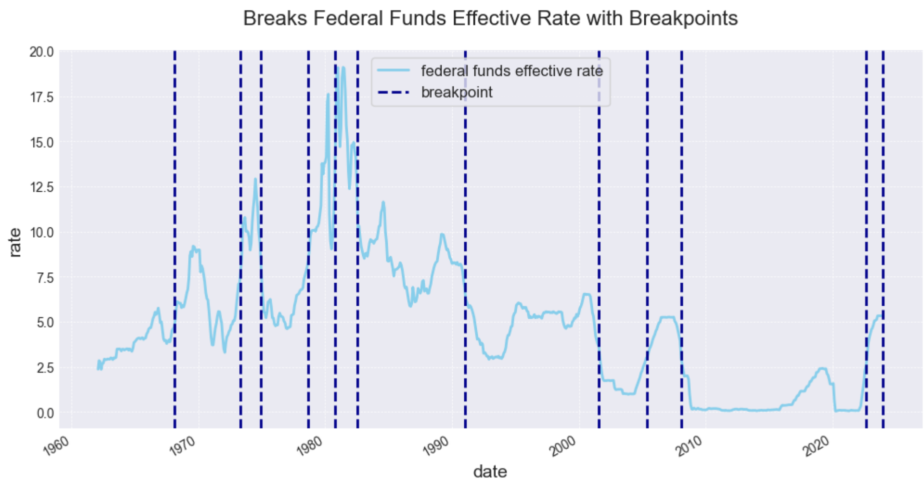


Figure 4.5: Breaks in the Federal Funds Effective Rate

The results from the change point detection method applied to the 10-year expected inflation are presented in Figure 4.6. A penalty of 2 is used for this analysis. The first observed break was around the mid-1980s, which was not anticipated. The second break, around 1986, closely matches our prediction. Although we expected a break around 1995, breaks were detected around 1992 and 1998. These breaks appear to be more accurate than a single break around 1995. Additionally, the break around 1998 captures the expected break around 2000, and the break around 2003 better divides the data. The anticipated breaks around the late 2000s, around 2008 and early 2020s align with the results.

Overall, the results align closely with our initial expectations, demonstrating the method's effectiveness in identifying significant changes in the 10-year expected inflation.

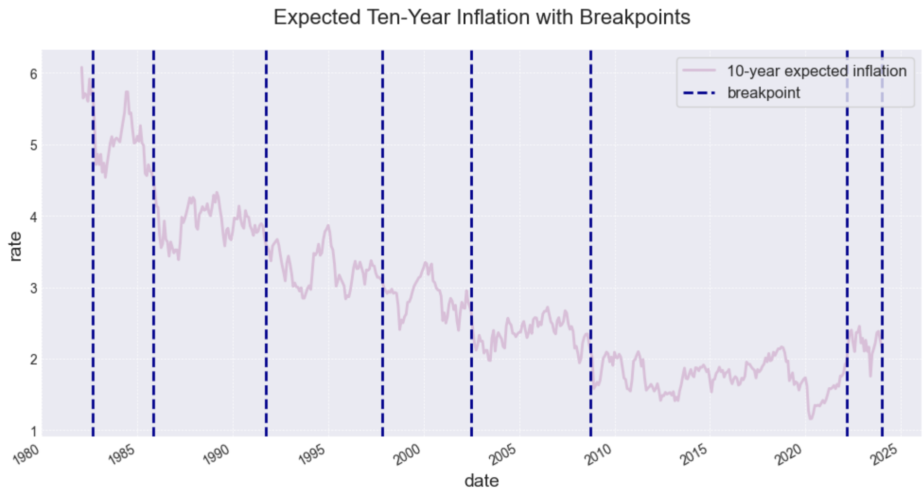


Figure 4.6: Breaks in Expected 10-Year Inflation

The results from the change point detection method applied to the Consumer Price Index (CPI) are shown in Figure 4.7. A penalty of 1 is used for this analysis. The first observed break is around the mid to late 1960s, which was not in line with our earlier expectations but seems logical given the differing average rates in both regimes. The following two breaks, around the mid-1970s, align with our expectations and capture this movement effectively. The fourth break, in the late 1970s, was not expected. The fifth break, in the early 1980s, aligns with our expectations. Although we anticipated a break around the late 1990s, no break was detected there. However, the expected break around the late 2000s was observed in 2008, around the financial crisis, with the model detecting two breaks there. The final breakpoints, around the early 2020s, also align with our expectations.

Overall, while the results align with our initial expectations, they are less consistent than those for the other datasets.

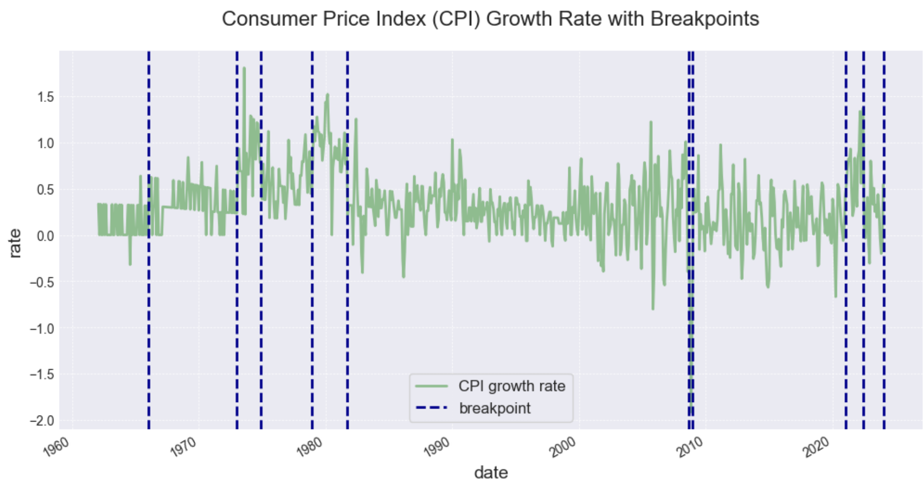


Figure 4.7: Breaks in the Consumer Price Index (CPI)

4.4 F1 Score

4.4.1 Comparison of RWC and DCC(1,1) GARCH

To compare the performance of the Rolling Window Correlation model and the DCC(1,1) GARCH model, we use the F1 score as a metric. The precision and recall are well-defined up to a margin of 240 because this method requires that the minimum distance between true breakpoints (DCC(1,1) GARCH breakpoints) is larger than the maximum margin. Since the minimum distance is 247, we use 240 as the maximum margin.

The F1 scores for various margins are presented in Table 4.3. Initially, at a margin of 30 days, the F1 score is already very high and further increases as the margin increases. This indicates that with a margin of 30 days, the Rolling Window Correlation model and the DCC(1,1) GARCH model exhibit a precision of 0.625, meaning that 62.5% of the breakpoints identified by the Rolling Window Correlation model are also detected by the DCC(1,1) GARCH model. Similarly, a recall of 0.556 implies that 55.6% of the breakpoints identified by the DCC(1,1) GARCH model are also recognized by the Rolling Window Correlation model. With a margin between 210 and 240, the F1 score is 0.824, indicating a very high score that demonstrates the robustness of the detected breaks. Figure 4.8 shows the DCC(1,1) GARCH correlation with the breaks found in the DCC(1,1) GARCH correlation and the breaks found in the RWC. This figure confirms a similarity between the breaks found in both time-varying correlations.

Table 4.3: F1 Score for RWC Breakpoints at Different Margins

Margin (days)	**precision**	**recall**	**F1 score**
30–120	0.625	0.556	0.588
150–180	0.750	0.667	0.706
210–240	0.875	0.778	0.824

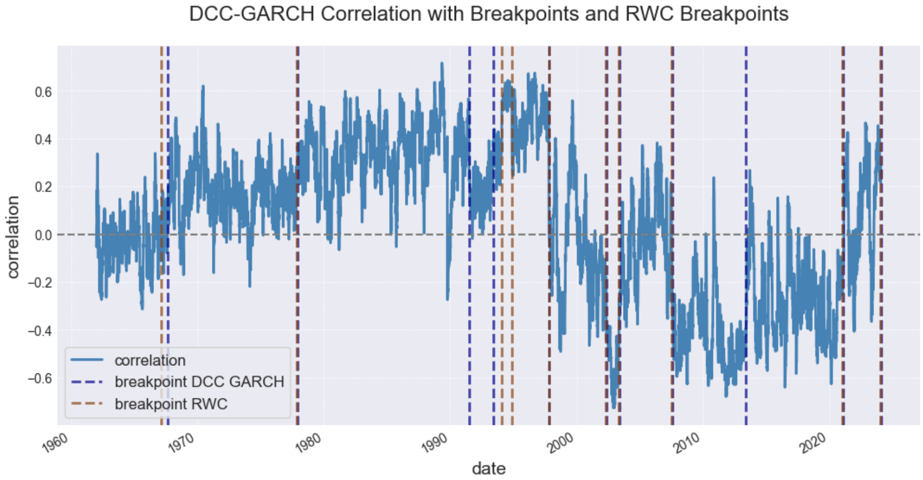


Figure 4.8: DCC GARCH correlation with Breakpoints and RWC Breakpoints

4.4.2 Macroeconomic Factors

The results of the F1 score for the federal funds effective rate, presented in Table 4.4, demonstrate a noticeable improvement as the margin increases. For a margin ranging from 30 to 120 days, the F1 score is 0.087[9], indicating relatively low performance in detecting the true breaks, which are the breaks found in the DCC(1,1) GARCH correlation data. However, when the margin increases to 210-240 days, the F1 score significantly improves to 0.348. This suggests that with a larger margin, the model's ability to accurately identify breaks in the correlation data improves, reflecting higher precision and recall. A graphical representation of the breaks found in the federal funds effective rate and the breaks detected in the DCC(1,1) GARCH correlation is presented in Figure 4.9. This figure illustrates that the federal funds effective rate does not explain the structural breaks identified in the DCC(1,1) GARCH correlation.

Choosing an appropriate margin is crucial for interpreting the influence of the federal funds effective rate on the correlation between stock and bond returns. A smaller margin suggests that the federal funds effective rate does not influence the correlation, as it fails to capture true breaks due to the narrow window. On the other hand, a larger margin could lead to the mistaken conclusion that the federal funds effective rate influences the correlation when this might merely be coincidental.

[9]Considering that change detection methods always place a break at the end of the dataset, all the macroeconomic factors in our case share at least one break with the breaks found in the DCC(1,1) GARCH correlation

Table 4.4: F1 Score for Fed Fund Rate at Different Margins

Margin (days)	precision	recall	F1 score
30–120	0.071	0.111	0.087
150–180	0.143	0.222	0.174
210–240	0.286	0.444	0.348

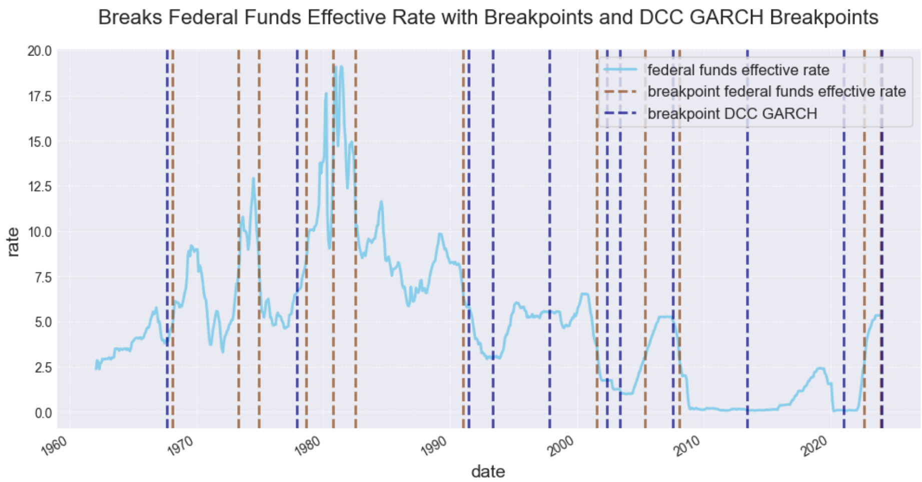


Figure 4.9: Breaks in the Federal Funds Effective Rate and Breaks in the DCC GARCH Correlation

The results of the F1 score for the expected 10-year inflation at different margins, presented in table 4.5, show a similar trend. For a margin of 30, the F1 score is 0.133, which is very low. However, the F1 score improves significantly when the margin is slightly increased. For a margin of 90-240 days, the F1 score becomes 0.4, indicating better detection of true breaks. This suggests that the model's performance in identifying breaks in expected inflation data also benefits from a larger margin. A graphical representation is presented in Figure 4.10, which shows a similarity between the structural breaks in the DCC(1,1) GARCH correlation and the breaks detected in the 10-year expected inflation. However, between 1980 and 1990, there are two breaks in the 10-year expected inflation that are not detected in the DCC(1,1) GARCH correlation.

Table 4.5: F1 Score for 10-Year Expected Inflation at Different Margins

Margin (days)	precision	recall	F1 score
30	0.143	0.125	0.133
60	0.286	0.250	0.267
90–240	0.429	0.375	0.400

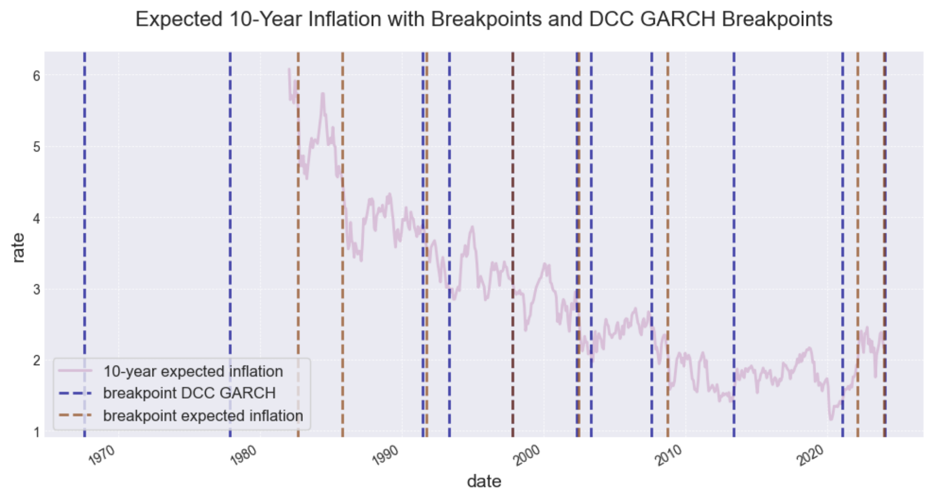


Figure 4.10: Breaks in the Expected 10-year Inflation and Breaks in the DCC GARCH Correlation

The F1 scores for the Consumer Price Index growth rate at various margins are presented in Table 4.6. The results indicate that the F1 score remains low across all margins, suggesting that CPI growth data does not influence the correlation between stocks and bonds. This is confirmed by the breaks shown in Figure 4.11.

Table 4.6: F1 Score for Consumer Price Index growth rate at Different Margins

Margin (days)	**precision**	**recall**	**F1 score**
30–240	0.111	0.100	0.105

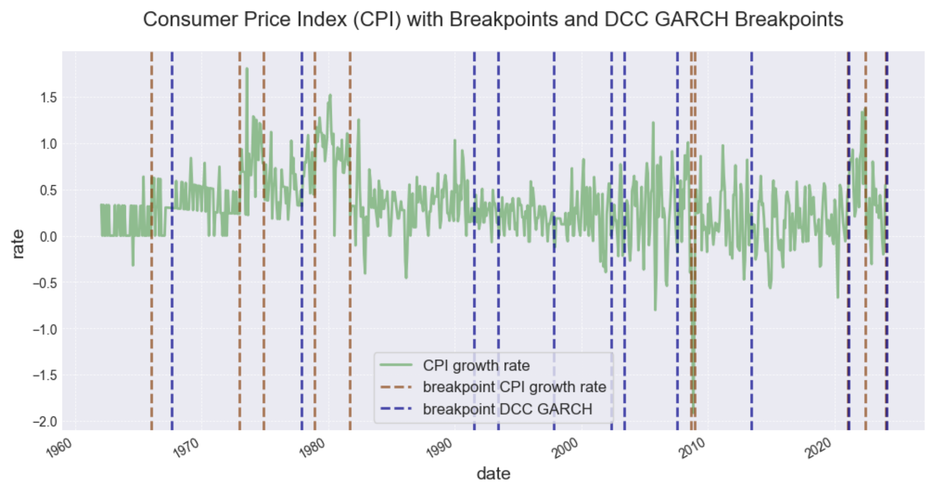


Figure 4.11: Breaks in the CPI Growth Rate and Breaks in the DCC GARCH Correlation

The F1 scores for the Federal Reserve Chairs at various margins are presented in Table 4.7. While it would be ideal to consider a margin of one year before and after a break due to the use of yearly data, the metrics do not allow for this approach because multiple breaks occur within a one-year interval. Nevertheless, the change in Federal Reserve Chairs does not appear to be a reliable indicator of structural breaks in the stock-bond correlation. This is confirmed by the breaks shown in Figure 4.12.

Table 4.7: F1 Score for Fedederal Reserve Chairs at Different Margins

Margin (days)	**precision**	**recall**	**F1 score**
30–120	0.167	0.125	0.143
150–240	0.333	0.250	0.286

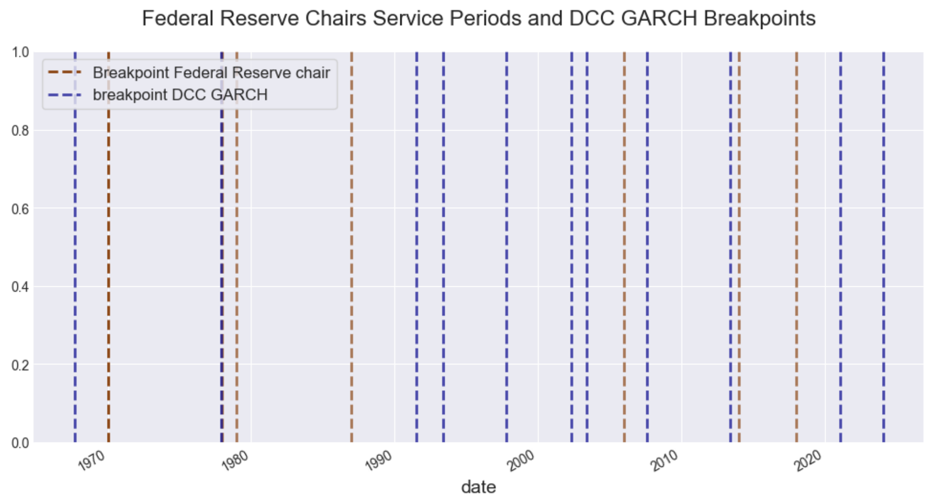


Figure 4.12: Switches of the chair of the Federal Reserve and Breaks in the DCC GARCH Correlation

Examining a margin of 240 days across all datasets reveals that the 10-year expected inflation most closely aligns with the actual breaks identified in the DCC(1,1) GARCH correlation data. The Fed Funds rate follows closely, exhibiting higher recall but lower precision. This means that the 10-year expected inflation is better in the sense that breaks in the 10-year expected inflation are also breaks in the DCC(1,1) GARCH correlation, and the Fed Fund rate has a higher recall because more breaks in the DCC(1,1) GARCH correlation correspond to breaks in the Fed Fund rate. The results vary with the choice of margin. For instance, with a margin of 30, the F1 score is highest for different periods of chairs of the Federal Reserve. However, for all other margins, the F1 score for the 10-year expected inflation remains the highest.

5 Conclusions

The main findings of this research show that the results of the Rolling Window Correlation and the Dynamic Conditional Correlation GARCH (DCC(1,1) GARCH) model appear similar at first glance, a conclusion supported by robustness checks. Nonetheless, the Rolling Window Correlation with a 30-day window exhibits more volatility than the DCC(1,1) GARCH correlation. Both results show significant changes in the correlation over time, suggesting that these correlations vary over time. This answers the first research question which asks if the stock-bond correlation is time-varying. This variability is consistent with prior studies (Engle, 2002; Andersson et al., 2008; Ang and Bekaert, 2002), demonstrating that correlations can fluctuate between positive and negative values. Additionally, in recent years, the stock-bond correlation is changing to a positive correlation again.

Our analysis further reveals that recent shocks to returns significantly affect the variance, and past conditional variances are highly predictive of current variances, indicating strong volatility clustering. While new shocks have a limited effect on altering the correlation, the persistence of these correlations remains high, suggesting that the correlation tends to be relatively stable over time.

Visual inspection suggests a degree of similarity when examining structural breaks in the time-varying correlations of the RWC correlations and the DCC(1,1) GARCH correlation. This is confirmed by the calculated F1 score. Since the breaks are very similar we can say that the detected breaks are very robust.

Due to its reduced noise, the DCC(1,1) GARCH model was used for further analysis to identify macroeconomic factors influencing the stock-bond correlation. Change point detection methods applied to the Consumer Price Index (CPI), federal funds effective rate, and 10-year expected inflation revealed multiple breaks in all three datasets. The switch of the Federal Reserve chairs was also considered. The F1 score was calculated for each factor to assess its ability to predict structural breaks in the stock-bond correlation. Among these, the 10-year expected inflation had the highest F1 score of 0.400. This indicates that it was the most predictive of structural breaks. This is in line with the result of Ilmanen (2003). All other macroeconomic factors had a lower F1 score. This leads to the conclusion that CPI, the federal funds effective rate, the switch of Federal Reserve chairs, and the 10-year expected inflation do not fully account for the observed breaks in the correlation data. This suggests that there are no structural breaks caused by macroeconomic factors. Which answers the second research question which asks if there are any macroeconomic factors that cause a structural break in the stock-bond correlation.

In summary, our analysis confirms the time-varying nature of stock-bond correlations, identifies that there are no macroeconomic factors tested in this thesis that can explain structural breaks in the stock-bond correlation, and highlights the use performance of the DCC(1,1) GARCH model which offers a robust tool for filtering noise and capturing the true dynamics of the stock-bond relationship, making it valuable for financial analysis, risk management, and policy decision-making.

6 Discussion

This study aimed to investigate the time-varying correlation between stock and bond returns and identify the macroeconomic factors influencing this correlation by detecting structural breakpoints. Understanding that stock-bond correlations are time-varying is crucial for investors, monetary policymakers, and risk managers. For investors, this insight helps construct more resilient portfolios that account for the dynamic nature of asset correlations. For risk managers, it allows for better assessing and mitigating risks associated with portfolio diversification and asset allocation. For policymakers, recognizing the factors that influence these correlations, even indirectly, can aid in better forecasting and managing economic conditions. For example, central banks could use these findings to anticipate and mitigate the impact of economic shocks on financial markets. For example, knowing that the federal funds effective rate does not cause structural breaks in the stock-bond correlation will help in deciding how to set the federal funds effective rate. Investors with a standard 60/40 portfolio also want to know if their allocation is negatively correlated, suggesting a hedge in a portfolio, or if they are positively correlated, which would increase the portfolio's risk.

This research is especially relevant this year due to the timing of the U.S. elections in November 2024. Former President Trump has been outspoken in his criticism of Federal Reserve Chair Jerome Powell, accusing him of intending to lower interest rates to benefit the Democrats (Goldman and Ross, 2024). Consequently, Powell might not be reappointed if Trump is elected. If the chairs of the Federal Reserve significantly influences structural breaks in the stock-bond correlation, the political and economic tensions could have substantial impacts on the financial markets. However, according to the results of this thesis, this is not the case.

This study has multiple strengths, such as its comprehensive approach to modeling the time-varying correlations using an advanced DCC(1,1) GARCH model and a large dataset spanning multiple decades. However, this approach also had some limitations.

One limitation is the window size of the rolling window correlation method since its effectiveness is highly dependent on this chosen window size. A different window size might yield better alignment with the breakpoints found by the DCC(1,1) GARCH model. The rolling window approach captures short-term variations but can be sensitive to the window length. Future research could explore optimal window sizes or adaptive window techniques that adjust based on data characteristics to enhance the accuracy of break detection. Comparing the practical implications of using Rolling Window Correlation and DCC(1,1) GARCH models shows different advantages and challenges. The rolling window approach is straightforward and computationally less intensive, making it suitable for real-time applications. However, the DCC(1,1) GARCH model offers a more sophisticated and dynamic approach to capturing time-varying correlations. One must weigh these trade-offs when choosing a method for correlation estimation in asset- and risk management.

Another limitation of this study is the assumption of normality for the error term in the DCC(1,1) GARCH model. This assumption does not hold, as seen in the normality test. This implication is very significant since the results could be completely different ff normality is assumed, when this does not hold. Section 7 presents a possible solution for this.

Determining which penalty to use in change-point detection methods is subjective, representing a significant shortcoming. As highlighted by Celisse et al. (2018), varying penalties can lead to different detection results. This subjectivity introduces biases and inconsistencies in identifying structural breaks, affecting the robustness of the results. Future research could look at developing more automated methods for penalty selection to enhance the reliability of change-point detection outcomes. The penalty term exhibits significant sensitivity to changes, as does the bandwidth parameter, γ. Variations in γ greatly influence the number of detected breakpoints. The choice of cost function, kernel, and optimization method also considerably impacts the results.

The F1 score, as used in this study, is limited in its ability to detect multiple breakpoints within a close margin of a true breakpoint. This limitation suggests a need for adapting the F1 score or developing new metrics that can better capture the multiplicity and proximity of breakpoints. Improved metrics provide a more nuanced evaluation of factors that could result in a structural break in the stock-bond correlation.

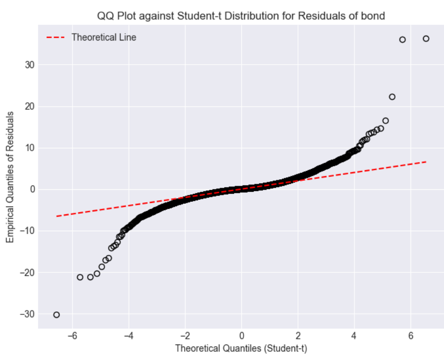
The findings of this thesis have broader implications for financial theory and practice. Understanding the dynamics of the stock-bond correlation and the factors driving structural breaks can inform asset allocation, risk management, and policy-making, as mentioned.

In summary, this thesis provides valuable insights into the time-varying correlation between bonds and stocks, the macroeconomic vectors influencing these correlations. Future research should continue to extend these models and explore their applications in broader financial contexts.

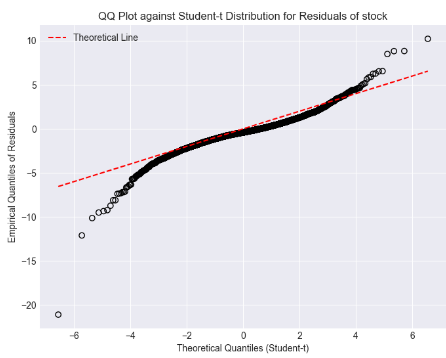
7 Proposal for Further Research

7.1 Different Distribution

In this thesis, we initially assume that the residuals follow a normal distribution. However, subsequent tests reveal that this assumption does not hold true, suggesting that another distribution might better fit the data. Orskaug (2009) discusses the possibility of achieving a better fit by modifying the log-likelihood functions. One alternative that might fit the data better is the Student-t distribution. Nevertheless, as shown in Figure 7.1, which displays Q-Q plots for the residuals of log bond and stock returns against the Student-t distribution, this distribution also fails to provide a perfect fit. The Q-Q plots reveal significant deviations at the tails, indicating the presence of heavy tails or outliers that the Student-t distribution does not fully capture. Figure 7.2 further highlights that for log bond returns with shorter maturities, the fit deviates even more from the Student-t distribution. The Q-Q plot for the bond with a 30-year maturity closely resembles that of the residuals of the log bond returns for the 10-year maturity. Therefore, future research may need to explore other distributions or more advanced modeling techniques to better represent the residuals.

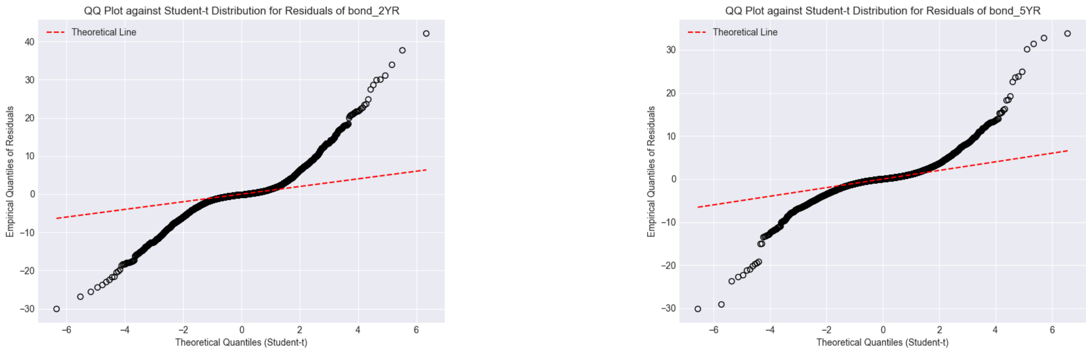


(a) Q-Q Plot for the Residuals of Log Bond Returns



(b) Q-Q Plot for the Residuals of Log Stock Returns

Figure 7.1: Q-Q Plots for Bond and Stock Residuals



(a) Q-Q Plot for the Residuals of Log Bond Returns (2 Year maturity)

(b) Q-Q Plot for the Residuals of Log Bond Returns (5 Year maturity)

Figure 7.2: Q-Q Plots for Bond Residuals with different maturities

7.2 Finding Breakpoints in Real Time

To make effective investment decisions, it is crucial to identify breakpoints in the data quickly and in real time. Future research should focus on assessing how rapidly these breakpoints can be detected using real-time data. One approach could be to apply the process of detecting breakpoints incrementally, using data up to various points in time and applying online change-point detection methods. We could then compare these results with the result of this thesis. This would allow us to determine whether breakpoints can be identified within days or if it takes several years.

If the exact date of the breakpoint is not necessary, a margin can be used to evaluate whether the newly detected date falls within an acceptable range of the actual breakpoint. This approach provides flexibility in the detection process and can help if precision is not very necessary.

7.3 Look at Different Data

Another area for future research could be to recalculate the DCC(1,1) GARCH correlation after specific events. For instance, the DCC(1,1) GARCH correlation could be calculated for periods corresponding to different Federal Reserve chairman's periods, analyzing the correlation during each chairman's term.

Additionally, this approach could be applied to periods of varying inflation rates. By comparing the DCC(1,1) GARCH correlations during periods of low inflation with those from periods of high inflation, researchers can also assess if inflation impacts the correlation.

Another idea for future research is to redo the analysis using different datasets, such as bond and stock indices from European countries or China. This would help determine if the findings of this paper hold across different markets and economic environments.

Bibliography

A'Hearn, B. and U. Woitek (2001): "More international evidence on the historical properties of business cycles," *Journal of Monetary Economics*, 47, 321–346.

Andersson, M., E. Krylova, and S. Vähämaa (2008): "Why does the correlation between stock and bond returns vary over time?" *Applied Financial Economics*, 18, 139–151.

Ang, A. and G. Bekaert (2002): "International asset allocation with regime shifts," *The review of financial studies*, 15, 1137–1187.

Arlot, S., A. Celisse, and Z. Harchaoui (2012): "A kernel multiple change-point algorithm via model selection," *arXiv preprint arXiv:1202.3878*.

Auger, I. E. and C. E. Lawrence (1989): "Algorithms for the optimal identification of segment neighborhoods," *Bulletin of mathematical biology*, 51, 39–54.

Baele, L., G. Bekaert, and K. Inghelbrecht (2010): "The determinants of stock and bond return comovements," *The Review of Financial Studies*, 23, 2374–2428.

Bautista, C. C. (2003): "Interest rate-exchange rate dynamics in the Philippines: a DCC analysis," *Applied Economics Letters*, 10, 107–111.

Bauwens, L., S. Laurent, and J. V. Rombouts (2006): "Multivariate GARCH models: a survey," *Journal of Applied Econometrics*, 21, 79–109.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of econometrics*, 31, 307–327.

——— (1990): "Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model," *The Review of Economics and Statistics*, 498–505.

Campbell, J. Y. and J. Ammer (1993): "What moves the stock and bond markets? A variance decomposition for long-term asset returns," *The Journal of Finance*, 48, 3–37.

Cappiello, L., P. Hördahl, A. Kadareja, and S. Manganelli (2006): "The Impact of the Euro on Financial Markets," *ECB working paper*.

Celik, S. (2012): "The more contagion effect on emerging markets: The evidence of DCC-GARCH model," *Economic Modelling*, 29, 1946–1959.

Celisse, A., G. Marot, M. Pierre-Jean, and G. Rigaill (2018): "New efficient algorithms for multiple change-point detection with reproducing kernels," *Computational Statistics & Data Analysis*, 128, 200–220.

Chen, C. W., M. K. So, and E. M. Lin (2009): "Volatility forecasting with double Markov switching GARCH models," *Journal of Forecasting*, 28, 681–697.

Chiang, T. C., B. N. Jeon, and H. Li (2007): "Dynamic correlation analysis of financial contagion: Evidence from Asian markets," *Journal of International Money and finance*, 26, 1206–1228.

Cho, J. H. and A. M. Parhizgari (2009): "East Asian financial contagion under DCC-GARCH," *International Journal of Banking and Finance*, 6, 17–30.

Connolly, R., C. Stivers, and L. Sun (2005): "Stock market uncertainty and the stock-bond return relation," *Journal of Financial and Quantitative Analysis*, 40, 161–194.

Cont, R. (2001): "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative finance*, 1, 223.

Czasonis, M., M. Kritzman, and D. Turkington (2020): "The Stock-Bond Correlation," *The Journal of Portfolio Management.*

David, A. and P. Veronesi (2001): "Inflation and earnings uncertainty and the volatility of asset prices: an empirical investigation," *Manuscript, Graduate School of Business, University of Chicago.*

Engle, R. (2002): "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, 20, 339–350.

Engle, R. F. and K. F. Kroner (1995): "Multivariate simultaneous generalized ARCH," *Econometric theory*, 11, 122–150.

European Central Bank (2022a): "Cross-asset correlations in a more inflationary environment and challenges for diversification strategies," *European Central Bank: Financial Stability Review.*

——— (2022b): "Financial Stability Review, November 2022," Accessed: 2024-06-06.

Forbes, K. J. and R. Rigobon (2002): "No contagion, only interdependence: measuring stock market comovements," *The journal of Finance*, 57, 2223–2261.

Garreau, D. and S. Arlot (2018): "Consistent change-point detection with kernels," .

Glynn, J., N. Perera, and R. Verma (2007): "Unit root tests and structural breaks: A survey with applications," .

Goldman, D. and A. Ross (2024): "Trump says he would not reappoint Fed Chair Jerome Powell," *CNN*, updated 11:04 AM EST, Fri February 2, 2024.

Gulko, L. (2002): "Decoupling," *Journal of Portfolio Management*, volume 28, 59.

Harchaoui, Z. and O. Cappé (2007): "Retrospective mutiple change-point estimation with kernels," in *2007 IEEE/SP 14th Workshop on Statistical Signal Processing*, IEEE, 768–772.

Hsu, D. (1984): "The behavior of stock returns: is it stationary or evolutionary?" *Journal of Financial and Quantitative Analysis*, 19, 11–28.

Ilmanen, A. (2003): "Stock-bond correlations," *The Journal of Fixed Income*, 13, 55.

Jackson, B., J. D. Scargle, D. Barnes, S. Arabhi, A. Alt, P. Gioumousis, E. Gwin, P. Sangtrakulcharoen, L. Tan, and T. T. Tsai (2005): "An algorithm for optimal partitioning of data on an interval," *IEEE Signal Processing Letters*, 12, 105–108.

Killick, R., P. Fearnhead, and I. A. Eckley (2012): "Optimal detection of change-points with a linear computational cost," *Journal of the American Statistical Association*, 107, 1590–1598.

Lee, M.-C., J.-S. Chiou, and C.-M. Lin (2006): "A study of value-at-risk on portfolio in stock return using DCC multivariate GARCH," *Applied Financial Economics Letters*, 2, 183–188.

Li, L. (2002): "Correlation of stock and bond returns," *Manuscript, Department lf Economics, Yale University.*

Lo, A. W. and A. C. MacKinlay (2011): *A non-random walk down Wall Street*, Princeton University Press.

Manera, M., M. McAleer, and M. Grasso (2006): "Modelling time-varying conditional correlations in the volatility of Tapis oil spot and forward returns," *Applied Financial Economics*, 16, 525–533.

Nazir, F., N. Mohamad, and H. Bahaludin (2021): "Comparison of Correlation for Asia Shariah Indices using DCC-GARCH and Rolling Window Correlation," *Applied Mathematics and Computational Intelligence*, 10, 203–219.

Orskaug, E. (2009): "Multivariate dcc-garch model:-with various error distributions," Master's thesis, Institutt for matematiske fag.

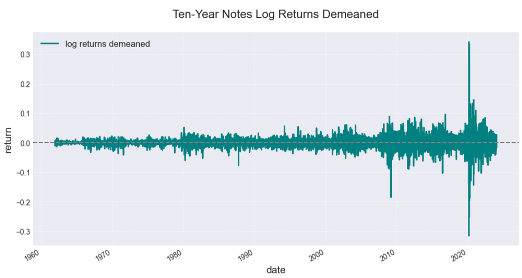
Pericoli, M. (2018): "Macroeconomics determinants of the correlation between stocks and bonds," *Bank of Italy Temi di Discussione (Working Paper) No*, 1198.

Shawe-Taylor, J. and N. Cristianini (2004): *Kernel methods for pattern analysis*, Cambridge university press.

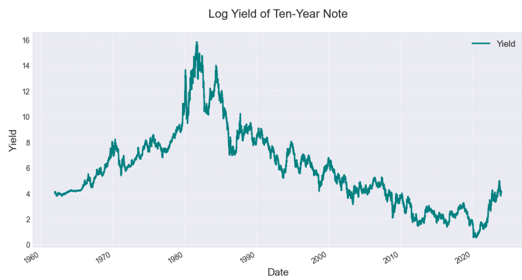
Truong, C., L. Oudre, and N. Vayatis (2020): "Selective review of offline change point detection methods," *Signal Processing*, 167, 107299.

A Data visualization

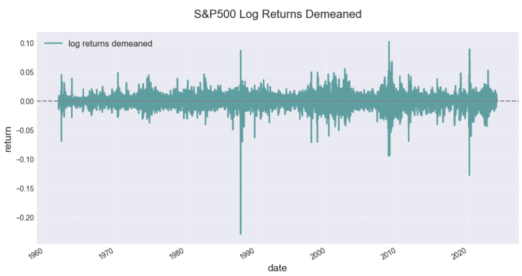
The following representation includes the stock data and bond data used. It displays the yields of the 10-year notes and the closing prices of the S&P 500, alongside the demeaned log returns of both datasets.



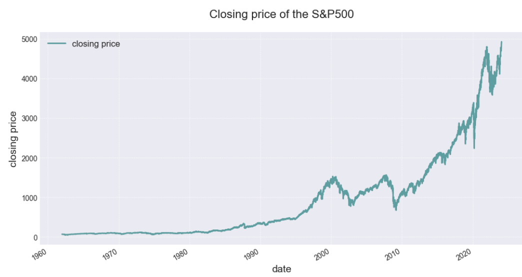
(a) Log Bond Returns Demeaned



(b) Yield Log



(c) Log Returns Demeaned Stock



(d) Close price Stock

Figure A.1: Bond and Stock data

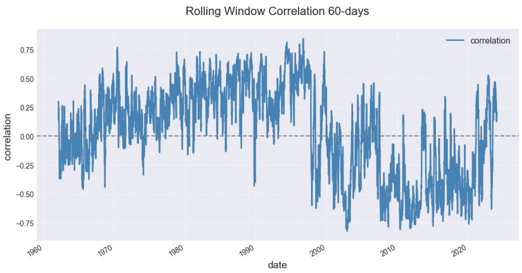
B Cost function: kernel

Equation B.1 shows the derivation of the cost function of a general kernel.

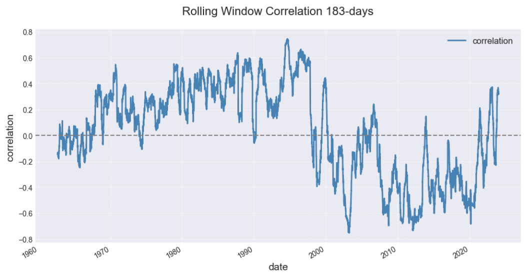
$$
\begin{aligned}
c_{\text{kernel}}(y_{a..b}) &= \sum_{t=a+1}^{b} \|\phi(y_t) - \bar{\mu}_{a..b}\|_{\mathcal{H}}^2 \\
&= \sum_{t=a+1}^{b} \langle \phi(y_t) - \bar{\mu}_{a..b}, \phi(y_t) - \bar{\mu}_{a..b} \rangle_{\mathcal{H}} \\
&= \sum_{t=a+1}^{b} \langle \phi(y_t), \phi(y_t) \rangle_{\mathcal{H}} - 2 \sum_{t=a+1}^{b} \langle \bar{\mu}_{a..b}, \phi(y_t) \rangle_{\mathcal{H}} + \sum_{t=a+1}^{b} \langle \bar{\mu}_{a..b}, \bar{\mu}_{a..b} \rangle_{\mathcal{H}} \\
&\text{Using that } \bar{\mu}_{a..b} = \frac{1}{b-a} \sum_{s=a+1}^{b} \phi(y_s)\text{, we get} \\
&= \sum_{t=a+1}^{b} \langle \phi(y_t), \phi(y_t) \rangle_{\mathcal{H}} - \frac{2}{b-a} \sum_{s,t=a+1}^{b} \langle \phi(y_s), \phi(y_t) \rangle_{\mathcal{H}} + \frac{1}{b-a} \sum_{s,t=a+1}^{b} \langle \phi(y_t), \phi(y_s) \rangle_{\mathcal{H}} \\
&= \sum_{t=a+1}^{b} \langle \phi(y_t), \phi(y_t) \rangle_{\mathcal{H}} - \frac{1}{b-a} \sum_{s,t=a+1}^{b} \langle \phi(y_s), \phi(y_t) \rangle_{\mathcal{H}} \\
&\text{Using 3.12, we get} \\
&= \sum_{t=a+1}^{b} k(y_t, y_t) - \frac{1}{b-a} \sum_{s,t=a+1}^{b} k(y_s, y_t).
\end{aligned}
\tag{B.1}
$$

C Rolling Window Correlation for different time windows

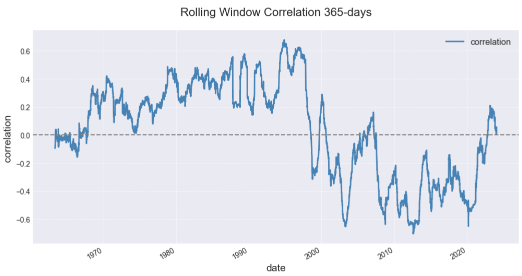
The following figures show the Rolling Window correlation calculated using different time windows.



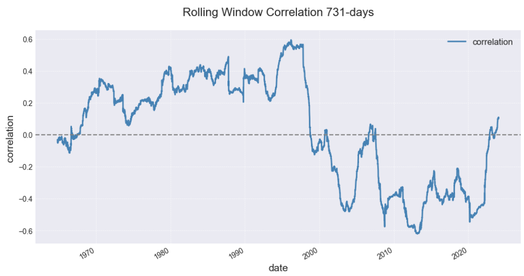
(a) Rolling Window Correlation for a 60 day window



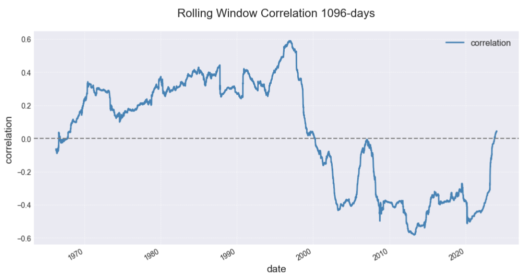
(b) Rolling Window Correlation for a 183 day window



(c) Rolling Window Correlation for a 365 day window



(d) Rolling Window Correlation for a 731 day window



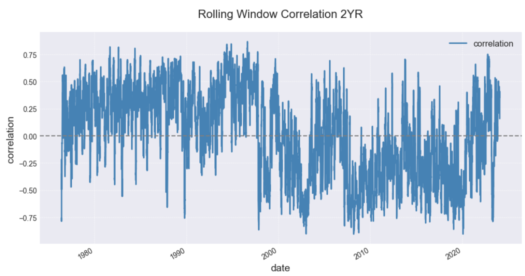
(e) Rolling Window Correlation for a 1096 day window

Figure C.1: Correlations between stocks and bonds using various time windows

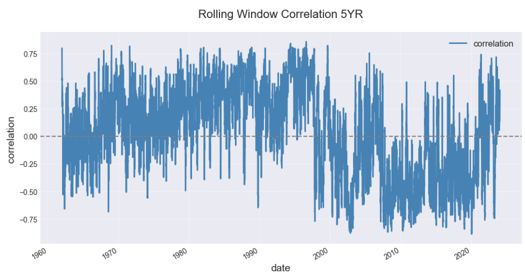
D Robustness

D.1 Rolling Window Correlation

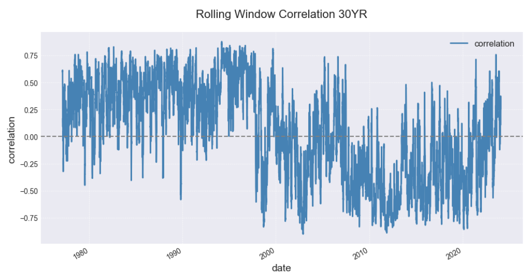
The following figures show the Rolling Window Correlation using a 30-day time window but using different datasets for the bonds. Namely bond data with a 2-year, 5-year and 30-year maturity. This is done as a robustness check.



(a) Rolling Window Correlation using the two-year note



(b) Rolling Window Correlation using the five-year note

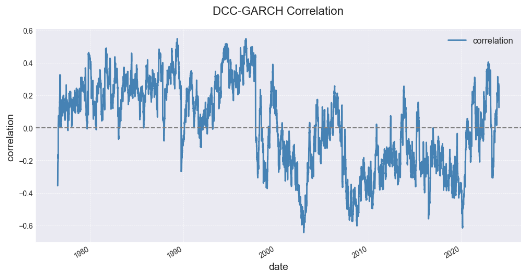


(c) Rolling Window Correlation using the thirty-year note

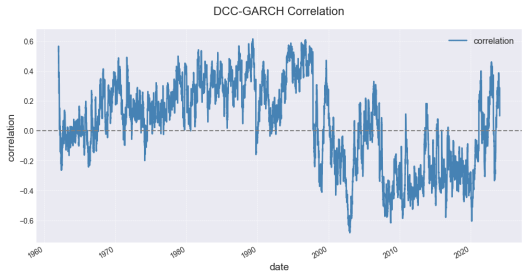
Figure D.1: Correlations between stocks and bonds using various bond data

D.2 DCC(1,1) GARCH

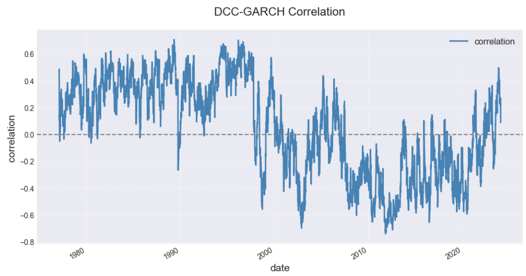
The same process is applied to the DCC(1,1) GARCH correlations. Below, the correlations for various bond datasets are presented along with the estimated parameters used to calculate these.



(a) DCC(1,1) GARCH Correlation using 2-year Bond data



(b) DCC(1,1) GARCH Correlation using 5-year Bond data



(c) DCC(1,1) GARCH Correlation using 30-year Bond data

Figure D.2: Correlations between stocks and bonds using various bond data

Table D.2: Parameter Estimates for Different Bond Maturities with Kappa and Theta

	Estimate	**T-Statistic**
2 YEAR		
κ	0.0216	9.596***
θ	0.9730	3.668***
5 YEAR		
κ	0.0239	6.790***
θ	0.9715	3.742***
30 YEAR		
κ	0.0322	6.275***
θ	0.9642	3.864***

Note: This table displays the parameter estimates and their corresponding t-statistics for kappa and theta for different bond maturities. T-statistics are based on a bootstrap of 50. Significance levels: $*** : p < 0.01$.

Table D.1: Parameter Estimates for Different Bond Maturities and Stocks Returns with T-Statistics

	Bond	**T-Statistic**	**Stock**	**T-Statistic**
2 YEAR				
γ	0.0011	182.48***	0.0040	70.69***
ϕ	0.1000	1.80*	9.9968e-2	1.40
ω	5.7385e-06	4.95e-2	9.778e-07	0.20
α	0.1982	4303.16***	0.1975	4265.06***
β	0.7977	277.21***	0.7991	346.82***
5 YEAR				
γ	4.77e-4	28.39***	3.147e-3	73.08***
ϕ	0.1000	1.02	9.9981e-2	1.01
ω	2.1900e-06	0.16	6.2100e-07	0.14
α	0.2113	1830.59***	0.2109	1826.22***
β	0.7985	148.75***	0.7999	176.52***
10 YEAR				
γ	2.41e-4	134.19***	0.0031	162.08***
ϕ	0.1000	1.47	9.9987e-2	1.01
ω	1.0229e-06	0.27	5.1253e-07	0.16
α	0.2167	6454.89***	0.2165	6453.38***
β	0.7997	569.15***	0.8002	612.37***
30 YEAR				
γ	7.80e-4	127.24***	4.399e-3	378.70***
ϕ	0.1000	1.07	9.9977e-2	1.84*
ω	3.7000e-06	0.16	1.2500e-06	0.55
α	0.1803	2042.07***	0.1803	2033.00***
β	0.7977	170.62***	0.7976	156.52***

Note: $* : p < 0.1, ** : p < 0.05, *** : p < 0.01$. This table displays parameter estimates and their corresponding t-statistics categorized under different bond maturities and stock.

E Additional results Jarque-Bera test

Furthermore, a Jarque-Bera test was performed to check the normality of the residuals in Andersson et al. (2008), given their assumption of normality, which is also assumed in this study. This is done for the whole length of the datasets as used in the paper of Andersson et al. (2008) in Table E.1. In Table E.2, the results are presented for the analysis using the time frame of Andersson et al. (2008) and in Table E.3 the results are presented for the the analysis using the entire dataset used in this thesis.

Table E.1: Jarque-Bera Test Results for Bond and Stock returns using the timeframe of Andersson et al. (2008)

	Bond	**Stock**
Statistic	1135.34	2356.80
p-value	2.9144e-247	0.0

Note: This table displays the Jarque-Bera test statistics and their corresponding p-values for the residuals of log returns for both bonds and stocks.

Table E.2: Jarque-Bera Test Results for Different Periods using the timeframe of Andersson et al. (2008)(Log-transformed Data)

Period	**Description**	**Jarque-Bera for Bonds**		**Jarque-Bera for Stocks**	
		Statistic	p-value	Statistic	p-value
1991-01-01 to 1999-12-31	Pre Dot-com bubble	688.38	3.32e-150***	3432.15	0.0***
2000-01-01 to 2002-12-31	Dot-com bubble burst	138.58	8.10e-31***	42.06	7.37e-10***
2003-01-01 to 2005-12-31	Post Dot-com bubble burst	70.40	5.17e-16***	59.67	1.10e-13***
1991-01-01 to 2000-12-31	Before 9/11	676.26	1.42e-147***	2728.45	0.0***
2001-01-01 to 2002-12-31	Around 9/11	48.05	3.69e-11***	28.78	5.64e-7***
2003-01-01 to 2005-12-31	After 9/11	70.40	5.17e-16***	59.67	1.10e-13***
1991-01-01 to 1995-12-31	Early 90s recession recovery	333.73	3.39e-73***	337.78	4.48e-74***
1996-01-01 to 2000-12-31	Economic expansion	302.73	1.83e-66***	620.01	2.32e-135***
2001-01-01 to 2002-12-31	Early 2000s recession	48.05	3.69e-11***	28.78	5.64e-7***
2003-01-01 to 2005-12-31	Recovery from early 2000s recession	70.40	5.17e-16***	59.67	1.10e-13***
1991-01-01 to 1994-12-31	Pre-1992 election	304.73	6.76e-67***	243.64	1.24e-53***
1995-01-01 to 1998-12-31	Pre-1996 election	451.41	9.51e-99***	2723.02	0.0***
1999-01-01 to 2002-12-31	Pre-2000 election	160.60	1.34e-35***	51.66	6.06e-12***
2003-01-01 to 2005-12-31	Pre-2004 election	70.40	5.17e-16***	59.67	1.10e-13***
1991-01-01 to 1993-12-31	1991-1993	15.18	5.05e-4***	211.99	9.26e-47***
1994-01-01 to 1996-12-31	1994-1996	462.42	3.86e-101***	137.84	1.17e-30***
1997-01-01 to 1999-12-31	1997-1999	106.71	6.72e-24***	478.45	1.27e-104***

Note: This table summarizes the results of the Jarque-Bera tests for bond and stock returns over various periods, using the overall time frame of Andersson et al. (2008). The Jarque-Bera statistic tests the null hypothesis that the data follows a normal distribution. Significance levels: $***: p < 0.01$.

Table E.3: Jarque-Bera Test Results for Different Periods using the whole dataset (Log-transformed Data)

Period	**Description**	**Jarque-Bera for Bonds**		**Jarque-Bera for Stocks**	
		Statistic	p-value	Statistic	p-value
1991-01-01 to 1999-12-31	Pre Dot-com bubble	684.05	2.88e-149***	3434.27	0.0***
2000-01-01 to 2002-12-31	Dot-com bubble burst	138.35	9.09e-31***	42.06	7.37e-10***
2003-01-01 to 2005-12-31	Post Dot-com bubble burst	70.17	5.78e-16***	59.67	1.10e-13***
1991-01-01 to 2000-12-31	Before 9/11	672.30	1.03e-146***	2729.80	0.0***
2001-01-01 to 2002-12-31	Around 9/11	47.95	3.88e-11***	28.78	5.64e-7***
2003-01-01 to 2005-12-31	After 9/11	70.17	5.78e-16***	59.67	1.10e-13***
1991-01-01 to 1995-12-31	Early 90s recession recovery	330.27	1.92e-72***	338.16	3.71e-74***
1996-01-01 to 2000-12-31	Economic expansion	302.13	2.47e-66***	619.99	2.35e-135***
2001-01-01 to 2002-12-31	Early 2000s recession	47.95	3.88e-11***	28.78	5.64e-7***
2003-01-01 to 2005-12-31	Recovery from early 2000s recession	70.17	5.78e-16***	59.67	1.10e-13***
1991-01-01 to 1994-12-31	Pre-1992 election	300.93	4.51e-66***	244.04	1.02e-53***
1995-01-01 to 1998-12-31	Pre-1996 election	450.33	1.63e-98***	2723.00	0.0***
1999-01-01 to 2002-12-31	Pre-2000 election	160.40	1.48e-35***	51.66	6.06e-12***
2003-01-01 to 2005-12-31	Pre-2004 election	70.17	5.78e-16***	59.67	1.10e-13***
1991-01-01 to 1993-12-31	1991-1993	14.54	6.98e-4***	212.47	7.28e-47***
1994-01-01 to 1996-12-31	1994-1996	462.08	4.59e-101***	137.83	1.18e-30***
1997-01-01 to 1999-12-31	1997-1999	106.37	7.98e-24***	478.44	1.28e-104***

Note: This table summarizes the results of the Jarque-Bera tests for bond and stock returns over various periods. The Jarque-Bera statistic tests the null hypothesis that the data follows a normal distribution. Significance levels: $***: p < 0.01$.