INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Dynamics in Stock Return Volatility and Spillover Effects: Does Firm Size Matter?

Francisco da Silva Antunes

Master in Finance

Supervisor:
PhD José Joaquim Dias Curto, Associate Professor with Aggregation, ISCTE-IUL

October, 2021

Department of Finance

Dynamics in Stock Return Volatility and Spillover Effects: Does Firm Size Matter?

Francisco da Silva Antunes

Master in Finance

Supervisor:
PhD José Joaquim Dias Curto, Associate Professor with Aggregation, ISCTE-IUL

October, 2021

Acknowledgements

Throughout the research and development of this dissertation, I have received continued encouragement and assistance, for which I am deeply grateful.

I would like to thank my supervisor José Joaquim Dias Curto, whose expertise, availability, and guidance helped me raising this work to a new level. I would also like to thank Beatriz Freire, for her readiness to take my concerns into account and critical insights throughout this journey.

I will always be grateful to Inês Lopes for her confidence, support, and love during all these years. Last but not least, I would particularly like to thank my parents, Helena and João, for their individual sacrifices and utmost dedication over the past twenty-three years.

Resumo

Esta dissertação investiga a relação entre a dimensão de uma empresa e a volatilidade das ações. Partindo de uma base sólida de conhecimentos que indicam a existência de um fator dimensão em Banz (1981) e Reinganum (1981), testamos dois temas conexos. Inicialmente, se a assimetria na volatilidade das rendibilidades está associada ao tamanho da empresa e ao setor em que opera. Posteriormente, a existência de efeitos de repercussão entre empresas com elevado e baixo valor de mercado. Ao comparar diferentes modelos ARCH, descobrimos que boas e más notícias geram impactos distintos na volatilidade, já que as especificações assimétricas são superiores ao modelo GARCH simétrico. Os resultados mostram ainda que esta dinâmica é mais significativa em grandes empresas, mas que pode variar conforme o setor. Este estudo revela também um elevado co-movimento entre grandes e pequenas empresas, já que ambas demonstram dinâmicas similares na volatilidade e uma forte correlação, especialmente em períodos de forte instabilidade financeira. Desta forma, através de um GARCH multivariado, acabamos por confirmar um comportamento assimétrico não só na volatilidade, mas também na correlação. Os modelos ARCH implementados modelam a volatilidade dos retornos do índice S&P500, do índice Russel 2000 e dos respetivos sectores para o período 2006-2020.

Classificação JEL:

C58, G10

Palavras-Chave:

Volatilidade, Dimensão da Empresa, Efeitos de Repercussão, GARCH Univariado, GARCH Multivariado.

Abstract

This dissertation explores the relationship between firm size and equity volatility. Working from a solid ground that indicates the existence of a size effect in Banz (1981) and Reinganum (1981), we test two related topics. At first, whether an asymmetric effect in returns' volatility is linked to firm size and firm's industrial sector. Secondly, the existence of spillover effects between large and small market capitalization firms. By comparing different ARCH type models, we find that good and bad news have different impacts on volatility, as the asymmetric specifications outperform the symmetric GARCH model. In addition, our empirical results show that such dynamic is stronger in large-cap firms, but it may vary according to the firm's sector. This study also signals a strong co-movement between large and small firms, as both display a similar dynamic in volatility and a strong correlation, especially in periods of financial turmoil. Using multivariate GARCH models, we end up unveiling an asymmetric behaviour not only in volatility but also in conditional correlation. All ARCH type models are employed to the returns' volatility of S&P500 Index, Russel 2000 Index, and their respective industry sectors over the 2006-2020 period.

JEL Classification:

C58, G10

Keywords:

Volatility, Firm Size, Spillover Effects, Univariate GARCH, Multivariate GARCH.

Table of Contents

List of Figures

x

List of Tables

Glossary

ACF - Autocorrelation Function

A-DCC - Asymmetric Dynamic Conditional Correlation

ADF - Augmented Dickey-Fuller

AG-DCC - Asymmetric Generalized Dynamic Conditional Correlation

AIC - Akaike Information Criteria

APARCH - Asymmetric Power Autoregressive Conditional Heteroscedasticity

ARCH - Autoregressive Conditional Heteroscedasticity

ARMA - Autoregressive Moving Average

BIC - Bayesian Information Criteria

DCC - Dynamic Conditional Correlation

EGARCH - Exponential Generalized Autoregressive Conditional Heteroscedasticity

GARCH - Generalized Autoregressive Conditional Heteroscedasticity

GJR - Glosten-Jagannathan-Runkle

HN – Harvey-Newbold

LM – Lagrange Multiplier

LR - Likelihood Ratio

PACF - Partial Autocorrelation Function

Russell 2000 - RUT

S&P500 - Standard & Poor’s 500

TARCH - Threshold Autoregressive Conditional Heteroscedasticity

1. Introduction

The interest in market volatility and correlation between assets derives from its impact on capital investment decisions, as a deeper understanding of the major components of risk enables investors to make better decisions.

As ever, when undertaking those same decisions, individuals/institutions face the intertemporal risk-return trade-off – i.e. the potential return for a given investment is directly linked to the level of risk that the investor can bear. This trading principle depends on numerous factors, being Banz (1981) and Reinganum (1981) the first to suggest firm size as a determinant factor of this, and to understand the important role firm size plays in Financial Economics.

Small firms are presumed to be riskier than large firms, as in the overall the available information disclosed by small firms is minimal and insufficient. Furthermore, they cannot benefit from economies of scale and inevitably operate with higher production costs, their securities are less liquid than large firms securities, and yet, unlike larger firms which go through a continuous growing process they are in their early days, the time at usually most of the firms tend to fail (Ben-Zion & Shalit, 1975).

The high number of papers about the topic revealed that these distinctive features between different-sized firms materialize in distinct behaviours. Those differences became evident when comparing how firms with a different market capitalization perform in terms of risk-adjusted returns, how their stock volatility moves throughout time, and how the volatility transmission between those firms helps or not to infer their future dynamics.

For instance, when comparing risk-adjusted returns, several authors proved the existence of a size effect, with the small-sized firms performing higher risk-adjusted returns than the larger firms. Others showed that this size effect is no longer valid, and some even suggested the reverse, with stock returns of large firms outperforming stock returns of small-sized firms (Asness et al., 2018; Horowitz, Loughran & Savin, 2000; Dimson & Marsh, 1999).

In terms of volatility dynamics, multiple papers evince a negative relation between stock returns and changes in returns volatility. Some studies suggest that this observation impacts more strongly on small-sized firms than large firms, and others indicate otherwise. (Chelley-Steeley & Steeley, 1996; Dzieliński, Rieger, & Talpsepp, 2018).

For the volatility transmission between different-sized firms, authors once more came up with different results. Some suggested that the volatility of large stocks is relevant in predicting the future dynamics of smaller firms but not the inverse (Harris & Pisedtasalasai, 2006), whereas others identified a bidirectional relation (Hasan & Francis, 1998).

Therefore, and considering this extensive existing literature, we intend to investigate volatility dynamics across different-sized firms. This study extends previous literature on two subjects.

Firstly, and unlike previous literature that is mainly focused on such comparison, we extend that research and analyse now not only for large and small firms but also for dynamics across different industry sectors. For this purpose, the Global Industry Classification Standard (GICS) is considered. This constitutes a valuable investment mechanism to capture the comprehensiveness, extent, and evolution of sectors and industries. To the end of this study, we analyse the index sectors of Standard & Poor's 500 (S&P500) and Russel 2000 (RUT) traded during the 2006-2020 period. The first index comprises the performance of large market capitalization equities, and the latter concerns small-capitalization firms. The industry sectors evaluated are the Consumer Discretionary, Consumer Staples, Energy, Financials, Health, Industrials, Information Technology, Materials, and Utilities. Additionally, we will also investigate for spillover effects between the S&P500 and RUT Indexes. For this topic, we will implement a recent class of models which preserves the simple representation of univariate GARCHs with an easy to compute the correlation estimator, the Dynamic Conditional Correlation models.

Secondly, and for both subjects, we exhibit here results that cover the dramatic transformation experienced in financial markets during the Coronavirus Pandemic, therefore presenting relevant insights for those interested in improving their expertise on stock market volatility and in predicting future values of this.

To address the Asymmetric Volatility Phenomenon, we will model the return series with an ARMA structure. Given its homoskedasticity limitation, and to conclude about the return's conditional volatility behaviour, we will incorporate GARCH statistical processes. The concerned model is the Asymmetric Power GARCH (APARCH) of Ding, Granger, and Engle (1993), a powerful statistical process that includes seven specific cases as restrictions of parameters of the model itself. Among them, we have the GARCH of Bollerslev (1986) which considers that positive and negative shocks have the same impact on volatility and the GJR-GARCH that incorporates an asymmetric coefficient.

For the Spillover Effects topic, we will extend the univariate framework to a multivariate GARCH approach. We will test two specifications, the Dynamic Conditional Correlation (DCC) model from Engle and Sheppard (2001) and the Asymmetric Dynamic Conditional

Correlation (A-DCC) model from Cappiello, Engle, and Sheppard (2006), an expansion of the former that allows for conditional asymmetries in correlation dynamics.

For the several GARCH models under analysis, we consider the Student's *t* for the conditional distribution of the errors, thus following the proposal in Bollerslev (1987), when he inferred from previous papers that returns are frequently well described by a unimodal symmetric distribution with fatter tails than the Normal.

The achieved results will contribute to a deeper understanding in diverse fields of Finance such as portfolio management, where an improved comprehension of volatility spillovers will be valuable for portfolio managers, institutions, and investors throughout the stocks' selection procedure and the asset allocation process; research, for individuals interested in the study of equity markets; financial applications and risk management, as a better understanding of volatility dynamics will be crucial in option pricing and value-at-risk models that are built from volatility forecasts; among many others.

The remainder of the proposal is laid out as follows. Section 2 describes the existing literature of the two main topics addressed here, the asymmetric volatility phenomenon and spillover effects between large and small firms. Section 3 details the models to be used throughout this dissertation and describes the numerous statistical tests to be implemented. Section 4 reports the empirical results obtained from modelling index returns volatility and correlation, using the specifications described in Section 3.

2. Literature Review

Over time, intellectuals have been working hard to attest and contribute to the efforts devolved before theirs by proposing new methods, suggesting different views, and bringing new topics into discussion.

During these years, a substantial number of studies covering multiple fields such as profitability, diversification, technological innovation, and capital markets have revealed that incorporating a firm size variable can be decisive, leading authors to attain better results on repeated occasions. The first studies on the subject were developed by Banz (1981) and Reinganum (1981). They identified a size effect, with small firms presenting substantially higher risk-adjusted returns than large firms over a long period. Afterwards, many other pundits published papers indicating the existence of this size effect, raising several explanations for the occurrence of such phenomenon, namely Roll (1981), Reinganum (1982), Stoll and Whaley (1983), among others.

The interest in this phenomenon has persisted to this day, with numerous studies following such findings and contributing to the better comprehension we have today.

Nonetheless, authors have demonstrated that previous notions on the topic are not as straightforward as we believed, presenting in more recent times different results for the size effect phenomenon. Meeting prior observations, Asness et al. (2018) found evidence of an effective and stable size premium after controlling key variables like profitability, stability, growth, and safety. Likewise, Acharya and Pedersen (2005) detected that size matters since small companies face more liquidity risk, which requires higher expected returns. However, some indicated that the tendency in stock returns of small firms to outperform the returns of larger firms is unwarranted (Horowitz, Loughran, & Savin, 2000). Others, as Dimson and Marsh (1999) and Al-Rjoub et al. (2005), even presented results for a reversal in the size effect phenomenon, with larger firms displaying substantially higher risk-adjusted returns than smaller companies.

Although, despite this intensive and continuous work on the subject, there are still multiple nuances that few have explored yet, namely the relationship between firm size and returns' volatility, the stock volatility's behaviour across different industry sectors, and the transmission mechanisms between large and small firms.

2.1. Asymmetric Volatility Phenomenon on Different-Sized Firms

Previous studies from Black (1976) and Christie (1982) documented that, on average, distinct shocks impact differently on return volatility, as volatility tends to fall in response to "good news" and rise in response to "bad news". According to their empirical studies, this asymmetry can be explained, at least in part, through the financial leverage concept, named as leverage effect. When the firm's equity value decreases, the weight of debt in the firm capital structure increases, *ceteris paribus*. Therefore, leaving the equity holders in a riskier position, as senior securities rank above common stock at the distribution of the firm's income and in case of bankruptcy.

Besides the leverage effect, which is not strong enough to be fully accountable for an asymmetry in conditional variance, Campbell and Hentschel (1992) proposed a volatility feedback hypothesis. They argued that a large piece of news (good or bad) increases the future expected volatility, which in turn increases the required rate of stock's return and subsequently leads to a decline in stock price. The difference is that in the case of negative news, the volatility effect amplifies the impact of bad news.

Given these findings of asymmetric behaviour in the conditional variance, some authors stepped forward and related this to the firm size, unveiling the existence of a "firm size" effect. On the one hand, we have pundits that identified stronger evidence of asymmetric volatility in small capitalization companies than in larger ones (Chelley-Steeley & Steeley, 1996; Henry & Sharma, 1999), presenting a result consistent with investors' expectations that shocks in small firms are more uncertain. On the other hand, we have authors as Dzieliński, Rieger, and Talpsepp (2018), who proved that stocks under a higher level of attention (typically large firms) exhibit a larger asymmetry in volatility. Their research suggests that an asymmetry in volatility is related to the quantity of coverage. According to them, bad news about a small firm might be noticed by investors, but bad news about a large company are likely to generate far greater attention. This greater coverage will lead investors to receive more conflicting information, especially in bad times, which will further increase volatility in the case of large companies. Their empirical research is grounded in previous literature of Andrei and Hasler (2015), who found out that asymmetry in volatility is driven by asymmetric attention. According to them, in the case of inattentive investors or investors that receive less information, the new information is only gradually assimilated into stock prices, as the learning process is slow. In contrast, vigilant investors immediately absorb new information into prices.

The study of such dynamics does not end right then and there. Future papers raised the question of how those different-sized firms behave among distinct industry groups. This question is a consequence of increasing economic integration and ongoing corporate globalization, that gradually remove country and regional factors from the decision process (Cavaglia, Brightman & Aked, 2000). Ferreira and Gama (2005) identified that overall industry risk already dominates the world and country risk, thus adverting for the prominent role that industry factors play in investment strategy and the need to take them into account.

In order to arrive to previous results, the conditional variance of stock returns was broadly represented by the GARCH family of statistical processes. Engle (1982) was the first to introduce this new class of stochastic processes, the Autoregressive Conditional Heteroscedastic (ARCH) processes. The original ARCH model was developed to improve the existing models, replacing the previous assumption of constant volatility for conditional volatility. Since then, its popularity has significantly increased, and numerous specifications have been introduced. The GARCH processes demonstrate a good capacity to estimate appropriate volatility forecasts, meet the main stylized facts for asset returns, and yet, the estimation of the model parameters is quite efficient.

Within this family, the Generalized ARCH of Bollerslev (1986), more specifically the GARCH (1,1), has become the most popular ARCH specification. The GARCH model is a generalization of the original ARCH in Engle (1982) that allows for past conditional variances in the current conditional variance equation. Taylor (2011) claims that the popularity of GARCH (1,1) can be easily explained. Firstly, this specification is easy to estimate as it includes only four estimates. Furthermore, it accounts for the main stylized facts of daily returns. At last, it produces volatility forecasts with similar accuracy to the ones generated from more complex specifications.

However, and as evidenced in Hansen and Lunde (2005), the GARCH (1,1) has revealed not being able to capture the asymmetric behaviour of returns volatility in financial markets, with this specification being outperformed by others that can accommodate it.

Empirical studies by Nelson (1991) and Glosten, Jagannathan, and Runkle (1993) showed that it is crucial to consider an asymmetric coefficient in financial time series models. In order to account for it, an immensity of univariate GARCH specifications have been proposed: the Exponential GARCH (EGARCH) model of Nelson (1991), the GJR-GARCH of Glosten, Jagannathan, and Runkle (1993), the Threshold ARCH (TARCH) model of Zakoian (1994), the Asymmetric Power Autoregressive Conditional Heteroscedasticity (APARCH) of Ding, Granger, and Engle (1993), among many others.

The APARCH model is a unique specification that will assume a primary role in this dissertation. This specification encompasses seven other special cases (among them the GARCH, GJR, and TARCH) and allows the power of conditional variance equation to be estimated from the data. It was proposed in Ding, Granger, and Engle (1993), while investigating the long memory property of returns - i.e., when returns distant in time display high autocorrelation between them. In line with previous results, they identified a lower serial correlation between stock market returns than between absolute returns. They found out that power transformations of absolute returns ($|r_t|^d$) have substantial high correlations up to 100 lags, proving that stock returns have long-term memory. According to their study, the largest autocorrelation is observed for values of d close to 1. From this evidence, they tested if other models could also generate a similar pattern to the autocorrelation in stock market returns, and it is rather interesting that they do.

2.2. Spillover Effects between Large and Small Firms

In addition to those findings, events of price and volatility transmission emerged. Over time, numerous studies have investigated the spillover effects between developed and emerging stock markets, among major asset classes, between different sectors and industries, and within some regions (Li & Giles, 2015; Liow, 2015; Elyasiani et al., 2015). These studies found evidence of spillover effects from the US market to emerging stock markets, pointed out equity as the main contributor of spillover effects among major asset classes, and showed strong linkages between different industries. However, the study of spillover effects between different-sized firms has not received that much attention.

Lo and MacKinlay (1990) and Boudoukh, Richardson and Whitelaw (1994) were among the first to conduct studies on this topic, revealing the existence of cross effects among different-sized firms. Interestingly, a number of those studies demonstrated that such cross-correlations are asymmetric, with the returns of large firms being used to explain the return of small firms, but not the inverse (Mech, 1993).

Subsequently, Conrad, Gultekin, and Kaul (1991) identified the existence of this same asymmetry in the volatility's predictability. They found that shocks to large firms are important in predicting the future volatility of small firms, whereas shocks to small firms have no impact on the future large firms' conditional variance. According to them, this asymmetry in volatility spillovers is consistent with a market in which the prices of large stocks react to new information immediately, and the prices of small stocks respond with a lag.

More recently, Harris and Pisedtasalasai (2006) reached the same conclusion. They wondered how much transmission mechanisms tell us about market efficiency and argued that this information needs to be considered in financial applications determined from volatility estimates such as portfolio optimization, pricing, value-at-risk models, and hedging strategies.

Nevertheless, some authors did not attain the same results. Hasan and Francis (1998) found evidence of volatility spillovers between different-sized firms, but in contrast to previous literature, their results showed a symmetry in the predictability of volatilities. According to the authors, their analysis from monthly collected data captures more of the dynamics among firms than previous studies grounded on weekly data.

To investigate these transmissions of volatility, Engle and Sheppard (2001) introduced the Dynamic Conditional Correlation (DCC) model. The DCC is a multivariate specification with conditional variance and correlation where the conditional covariance is decomposed into conditional standard deviations and a time-varying correlation matrix, so that univariate and multivariate dynamics are dissociated. They were pioneers in developing a class of multivariate GARCH models capable of estimating large time-varying covariance matrices, as until recently practitioners used to assume a constant correlation in their studies. The DCC model meets the conclusions of most empirical tests carried out to verify such assumption, with pundits demonstrating that correlation increases in periods of high volatility and that autocorrelation's magnitude and persistence are affected by volatility (Billio, Caporin, & Gobbo, 2006). In essence, this parametrization congregates the simple interpretation and empirical success of univariate GARCH models, with ease to compute and interpret the dynamic correlation estimator.

In the meantime, multiple studies identified an asymmetric pattern in the cross-correlation between large and small firms. According to Yu and Wu (2001), this asymmetry may be explained through a group of economic factors such as market frictions, lagged information transmission, and institutional interest. Nevertheless, they found additional evidence suggesting that a large part of this asymmetric pattern may in fact be attributed to the differential quality of information between large and small companies. Large companies display a higher sensitivity of stock prices to market information and an increased quality of cash flows information.

To account for this issue, Cappiello, Engle, and Sheppard (2006) proposed an alternative multivariate GARCH model, the Asymmetric Generalized Dynamic Conditional Correlation

(AG-DCC). Unlike the original model, this specification permits now for a conditional asymmetry in volatility and correlation. According to its authors, this is an extension of the standard DCC-GARCH along two dimensions: to allow for asset-specific news and smoothing parameters, and conditional asymmetries in correlations. In their literature, authors show the importance of this new model by displaying empirical results that unveil an asymmetric dynamic in equity returns' conditional correlation, with this model performing better than all symmetric specifications.

Further in this dissertation, we will describe and implement the statistical models above-mentioned, reaching a better comprehension of volatility dynamics and transmission mechanisms between large and small firms.

3. Methodology

In this dissertation, we address two main topics: the asymmetric volatility phenomenon on distinct capitalization firms and the existence of spillover effects between the S&P500 and RUT Indexes.

3.1 Asymmetric Volatility Phenomenon

An asymmetric dynamic in returns' volatility will be tested through two complementary approaches: an in-sample and an out-of-sample methodology. Whereas the in-sample fit of a model is appreciated by practitioners, as it works well in numerous situations and helps them to acquire theoretical insights on the topics of interest, the latter allows them to evaluate the out-of-sample predictive power of the models.

In the out-of-sample methodology, the data is split into a training set for initial parameter estimation and model selection, and a validation set to evaluate the forecasts of one-day-ahead volatility.

In this paper, the first nine years will be considered as the training set (roughly 2500 observations), while the last four years will correspond to the validation set (approximately 1000 observations). Hence, we follow the insights of Hansen and Timmermann (2012), who recommend practitioners to adopt a broad out-of-sample period to strengthen the power of the forecast evaluation test.

To compare the models proposed, we will implement the Likelihood Ratio Test for the in-sample results and the Harvey-Newbold (HN) Test for the out-of-sample analysis. Before their implementation, there are still multiple steps to go through.

3.1.1 Normality and Stationarity in Index Returns

The importance of fitting the data under scrutiny with the right curve is quite significant, as it leads us to achieve more realistic results. To find the distribution that closer describes the data, we will test the normality of all observations.

The Jarque-Bera test is valid for large samples, testing the index returns normality through a simple computation and an asymptotically efficient procedure (Jarque & Bera, 1987).

The test statistic is asymptotically distributed as $\chi^2_{(2)}$, and defined from the estimates of Skewness ($\widehat{\sqrt{b_1}}$) and Kurtosis ($\widehat{b_2}$):

$$JB = n * \left(\frac{(\widehat{\sqrt{b_1}})^2}{6} + \frac{\left(\widehat{b_2} - 3\right)^2}{24} \right) \tag{1}$$

where $\widehat{\sqrt{b_1}} = \widehat{\mu_3}/\widehat{\mu_2^{3/2}}$ and $\widehat{b_2} = \widehat{\mu_4}/\widehat{\mu_2^2}$.

For a α significance level, the null hypothesis (normality in index returns) is rejected if JB > $\chi^2_{(2)}(\alpha)$.

Additionally, we test the time series stationarity. To that end, the Augmented version of the Dickey-Fuller test will be considered. Whereas the original version is built on a simple autoregression that can cause the errors to be autocorrelated, the Augmented Dickey-Fuller (ADF) test allows us to clean up any serial correlation by including lagged terms. In this analysis, the number of lagged terms to incorporate in the ADF statistical test are determined by minimizing a given Information Criteria.

To attain accurate results, we must choose the most appropriate specification from the class of models investigated by Dickey and Fuller (1979):

$$\Delta y_t = \gamma y_{t-1} + \sum_{j=2}^{p} \beta_j \Delta y_{t-j+1} + \varepsilon_t \tag{2}$$

$$\Delta y_t = \beta_0 + \beta_1 \mathrm{t} + \gamma y_{t-1} + \sum_{j=2}^{p} \beta_j \Delta y_{t-j+1} + \varepsilon_t \tag{3}$$

$$\Delta y_t = \beta_0 + \beta_1 \mathrm{t} + \gamma y_{t-1} + \sum_{j=2}^{p} \beta_j \Delta y_{t-j+1} + \varepsilon_t \tag{4}$$

The first model is not a plausible case to describe the stock returns since it is hard to believe that the analysed data was generated by a process where β_0 = 0. Moreover, there is no evidence of stock returns exhibiting a deterministic trend, thus model (3) is also not an option. Therefore, by a process of elimination, the ADF test will be implemented including an intercept in linear regression.

The ADF test assumes the following hypothesis:

$$\begin{cases} H_0: \gamma = 0 \\ H_1: \gamma < 0 \end{cases}$$

Under the null hypothesis, we test if the series contains a unit root – i.e., if the time series is non-stationarity. In the case of rejecting the null hypothesis, there is statistical evidence of stationarity in the time series.

3.1.2 Conditional Mean

To reach a greater understanding of the time series, we will model the daily returns through an Autoregressive Moving Average (ARMA) process. This statistical process presents a good specification for the conditional mean equation, based on past realizations. The ARMA(p,q) model incorporates p autoregressive terms and q lagged moving-average terms, as defined below:

$$y_t = \mu + \varphi_1 y_{t-1} + \ldots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \tag{5}$$

where y_{t-i} is a return at time $t - i$ and ε_t is a white noise process.

The coefficients $\varphi_1,\ldots, \varphi_p$ reveal how y_t is related to its past values $y_{t-1},\ldots, y_{t-p}$, and the coefficients $\theta_1,\ldots, \theta_q$ reveal how y_t is related to past random shocks.

In order to find the accurate conditional mean equation of each time series, two autocorrelation functions are considered: Total Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).

While Pure Autoregressive models are characterised by an ACF that decays towards zero and a PACF with p statistically significant spikes, the Pure Moving Average processes is recognized by an ACF with q statistically significant lags and a PACF that tails off towards zero. In an ARMA(p,q) model, the order of each term is extracted from the PACF and ACF, with the former decaying after lag p and the latter after lag q (Enders, 2015).

Lastly, after estimating the respective models, we will confirm whether those were able to capture the correlation embodied in each time series, thus testing if the residuals present no autocorrelation. To this end, we implement the Portmanteau Test suggested in Ljung and Box (1978), the Ljung-Box test:

$$\tilde{Q} = n(n+2)\sum_{k=1}^{m}\frac{\rho_k^2}{n-k} \tag{6}$$

where n is the number of observations, m is the number of lags in the null hypothesis, and ρ_k is the autocorrelation at lag k.

For a α significance level, we reject the null hypothesis if $\tilde{Q} > \chi^2_{1-\alpha,m}$, where:

$$\begin{cases} \mathrm{H_0}: \rho_1 = \ldots = \rho_m = 0 \\ \mathrm{H_1}: \exists \rho_j \neq 0 \end{cases}$$

If the null hypothesis is rejected, there is statistical evidence of serial correlation. Otherwise, the time series was generated by a non-autocorrelated process.

3.1.3 Conditional Heteroscedasticity

Considering that homoscedasticity is implied in the conditional distribution of ARMA processes, we test for the existence of a very common feature in financial returns, the existence of conditional heteroscedasticity.

For this purpose, we apply the Lagrange Multiplier (LM) test proposed in Engle (1982), which effectively allows us to test for the existence of autoregressive conditional heteroscedastic (ARCH) effects.

Considering that the squared residual is regressed on a constant and q lagged values:

$$e_t^2 = \hat{\alpha}_0 + \sum_{i=1}^{q}\hat{\alpha}_i e_{t-i}^2 + u_t \tag{7}$$

This test contemplates the following hypothesis:

$$\begin{cases} \mathrm{H_0}: \alpha_0 = \alpha_1 = \ldots = \alpha_m = 0 \\ \mathrm{H_1}: \exists\ \alpha_j \neq 0, \text{for } j = 1, 2, \ldots, m \end{cases}$$

The null hypothesis states the absence of an ARCH effect. Hence, if this is rejected, there is statistical evidence of conditional heteroscedasticity in residuals of the ARMA process.

3.1.4 Univariate Conditional Heteroscedasticity Models

To characterize the time-varying volatility, a common feature in economic and financial data, the ARCH Model and an extension of this, the GARCH Model, were introduced. Despite describing some significant features such as volatility clustering and heavy tails, there are others that neither of them can accommodate. Among them, the asymmetric dynamic in returns volatility.

Since these two were built, and in an attempt to account for additional features, numerous GARCH models and variants of these have been proposed. For the purpose of this study, we consider three of those specifications, confronting symmetric and asymmetric dynamics in volatility.

In here, all specifications are estimated through the Maximum Likelihood Method. The lag length of the models will be restricted to p = q = 1, following the research of Hansen and Lunde (2005). They were able to empirically demonstrate that models with more lags rarely outperform the (1,1) specification.

3.1.4.1 Asymmetric Power Autoregressive Conditional Heteroskedasticity Model (APARCH)

The model considered is the Asymmetric Power GARCH (APARCH) of Ding, Granger, and Engle (1993). This model encompasses seven other special cases as restrictions of parameters of the model itself. Unlike GJR and GARCH specifications, it allows the power (δ) of the heteroscedasticity equation to be estimated from the data under review. When considering this transformation, individuals can linearize otherwise non-linear models.

Then, when considering a linear regression:

$$y_t = \beta x_t + u_t \tag{8}$$

$$u_t = \varepsilon_t \sigma_t \tag{9}$$

where y_t is the dependent variable and x_t a vector with explanatory variables.

The APARCH structure is as follows:

$$\sigma^{\delta}{}_{t} = \omega + \sum_{i=1}^{p} \alpha_i (|u_{t-i}| - \gamma_i u_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j \sigma^{\delta}{}_{t-i} \qquad (10)$$

where, $\omega > 0, \delta > 0,\ \alpha_i \gg 0$ for $i = 1, \ldots, p, \beta_j \gg 0$ for $j = 1, \ldots, q,$
$-1 < \gamma_i < 1$ for $i = 1, \ldots, p$

The APARCH model encompasses the following special cases:

1) The Engle's ARCH(p) if $\delta = 2, \gamma_i = 0$ for $i = 1, \ldots, p$, and $\beta_j = 0$ for $j = 1, \ldots, q$.
2) The GARCH model of Bollerslev, when restricting $\delta = 2$ and $\gamma_i = 0$, for $i = 1, \ldots, p$.
3) The Taylor/Schwert's GARCH, if consider $\delta = 1$ and $\gamma_i = 0$, for $i = 1, \ldots, p$.
4) The GJR GARCH model of Glosten, Jagannathan, and Runkle, when $\delta = 2$.
5) The TARCH model of Zakoian, by restricting $\delta = 1$ and $\beta_j = 0$, for $i = 1, \ldots, q$.
6) The Higgins and Bera's NARCH model, if $\gamma_i = 0$ for $i = 1, \ldots, p$, and $\beta_j = 0$ for $j = 1, \ldots, q$.
7) The Log-ARCH model of Geweke and Pantula, when $\delta \rightarrow 0$.

3.1.4.2 Generalized Autoregressive Conditional Heteroskedasticity Model (GARCH)

To analyse for a symmetric dynamic in returns volatility, we test for the most applied ARCH specification in empirical research, the GARCH model. Bollerslev (1986) proposed a more comprehensive specification than the original ARCH model, where the conditional variance depends on the *q* most recent squared residuals and *p* latest conditional variances.

The GARCH model is one of the seven special cases encompassed in the APARCH model, and defined as:

$$\sigma^{2}{}_{t} = \omega + \sum_{i=1}^{p} \alpha_i u_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 \qquad (11)$$

where, $\omega > 0, \alpha_i \geq 0$ for $i = 1, \ldots, q$ and $\beta_j \geq 0$ for $j = 1, \ldots, p$

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$$

These constraints guarantee that the conditional variance is not negative, and the process is stationary.

3.1.4.3 GJR-GARCH Model

Conversely, to analyse asymmetric volatility movements (when positive and negative shocks have a different impact on the returns' volatility), we consider the APARCH and GJR-GARCH models. The latter is also a special case of the APARCH model when we restrict δ to 2, and:

$$\alpha_i^* = \alpha_i(1-\gamma_i)^2 \tag{12}$$

$$\gamma_i^* = 4\alpha_i\gamma_i \tag{13}$$

Glosten, Jagannathan, and Runkle (1993) demonstrated that the standard GARCH is misspecified. In view of improving this original model, they developed the GJR-GARCH that incorporates an asymmetric coefficient in the conditional variance (γ_i^*). The model is defined as follows:

$$\sigma^2{}_t = \omega + \sum_{i=1}^{p} \alpha_i^* u_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-i}^2 + \sum_{i=1}^{p} \gamma_i^* u_{t-i}^2 I_{t-i} \tag{14}$$

Asymmetry is incorporated in the GJR-GARCH model by weighting u_{t-i}^2 differently for positive and negative residuals, through the indicator variable:

$$I_t = \begin{cases} 1, & \text{if } u_t < 0 \\ 0, & \text{otherwise} \end{cases}$$

3.1.5 Likelihood Ratio Test

To determine which model better fits the data for in-sample evidence, we measure their goodness of fit by confronting the two special cases against the unrestricted model, the APARCH.

To this end, the Likelihood Ratio (LR) test is implemented. It compares the maximised value of the Log-Likelihood function between two nested models: the unrestricted model (L_u) and the restricted case (L_r).

The LR test statistic is asymptotically Chi-squared distributed, where m equals the number of restrictions imposed for each special case. The LR test is defined as:

$$\text{LR} = -2(L_r - L_u) \sim \chi^2_{(m)} \tag{15}$$

In the null hypothesis, the LR test states that the special case is a better fit for the data than the unrestricted model. It is rejected for a test statistic higher than a Chi-squared percentile with m degrees of freedom, a percentile that will vary according to the α significance level considered (Brooks, 2019).

3.1.6 Harvey-Newbold Test

To assess the out-of-sample predictability, we implement the HN test. This constitutes a test for Equal Predictive Ability, thus following the indication in Hansen (2005).

Harvey and Newbold (2000) proposed a test for multiple forecast encompassing. The HN test compares competing forecasts by testing whether one forecast encompasses the others - i.e. the inferior forecasts do not contain any useful information not already present in the superior forecast.

To test the null hypothesis that one forecast encompasses the others, we consider $(f_{1t}, \dots, f_{Kt})$ as K competing forecasts and A_t as the Actual Quantity, with:

$$e_{1t} = \lambda_1(e_{1t} - e_{2t}) + \lambda_2(e_{1t} - e_{3t}) + \dots + \lambda_{K-1}(e_{1t} - e_{Kt}) + \varepsilon_t \tag{16}$$

where $e_{it} = A_t - f_{it}$, ε_t is the error of the combined forecast, and $0 \leq \lambda_i \leq 1$.

The null hypothesis that f_1 encompasses $f_2, \dots, f_K$ is:

$$\text{H}_0\text{: } \lambda_1 = \lambda_2 = \cdots = \lambda_{K-1} = 0$$

The regression-based test for multiple forecast encompassing is an F-test of the joint significant parameters in e_{it}. The regression (16) can be written in general form as:

$$y_t = X_t'\beta + \varepsilon_t \tag{17}$$

where $y_t = e_{1t}, \beta = [\lambda_1\ \lambda_2 \ldots \lambda_{K-1}]', and\ X_t = [(e_{1t} - e_{2t})(e_{1t} - e_{3t})\ldots(e_{1t} - e_{Kt})]'$

Nonetheless, Harvey and Newbold (2000) strongly recommend the implementation of the Modified Diebold-Mariano-Type Test (MS^*) in practical applications, given its good size and reasonable power in large samples. They modified the original version to test for forecast encompassing, with the test statistic taking the form:

$$MS^* = (K-1)^{-1}(n-1)^{-1}(n-K+1)\bar{d}'\hat{V}^{-1}\bar{d} \tag{18}$$

where $\bar{d} = [\bar{d}_1\ \bar{d}_2 \ldots \bar{d}_{K-1}]'$, $\bar{d}_i = n^{-1}\sum d_{it}$, $d_{it} = e_{1t}(e_{1t} - e_{i+1,t})$ and $\hat{V}$ is the sample covariance matrix.

For the test statistic (18), we consider $F_{K-1,n-K+1}$ critical values.

3.2 Spillover Effects

Over time, volatility moves together across assets and markets. At sometimes closer, other times farthest, but the recognition of such characteristic through a Multivariate model is an opportunity to extend our knowledge and enhance the decision process.

In this paper, index returns are first modelled into a VAR framework. Furthermore, to address the time-varying correlations we expand the univariate analysis to a multivariate approach. To this end, we consider two Multivariate GARCH models of conditional variances and correlations, thus modelling conditional variances and correlations separately.

3.2.1 Vector Autoregressive Models (VAR)

Among the existent multivariate stochastic processes and given the VAR's resemblance with ordinary regression models, it turns to be the most applied model in practice. The VAR(p) model captures the linear interdependencies among multiple time series and is described by:

$$y_t = v + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t \tag{19}$$

where $y_t = (y_{1t}, \ldots, y_{Kt})'$ is a k-vector of stationary variables; $v, A_1, \ldots, A_p$ are matrices of coefficients to be estimated; and ε_t is a vector of innovations.

To implement a VAR framework, the time series must be stationary. To that end, the analysis is conducted using daily returns instead of index prices, thereby only capturing short-run dependencies between each time series.

The diagnostic testing will be implemented again, this time at a multivariate level. To test for normality in index returns, we implement the multivariate Jarque-Bera test. Moreover, to conclude about the model's stability, we implement the CUSUM test based on recursive residuals (Brown, Durbin & Evans, 1975). Through the multivariate Portmanteau (asymptotic) test, we can assess for serially correlated errors, and finally, we run the multivariate ARCH-LM test to study for the presence of conditional heteroscedasticity.

This model will be used to investigate and describe the relationship between the two return time series under analysis. For that purpose, we implement the Pairwise Granger Causality test.

3.2.2 Dynamic Conditional Correlation (DCC) Model

To determine the DCC GARCH parameters, we follow a two-stage procedure. Firstly, we estimate the univariate GARCH models for each index individually, and subsequently, using the transformed residuals of this first stage, we estimate the parameters of the dynamic correlation.

The time varying correlation matrix (R_t) is the following:

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{20}$$

with,

$$Q_t = (1 - a - b)\bar{Q} + a\varepsilon_{t-1}\varepsilon_{t-1}^T + bQ_{t-1} \tag{21}$$

where, Q_t^* is a diagonal matrix with the square root of the diagonal elements in Q_t.

The parameters a and b are conditioned to guarantee non-negativity and stationarity.

ε_t are the disturbances standardized by their conditional standard deviation.

$\bar{Q}$ equals $E[\varepsilon_t\varepsilon_t^T]$. It is the unconditional covariance matrix of the ε_t.

Additionally, to ensure that R_t is positive definite, we must guarantee that Q_t is positive definite.

The determined model is used to estimate the conditional covariance. The covariance matrix (H_t) is decomposed into a time-varying conditional standard deviation matrix (D_t) from the univariate GARCH models, and a time varying correlation matrix, R_t:

$$H_t \equiv D_t R_t D_t \tag{22}$$

$$h_{i,t} = \alpha_{i,0} + \sum_{q=1}^{Q_i} \alpha_{i,q}\, r_{i,t-q}^2 + \sum_{p=1}^{P_i} \delta_{i,p}\, h_{i,t-p} \tag{23}$$

where $h_{i,t}$ are the elements in the diagonal matrix D_t for *i = 1,2, ..., n,* and in accordance with the non-negativity and stationarity restrictions of the GARCH specification.

To guarantee that H_t is positive definite, the starting value of Q_t (Q_0) has to be positive definite. Additionally, two sufficient conditions for Q_t to be positive definite and stationary is to attain non-negative values of a and b, and that $a^2 + b^2 < 1$.

From the covariance matrix, we infer the predictability of volatility between distinct market capitalization firms.

3.2.3 Asymmetric Dynamic Conditional Correlation (A-DCC) Model

The A-DCC is a special case of the AG-DCC model, where original parameter matrices are replaced by scalars. To estimate such specification, we will adopt a two-stage procedure identical to the one presented in section 3.2.2, for the DCC-GARCH model. Hence, after estimating the several univariate GARCH models, the standardized residuals are used to estimate the correlation parameters. The difference between the two specifications is in the correlation evolution equation implemented in the second stage of the A-DCC's estimation, considering now:

$$Q_t = (\bar{Q} - a^2\bar{Q} - b^2\bar{Q} - g^2\bar{N}) + a^2\varepsilon_{t-1}\varepsilon_{t-1}^T + g^2 n_{t-1} n_{t-1}^T + b^2 Q_{t-1} \tag{24}$$

where,

$$n_t = I[\varepsilon_t < 0] \circ \varepsilon_t$$
$$\bar{N} = E[n_t n_t^T]$$

A sufficient condition for Q_t to be positive definite is that $(\bar{Q} - a^2\bar{Q} - b^2\bar{Q} - g^2\bar{N})$ is positive semi-definite, with $a^2 + b^2 + \delta g^2 < 1$ for $\delta =$ maximum eigenvalue $[\bar{Q}^{-1/2}\bar{N}\bar{Q}^{-1/2}]$.

This modification of the original DCC model was proposed in Cappiello, Engle, and Sheppard (2006), incorporating now two additional dimensions: asset-specific news and smoothing parameters, and a conditional asymmetry in the correlation dynamics.

3.3 Information Criteria

Along the several stages of this study, numerous decisions are based on a given Information Criteria. Hence, the choice of which Criteria to consider is of major importance.

While selecting the number of lag terms to incorporate in the structure identification and in the multiple statistical tests, we follow the results in Liew (2004). He concluded that the accuracy of each Information Criteria improves as the sample size grows, with the Akaike Information Criteria (AIC) outperforming the Bayesian Information Criteria (BIC) for a sample size lower than 480 and the other way around for a sample size higher than 480. Thus, as we investigate time series with a sample size substantially higher than 480, we will rely on BIC results.

4. Empirical Study

Our analysis uses price data from all indexes for a period of fourteen years, from 2006 until 2020. Hence, we analyse the models over a representative time window, including in such range the recent period of the Coronavirus pandemic and the 2007-2008 Global Financial Crisis, the two periods associated to the highest market volatility values ever registered in the CBOE Volatility Index (VIX).

The data series are collected from a Bloomberg terminal and expressed in US dollars (USD).

The daily adjusted closing prices are converted into daily log returns:

$$\mathrm{r}_{i,t} = \ln\left(\frac{\mathrm{P}_{i,t}}{\mathrm{P}_{i,t-1}}\right) = \ln\left(\mathrm{P}_{i,t}\right) - \ln\left(\mathrm{P}_{i,t-1}\right) \tag{25}$$

where $\mathrm{P}_{i,t}$ is the closing price of index *i* at moment *t*, and *t* = 1, 2, …, T.

4.1 Graphical Examination of Stock Prices and Index Returns

The best way to start any time series analysis is to plot it against time and evaluate it graphically. Thus, the daily adjusted closing prices of the S&P500 and RUT Indexes, and the transformed series of daily log returns are presented in Figures 1, 2 and 3, respectively.

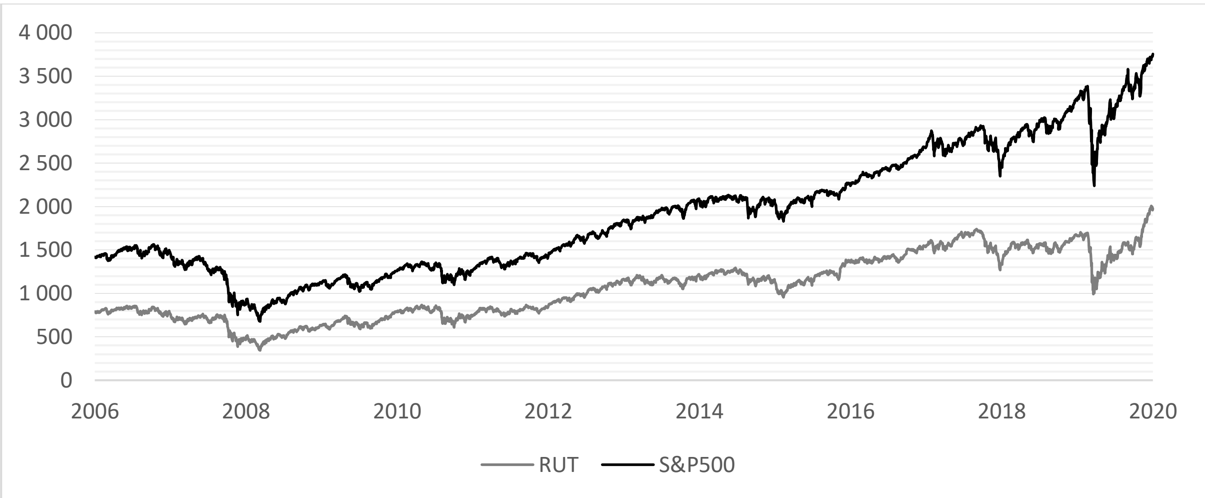

Figure 1 – Daily Stock Prices from 2006 - 2020

During the last fourteen years, the two indexes exhibited similar stock price movements. Between the end of 2006 and the end of 2008 both present a downward trend, signalling that regardless of the firms' capitalization value none could escape the 2007-2008 Global Financial Crisis. From then on, and in line with the subsequent recover in economic activity, they display a marked growing trend, finishing 2020 in historical highs. Nevertheless, in the years in between, both suffer significant drops in index levels, with emphasis on the recent Coronavirus Pandemic Crisis. Even so, those sharp declines have been followed by a quick and sustained rebound, a V-Shaped Recovery, thus not affecting the overall trend.

A stochastic trend like the one in Figure 1 is a common source of nonstationary, keeping us from employing econometric models and investigate for relationships between different time series. In order to attain stationary time series, where historical relationships can be generalized into the future, we converted closing prices into log returns.

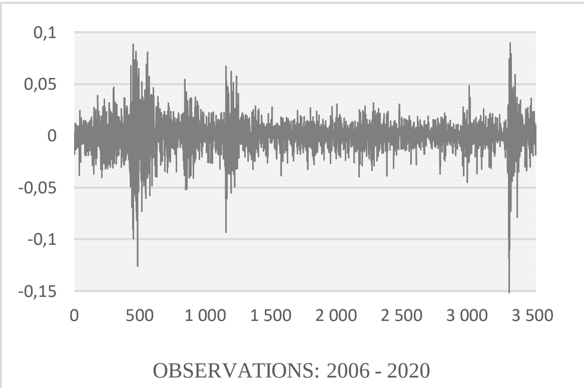

Figure 2 – Russel 2000 Index Returns

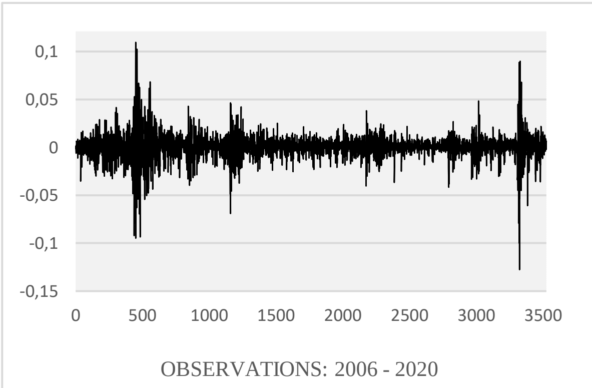

Figure 3 – S&P500 Index Returns

Figures 2 and 3 exhibit the S&P500 and RUT index returns. The previous transformation fully eliminated the stochastic trend, with observations converging now to a constant value. Contrarily to stock prices representation, shocks now seem to be temporary, with their effects dissipating over time and time series reverting to a long-term constant level.

Additionally, it is relevant to note that the daily returns in Figures 2 and 3 exhibit volatility clustering, with records of high volatility during and after financial crisis, and low volatility in periods of growth and stability. Those observations are thus in line with the existing literature in Mandelbrot (1963), where the author argues that large changes tend to be followed by large changes, and small changes tend to be followed by small changes.

Nonetheless, we can find distinctive characteristics between the two indexes. In fact, for the period considered, the RUT index seems to present a higher dispersion and appears to vary more than the S&P500. According to the figures above, we can also realise that small firms exhibit more extreme returns than large firms.

To describe volatility clustering and other related effects, we will then employ three GARCH specifications.

4.2 Asymmetric Volatility Phenomenon Analysis

4.2.1 Properties of Returns and Preliminary Tests

Table 1 reports preliminary statistics on key features of S&P500 and RUT index returns, whereas sector indexes' results are exhibited in Annex A.

The mean returns are all close to zero and predominantly positive, exhibiting a positive trend in price movements over time. The Energy and Financial Sectors present the lowest mean returns, with the former for small firms and the latter for large firms being the exception here, exhibiting a negative mean return. On the other hand, the Technological Sector presents the highest mean returns in both cases. On average, the returns of small firms exceed the returns of large firms, therefore meeting previous expectations of a higher return on small firms to compensate the investors for a higher risk.

Moreover, we addressed the extreme results of index returns rather than focusing only on the average. Table 1 shows that all indexes present negative skewness, therefore demonstrating a higher frequency of incurring small gains and a few extreme losses. A special reference for the Energy and Materials sectors, and large-cap firms of the Industrials sector, since all these present more negative extreme returns on the left tail than the remaining sectors. Regarding kurtosis, results meet previous expectations of finding leptokurtosis in financial data, with large capitalization firms exhibiting more extreme results than small firms.

Therefore, and it comes as no surprise, the Jarque-Bera Test rejects the normality of the returns for every single index in analysis. Due to the characteristics of the empirical distribution, namely the leptokurtosis, we will perform the Student's T distribution to model the conditional distribution of the errors.

Additionally, to guarantee that we can proceed with conditional mean estimation, we tested index returns' stationarity through the ADF test. For that purpose, we considered the "urca" package from the RStudio programming language. Empirical results support previous graphical evidence by rejecting the null hypothesis for a significance level of 1%, thus indicating clear evidence of stationarity in each time series. Therefore, converting the daily closing prices into daily log returns allowed us to remove the stochastic trend observed in Figure 1 and find a stationary series that we can work with.

Afterwards, in order to estimate the autoregressive and moving-average terms that better fit the data, we implemented the "auto.arima" function from the "forecast" package, as it automatically returns the best ARMA model for a given Information Criteria. After incorporating additional terms in the suggested processes, we were able to remove autocorrelation from most indexes. The results of the Ljung-Box test, for up to tenth order serial correlation in squared returns, are presented in Table 1. As the p-value of each return series is higher than a 5% significance level, we do not reject the null hypothesis and consequently verify the adequacy of the fit. Nevertheless, there is an exception. In the Financial Sector for large-cap firms, we were unable to capture all information embodied in the time series. We end up rejecting the null hypothesis, even after including additional terms in the conditional mean equation. For this case, we decided to proceed with the structure proposed from "auto.arima" function, being this the best model according to BIC, no matter the result of the test.

Lastly, as financial returns present significant evidence of volatility clustering, we investigated the squared residual series from each ARMA specification. Figure 4 plots the ACF for the squared residual series of the RUT and SP500 indexes. In both, we can visually confirm that there are still dependencies left in the data. To accommodate autocorrelation in squared returns, we will model indexes' volatility, thereby allowing for conditional heteroscedasticity.

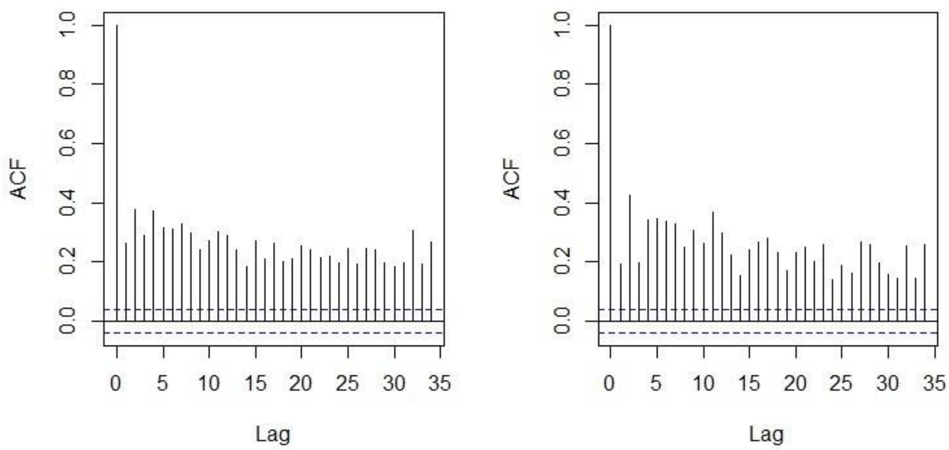

Figure 2 – ACF of RUT Squared Residuals (on the left), and ACF of SP500 Squared Residuals (on the right)

Additionally, to test for the presence of conditional heteroscedasticity we implemented the ARCH LM test. The test statistic results are straightforward and conclusive, they signal the presence of time-varying volatility as the null hypothesis is rejected for a 1% significance level and for any given order. Such dynamic will be captured by incorporating a GARCH structure in the processes, therefore, an ARMA-GARCH model must be estimated (Curto & Pinto, 2012).

Table 1 – Summary and Diagnostic Statistics of Indexes Returns

	Russell 2000 Index	**S&P500 Index**
Mean	0,00022	0,00018
Median	0,00101	0,00062
Maximum	0,08860	0,10957
Minimum	-0,12612	-0,09470
Std. Dev.	0,01670	0,01320
Skewness	-0,356	-0,326
Kurtosis	8,322	12,916
Jarque-Bera	3024*	10361*
ADF (intercept)	-37,828*	-39,822*
Ljung-Box	8,454	8,810
ARCH-LM	296*	301*

* Denote significance at 1% level.

4.2.2 In-Sample Estimation and Empirical Results

Table 2 provides in-sample results for the RUT and S&P500 Indexes, whereas all sector indexes results are reported in Annex B. The coefficients from conditional mean equations and the GARCH coefficients from conditional variance models are also reported in Table 2. To find those estimates, we employed the "rugarch" package from the R programming language.

Overall, the conditional variance estimates are statistically significant. This is an evidence regardless of the GARCH specification considered, thereby matching the LM test results and supporting our decision to deal with conditional heteroscedasticity in index returns.

Large values of α_1 and β_1 increase the conditional variance, but they do so in different forms. A larger α_1 means a higher response of conditional volatility to new information, since a shock (ε_t) has a significant impact on u_t^2 and ${\sigma^2}_{t+1}$. On the other hand, a larger β_1 means a stronger autoregressive persistence in conditional variance.

In this dissertation, α_1 and β_1 coefficients satisfy the properties of nonnegativity and stationarity in the processes, being always equal or higher than 0 and their sum a value below 1, respectively. Nonetheless, that sum is always close to 1, which signals a significant amount of volatility persistence with past volatility strongly affecting current volatility.

In the GARCH model, α_1 and β_1 estimates are statistically significant. They exhibit different dynamics across different capitalization firms and sector indexes. Large firms tend to present greater values of α_1 and lower values of β_1, thus suggesting that a market event has a greater immediate impact on large-cap firms than small firms. Since the volatility of prices is related to the rate of information that reaches the markets (Ross, 1989), a different dynamic between large and small firms is consistent with a reality in which the prices of large companies react to new information immediately, while the prices of small ones respond with a lag.

In the APARCH and GJR specifications, all β_1 estimates are statistically significant. Once again, we find evidence of a higher autoregressive persistence in small firms' volatility than in larger companies. Therefore, this conclusion can be generalized as it does not depend on a given model, it is verified in every GARCH specification.

In its turn, if we analyse at the sector index level, we understand that such features are more evident in some cases than in others. For instance, those characteristics are stronger in Consumer Staples and Financial Services, with large firms exhibiting high significant values of α_1 and lower values of β_1. On the other hand, such dynamics are residual or non-existent in the Energy Sector, with shocks impacting equally irrespective of the firms' capitalization value.

Table 2 – In-Sample Estimation and LR Test Results

	Russell 2000 Index			**S&P500 Index**		
	APARCH	GARCH	GJR	APARCH	GARCH	GJR
c	0,000	0,001*	0,000*	0,000*	0,001*	0,001*
θ_1	-0,058*	-0,058*	-0,053*	-0,063*	-0,073*	-0,066*
θ_2	0,004	-0,014	-0,010	-0,003	-0,020	-0,013
θ_3	0,010	-0,009	0,006	0,004	-0,020	-0,008
θ_4	-0,027	-0,037	-0,028	-0,003	-0,027	-0,011
θ_5	-0,024	-0,036	-0,034	-0,026#	-0,043#	-0,037#
ω	0,000	0,000	0,000	0,000	0,000	0,000
α_1	0,072*	0,094*	0,000	0,105*	0,137*	0,000
β_1	0,923*	0,892*	0,902*	0,899*	0,859*	0,866*
γ_1	0,985*		0,153*	0,999*		0,231*
δ_1	0,954*			0,917*		
LR		90,35*	28,10*		133,38*	37,79*

*, # Denote significance at 1% and 5% level, respectively. φ_i and θ_i are the conditional mean equation parameters. ω, α_1, β_1, γ_1, and δ_1 are the conditional variance parameters.

By implementing two asymmetric specifications, we intend to investigate if the empirical results of Black (1976) and Christie (1982) are still valid, relating them subsequentially to firm size and industry sector.

In the GJR specification, the asymmetric coefficient is statistically significant and always different from 0, leading us to conclude that the impact of news is asymmetric. To understand the sign of such asymmetry, we must recognize that while good news has an α_1 impact on volatility, bad news has an $\alpha_1 + \gamma_1$ impact. In this dissertation, the in-sample results show that γ_1 is statistically significant and higher than 0. Hence, our results reveal that bad news lead to more stock volatility than good news.

For the APARCH model and regardless of the stock index, the estimate for the asymmetric coefficient is also statistically significant and positive. Once more, we meet the expectation of financial markets becoming more volatile after a negative shock than a positive one. Therefore, our results are in accordance with most previous literature, supporting the existence of an asymmetric dynamic in returns volatility.

Afterwards, and before concluding if such effect is influenced by firm size and its industry sector, we will investigate which specification better fits time-varying volatility.

According to the LR test, the asymmetric models outperform the symmetric GARCH. The null hypothesis is rejected in almost all indexes, thus being the special cases a worst fit for the data than the APARCH model. There is a single exception here, the GJR specification is superior to the unrestricted model in the Utility Sector, but only for large firms. Nonetheless, both APARCH and GJR specifications incorporate an asymmetric coefficient which highlights the importance of including this additional term in the conditional variance equation. From the LR test results, we conclude that APARCH is the model more prone to have generated the data.

Thus far, we have validated that all firms exhibit an asymmetric behaviour in returns' volatility. However, Table 2 shows that large-capitalization companies tend to be more affected by the arrival of bad news than small ones. Here, we meet the expectations of Dzieliński, Rieger, and Talpsepp (2018), with firms under a greater level of attention (usually large firms) showing a stronger asymmetry in volatility. According to them, this asymmetry in volatility is driven by asymmetric attention.

Despite this generalized behaviour, it is relevant to highlight that dynamics tend to change from sector to sector. In the Energy, Financials, Materials, and Utility Sectors, small firms exhibit a stronger asymmetric dynamic than larger firms, countering the overall dynamic. Moreover, in the Consumer Staples and Health Sectors we have detected an equally strong asymmetry regardless of the firm's capitalization value.

Lastly, it is also worth highlighting that the power parameter (δ) in the APARCH specification is always close to the unity and consequently far from 2. Therefore, a power transformation in the conditional heteroscedasticity equation around 1 appears to be a better fit for the data, as this specification outperforms the other models in analysis (where δ equals 2) in almost all cases.

4.2.3 Out-of-Sample Estimation and Empirical Results

The out-of-sample methodology is conducted through the HN test, which confronts three competing forecasts with the Actual Quantity. The volatility forecasts for each conditional variance model are estimated in RStudio, by running a rolling window forecast from the "ugarchforecast" function. For the Actual Quantity, and since we are evaluating a variable that is unobservable even *ex-post*, we use an unbiased estimator. For the variable of interest in this paper, volatility, the squared return (assuming a zero mean) on an asset over a given period is proposed as a conditionally unbiased estimator (Patton, 2011). The comparison between each volatility forecast and the squared returns was made after coding the HN test in EViews, a statistical package for Windows.

In Table 3, we present the HN test results and the associated probabilities. For this test statistic, we do not reject the null hypothesis for a 1% significance level in sixteen of the twenty indexes. In all those cases, the forecasts of GARCH and GJR specifications are inferior to the APARCH model forecasts, thereby not providing additional insights apart from those already included in the superior forecast. This rejection may also be due to the competing forecasts being very similar or a wide variability in the sample.

However, the null hypothesis is rejected in the Consumer Discretionary sector, Technological sector for large-cap firms, and S&P500 Index. In those cases, a combination of the GARCH and/or GJR forecasts with those of APARCH will lead to enhanced forecast performance. Nevertheless, across those four indexes, the significance level associated to the test statistic is still higher in the APARCH specification than in the two special cases.

Therefore, and with such knowledge in hand, we conclude that APARCH forecasts are more likely to encompass GARCH and GJR forecasts than the other way around.

Table 3 – Harvey-Newbold Test Results

Index	APARCH	GARCH	GJR
RGUSDS	**15,455 [0,0001]**	44,400 [0,0000]	24,986 [0,0000]
RGUSSS	**0,199 [0,6553]**	24,938 [0,0000]	8,643 [0,0034]
RGUSES	**0,881 [0,3481]**	13,442 [0,0003]	6,841 [0,0090]
RGUSFS	**0,397 [0,5289]**	12,110 [0,0005]	3,345 [0,0677]
RGUSHS	**4,114 [0,0428]**	53,874 [0,0000]	17,640 [0,0000]
RGUSPS	**1,474 [0,2250]**	22,606 [0,0000]	9,688 [0,0019]
RGUSTS	**6,471 [0,0111]**	41,002 [0,0000]	22,331 [0,0000]
RGUSMS	**0,064 [0,8004]**	12,541 [0,0004]	2,741 [0,0981]
RGUSUS	**0,774 [0,3791]**	10,503 [0,0012]	2,041 [0,1534]
RUT	**2,333 [0,1270]**	31,956 [0,0000]	13,896 [0,0002]
S5COND	**9,779 [0,0018]**	36,182 [0,0000]	24,540 [0,0000]
S5CONS	**0,512 [0,4746]**	13,411 [0,0003]	1,171 [0,2794]
S5ENRS	**2,546 [0,1109]**	22,260 [0,0000]	10,761 [0,0011]
S5FINL	**0,186 [0,6662]**	7,366 [0,0068]	5,001 [0,0256]
S5HLTH	**2,493 [0,1147]**	25,933 [0,0000]	13,849 [0,0002]
S5INDU	**3,095 [0,0788]**	15,209 [0,0001]	10,077 [0,0015]
S5INFT	**9,958 [0,0016]**	39,127 [0,0000]	26,202 [0,0000]
S5MATR	**4,559 [0,0330]**	24,340 [0,0000]	8,557 [0,0035]
S5UTIL	**5,654 [0,0176]**	10,264 [0,0014]	5,741 [0,0168]
S&P500	**10,703 [0,0011]**	37,794 [0,0000]	31,503 [0,0000]

4.3 Spillover Effects Analysis

4.3.1 Pairwise Granger Causality Test

To study the existence of spillover effects, we extended the already discussed univariate analysis to a multivariate model, as the goal now is to establish the relationship between large and small firms.

Firstly, and before proceeding with the correlation analysis, we investigated how much of the current S&P500 level can be explained by past values of the RUT and vice versa. It is not correct to state that correlation necessarily means causation -i.e. just because two assets are correlated it is not implied that they cause one another. In econometrics, there are multiple examples where two subjects are closely related, but that in the end, their association turned out to be spurious. For a better comprehension of the causality between the pair of return time series in analysis, we computed the Pairwise Granger Causality Test using 5 lags (a usual trading week) in the regressions:

$$RUT_t = \alpha_0 + \alpha_1 RUT_{t-1} + \cdots + \alpha_5 RUT_{t-5} + \beta_1 S\&P500_{t-1} + \cdots + \beta_5 S\&P500_{t-5} \quad (21)$$

$$S\&P500_t = \alpha_0 + \alpha_1 S\&P500_{t-1} + \cdots + \alpha_4 S\&P500_{t-5} + \beta_1 RUT_{t-1} + \cdots + \beta_5 RUT_{t-5} \quad (27)$$

For both regressions, the results in Table 4 show a p-value lower than a 5% significance level, which leads us to reject the null hypothesis. Therefore, there is a feedback effect between the two indexes, with the S&P 500 Granger causing the RUT and the RUT Granger causing the S&P 500.

Table 4 – Pairwise Granger Causality Test Results

	F-Test	P-value
H0: S&P500 does not Granger-cause RUT	6,225	0,0000094
H0: RUT does not Granger-cause S&P500	5,516	0,0000459

After guaranteeing that this relationship is meaningful and the stationarity in index returns (in Annex C), we proceeded with VAR model estimation according to BIC.

4.3.2 Diagnostic Testing

Multiple diagnostic tests were computed to conclude about the VAR model's features.

The multivariate version of the Jarque-Bera test is applied to the residuals of the VAR model estimated. As in the univariate analysis, we reject the null hypothesis of the test-statistic for any significance level (see Table 5). Hence, the process is not Multivariate Normal distributed. To accommodate fatter tales in the distribution, we will consider the Multivariate Student's T distribution from now on.

Furthermore, to detect structural changes in linear regression relationships, we implement a test from the generalized fluctuation test framework, the CUSUM test based on recursive residuals. As the upper and lower critical lines were never exceeded (Figure 5), we conclude that the system is stable over time.

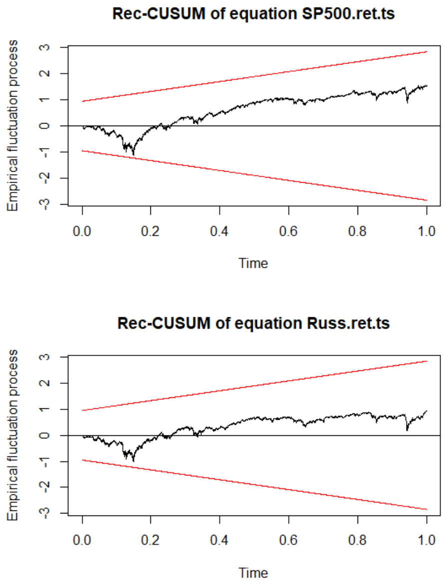

Figure 3 – CUSUM Test based on Recursive Residuals

Since stability in the VAR process is guaranteed, we proceed with diagnosis and implement the asymptotic Portmanteau statistic for testing serial correlation up to the tenth order. According to BIC, the number of p terms to include in the VAR model are 2, however, these terms were not enough to guarantee the absence of serially correlated disturbances. Therefore, as serial correlation is an undesirable characteristic of the model, we successively increased the p term's value during VAR estimation until finding a specification that does not exhibit autocorrelation. The test statistic is presented in Table 5 and is associated to a p-value higher than a 5% significance level. Since we do not reject the null hypothesis, we verify the adequacy of the fit for up to tenth order serially correlated disturbances.

Lastly, we computed the multivariate ARCH-LM test. Once again, we reject the null hypothesis, with the result in Table 5 indicating the existence of time-varying volatility. To better comprehend those developments in financial markets, we will introduce the two multivariate GARCH models already described in the literature review.

Table 5 – Diagnostic Statistics

	VAR Model
Jarque-Bera	22.094*
Pt.asymptotic	8,0329
ARCH-LM	2.117*

* Denote significance at 1% level.

4.3.3 DCC and aDCC Model Estimation and Empirical Results

A constant conditional correlation approach was already discarded by practitioners. Therefore, we implement here specifications that allow time-varying conditional correlations.

Differently from what was adopted in the univariate analysis, in the multivariate case we will not split the time series, proceeding now with a full-sample analysis. For that purpose, we executed the same summary and preliminary statistics as in the univariate methodology for the in-sample part. Once again, the results exhibit no serial correlation but display ARCH effects (Annex C).

The first step of DCC and aDCC model building is to fit a univariate GARCH specification to the time series. The empirical results in 4.2.2 showed that the APARCH model outperforms the other specifications under analysis. As a result, we consider the APARCH specification when modelling the time-varying volatility of S&P500 and RUT indexes. Using the "rmgarch" package from R software, we were able to compute the APARCH parameter estimates. The ARCH, GARCH, and asymmetric estimates are all statistically significant (Annex D), thus validating our decision to have dealt with conditional heteroscedasticity and asymmetric volatility in index returns.

Afterwards, we estimated the two parametrizations of dynamic conditional correlations. In both cases, the DCC joint estimates (DCC(a) and DCC(b)) are non-negative and statistically different from zero, confirming the time-varying nature of conditional correlations. Hence, we meet previous observations of Billio, Caporin, and Gobbo (2006), where they rejected the assumption of a constant dynamic in conditional correlations.

The attained estimates are presented in Table 6. The news term (DCC(a)) shows that the DCC model exhibits a more sudden reaction to new information than the asymmetric specification. Nevertheless, the latter displays a higher persistence after a market event since this specification displays a higher DCC(b) estimate.

Table 6 – Models Estimates and BIC values for DCC and aDCC Parametrizations

	DCC(1,1)	**aDCC(1,1)**
DCC(a)	0,070*	0,055*
DCC(b)	0,906*	0,909*
DCC(g)		0,025*
BIC	-14,067	-14,066

* Denote significance at 1% level.

Furthermore, there is evidence of an asymmetric response in correlations to joint bad news (both returns being negative), as the asymmetric parameter in the aDCC is statistically significant for a 1% significance level. Therefore, the correlation between the S&P500 and RUT will be higher after a joint negative shock than after a positive one with the same magnitude.

Lastly, we computed the half-life for the DCC specification – i.e. the expected time at which a shock to correlation is halfway dissipated. For the relationship in analysis, and following the approximation suggested in Engle and Sheppard (2001), we find that shocks to correlation are highly persistent, with a half-life of almost 29 weeks.

4.3.4 Graphical Examination of Dynamic Conditional Correlations

The estimation of these two multivariate models contributed to an improved comprehension of the relationship in analysis. We found evidence of an asymmetry in the correlation between large and small firms, but the magnitude of this association is still missing. To better understand the degree of this relationship, we extracted the plots of aDCC conditional variances, and DCC and aDCC time-varying conditional correlations. The three plots are depicted in Figures 6, 7, and 8.

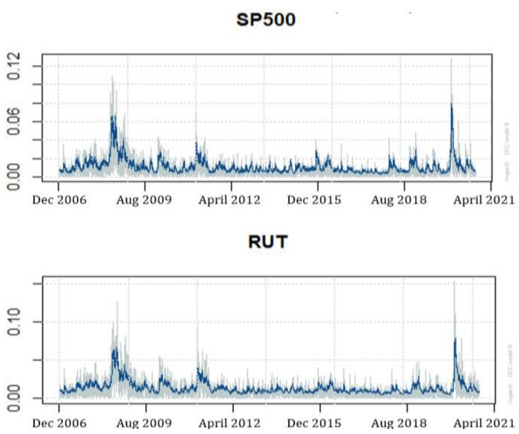

Figure 6 – aDCC Conditional Variance (in blue) vs |returns|

From Figure 6, we verify that the two indexes exhibit a similar conditional variance for the period in analysis. They tend to record sharp increases in periods of greater uncertainty, as during the 2007-2008 Financial Crisis and the peak of Coronavirus Pandemic, and low and relatively steady volatility in calm periods.

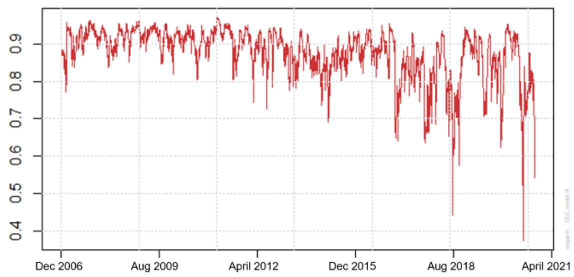

Figure 7 – DCC of SP500 and RUT

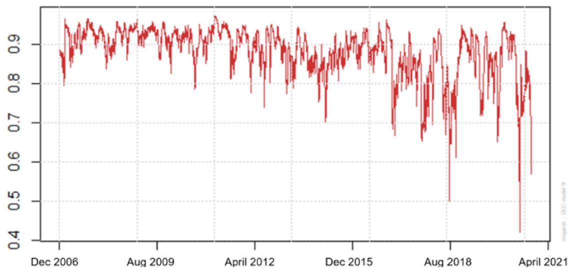

Figure 8 – aDCC of SP500 and RUT

In terms of conditional correlation, the DCC and aDCC specifications display a similar behaviour for the 2006-2020 period. The only substantial difference is their absolute values, with the DCC model exhibiting a more extensive range of values. In Figure 7, the DCC model's correlation falls below 0.4 and 0.5 during mid-2018 and mid-2020, whereas the aDCC's estimated correlation never dropped below those values for the same periods (Figure 8).

From Figures 7 and 8, we identify a clear and positive correlation between the two indexes, showing that RUT and S&P500 move together over time and strengthening the spillover hypothesis tested in this paper. Notice that they signal a strong contagion across the two indexes with correlations above 0.8 for much of the time, and an even greater co-movement during periods of financial turmoil. Therefore, there is small scope for those interested in a diversification strategy. If the goal of an investor is to diversify, that strategy will not be successfully reached by incorporating different-sized firms in the same portfolio.

The estimated correlations unveiled a strong association between the two indexes during periods of financial turmoil. For example, after reaching a record peak on 19th February, the S&P500 fell to a low of 2,237 by March 23, a decline of 34%. For that same period, the RUT experienced an even sharper decline, falling almost 41%. Note that, during this period, the correlation between the two indexes increased significantly and was never below 0.93.

In this dissertation, we were able to demonstrate that during periods of high turbulence is recorded a higher correlation than in periods of relative optimism. Thus, in addition to an asymmetric dynamic in conditional volatility, we ended up finding that there is also an asymmetric response in conditional correlation. Therefore, correlation between the S&P500 and the RUT may be greater for a downside move than for a positive shock of the same magnitude.

5. Conclusion

Everyone is exposed to any kind of risk. In this paper, we provide insights acquired throughout this study, helping investors to learn more about risk and to avoid costly decisions.

This dissertation expands previous investigation on the risk-return trade-off topic. It follows the empirical research of Banz (1981) and Reinganum (1981), the pioneers in understanding the decisive role that firm size takes in Financial Economics. Through the analysis of market volatility and correlation dynamics over different-sized firms, we were able to improve the awareness of investors regarding risk.

We examined two well-known indexes for the 2006-2020 period: the S&P500 index, considered as a measurement of 500 of the largest firms listed in the USA stock exchanges, and the RUT, the most used benchmark of small-capitalization firms.

For the first topic under review, the existence of an asymmetric volatility phenomenon in large and small firms, we additionally investigated for specific dynamics across the sector indexes of S&P500 and RUT. Our empirical analysis matches the existing literature of Black (1976) and Christie (1982), as we verify that regardless of the firm's capitalization value and sector, companies exhibit an asymmetric dynamic in volatility.

The in-sample results show that, based on the LR test, the APARCH specification outperforms the special cases in study, thus signalling that a model with an asymmetric term is a better fit for the data. Furthermore, we identify that overall, large companies are more affected by the arrival of bad news than small firms. Hence, we meet the results of Dzieliński, Rieger, and Talpsepp (2018), with firms under a greater media coverage (usually large firms) displaying a stronger asymmetric dynamic in volatility. They proved that an asymmetry in volatility is a result of asymmetric attention, with bad news about large companies generating far more attention than bad news about small firms.

At the sector group level, the results differ from sector to sector. While Consumer Discretionary, Industrials and Information Technology follow the overall dynamic, in the Energy, Financials, Materials, and Utility sectors the asymmetric behaviour is stronger for small companies. At the same time, in Consumer Staples and Health sectors the asymmetric behaviour is not related to the capitalization value, the phenomenon is equally strong for large and small firms. These results highlight the value of a more detailed analysis at the sector group level, as it becomes evident that industry factors play an important role in investment strategy.

Regarding the out-of-sample results and according to the HN test, the APARCH forecasts are more likely to encompass GARCH and GJR forecasts than otherwise, as they present the highest significance levels across all indexes in study. Therefore, the APARCH model is more capable to predict the returns' volatility in financial markets than GJR and/or GARCH specifications.

Lastly, since the current S&P500 level can be explained by past values of RUT and the other way around, we found out that there is a meaningful relationship between large and small firms. Consequently, to improve our knowledge about this relationship, we implemented a multivariate approach. The extension of a univariate framework enabled us to assess the time-varying correlation and to verify the existence of spillover effects between different-sized firms.

Our empirical research signals a strong co-movement between RUT and S&P500 indexes. The conditional variance of RUT describes a similar dynamic to S&P500's, and the time-varying correlation between the two is greater than 0.8 for most of the time. This association becomes even stronger during periods of financial turmoil, as conditional variance increases significantly and equally for both indexes, and correlation reaches a higher level than during periods of positivity and calm in financial markets. In the end, we came up with evidence of an asymmetric dynamic not only in volatility but also in conditional correlation.

In this paper, the empirical research relies on firms incorporated in the USA stock exchanges, which can ultimately be interpreted as a limitation. Nevertheless, we follow Ferreira and Gama (2005) insights of small country risk in more recent times, a consequence of increasing economic integration and corporate globalization. Therefore, an investigation from the world's largest economy can be widely accepted.

To account for sector risk, we investigated the asymmetric volatility phenomenon over different sector groups, but the same was not applied for the study of spillover effects. We believe that future research on the topic would be of great interest, allowing us to comprehend the transmission mechanisms between the several sectors and the predominant role that some might have over the others. Furthermore, implementing GACRH specifications that model more than two limiting regimes would be of great interest. The non-application of such models is a limitation of this study, explained by the difficulty to find fitting and forecasting methods and the complexity to create them. For instance, if future research accommodates an example of these, as the Flexible Coefficient GARCH of Medeiros and Veiga (2009), it would be able to differentiate the impact of "very good/bad" news from "good/bad" news and still contemplate a middle regime for calm periods. This way, we would be closer to understand how volatility moves throughout time.

In sum, the attained results will have a practical contribution and an updated response for multiple fields. For instance, they can be used in multiple financial applications such as option pricing, allowing investors to attain an improved estimation. The achieved results can also be reckoned in all other applications that rely on volatility forecasts, such as value-at-risk measurement. All those applications should consider an asymmetric dynamic in returns' volatility, a different degree of asymmetry between different-sized firms and take into account the distinctive features of the firm's sector under analysis.

At last, the evidence of transmission mechanisms between large and small firms must be considered in portfolio management. The recognition of spillover effects will be extremely useful for the stocks' selection and asset allocation processes, leading investors to look for assets that increase the diversification benefit in place of a combination of large and small firms.

References

Acharya, V., Pedersen, L.H. (2005). Asset pricing with liquidity risk. *Journal of financial economics*, *77*, 375-410.

Al-Rjoub, S. A., Varela, O., & Kabir Hassan, M. (2005). The size effect reversal in the USA. Applied Financial Economics, 15(17), 1189-1197.

Andrei, D., & Hasler, M. (2015). Investor attention and stock market volatility. *The review of financial studies*, *28*(1), 33-72.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T. J., & Pedersen, L. H. (2018). Size matters, if you control your junk. *Journal of Financial Economics*, *129*(3), 479-509.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, *9*(1), 3-18.

Ben-Zion, U., & Shalit, S. S. (1975). Size, leverage, and dividend record as determinants of equity risk. *The Journal of Finance*, *30*(4), 1015-1026.

Billio, M., Caporin, M., & Gobbo, M. (2006). Flexible dynamic conditional correlation multivariate garch models for asset allocation. *Applied Financial Economics Letters*, *2*(02), 123-130.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 meetings of the business and economic statistics section of the American Statistical Association*, 177-181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, 542-547.

Boudoukh, J., M. Richardson and R. Whitelaw (1994). A Tale of Three Schools: Insights on Autocorrelations of Short-Horizon Returns. *Review of Financial Studies, 7*, 539–73.

Brooks, C. (2019). *Introductory econometrics for finance*. Cambridge university press.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, *37*(2), 149-163.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of financial Economics*, *31*(3), 281-318.

Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial econometrics*, *4*(4), 537-572.

Cavaglia, S., Brightman, C., & Aked, M. (2000). The increasing importance of industry factors. *Financial Analysts Journal, 56*(5), 41-54.

Chelley-Steeley, P. L., & Steeley, J. M. (1996). Volatility, leverage and firm size: the UK evidence. *The Manchester School*, *64*(S1), 83-103.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of financial Economics*, *10*(4), 407-432.

Conrad, J., Gultekin, M. N., & Kaul, G. (1991). Asymmetric predictability of conditional variances. *The Review of Financial Studies*, *4*(4), 597-622.

Curto, J., & Pinto, J. (2012). Predicting the financial crisis volatility. *Economic Computation And Economic Cybernetics Studies and Research Journal*, *46*(1), 183-195.

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

Dimson, E., & Marsh, P. (1999). Murphy's law and market anomalies. *The Journal of Portfolio Management*, *25*(2), 53-69.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, *1*(1), 83-106.

Dzieliński, M., Rieger, M. O., & Talpsepp, T. (2018). Asymmetric attention and volatility asymmetry. *Journal of Empirical Finance*, *45*, 59-67.

Elyasiani, E., Kalotychou, E., Staikouras, S. K., & Zhao, G. (2015). Return and volatility spillover among banks and insurers: Evidence from pre-crisis and crisis periods. *Journal of Financial Services Research*, *48*(1), 21-52.

Enders, W. (2015). Applied econometric time series fourth edition.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Engle, R. F., & Sheppard, K. (2001). *Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH*. Working Paper no. 8554, National Bureau of Economic Research.

Ferreira, M. A., & Gama, P. M. (2005). Have world, country, and industry risks changed over time? An investigation of the volatility of developed stock markets. *Journal of Financial and Quantitative Analysis*, 195-222.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779-1801.

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, *23*(4), 365-380.

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?. *Journal of applied econometrics*, *20*(7), 873-889.

Hansen, P. R., & Timmermann, A. (2012). Choice of sample split in out-of-sample forecast evaluation.

Harris, R. D., & Pisedtasalasai, A. (2006). Return and volatility spillovers between large and small stocks in the UK. *Journal of Business Finance & Accounting*, *33*(9-10), 1556-1571.

Harvey, D., & Newbold, P. (2000). Tests for multiple forecast encompassing. *Journal of Applied Econometrics*, *15*(5), 471-482.

Hasan, I., & Francis, B. B. (1998). Macroeconomic factors and the asymmetric predictability of conditional variances. *European Financial Management*, *4*(2), 207-230.

Henry, O. T., & Sharma, J. (1999). Asymmetric conditional volatility and firm size: evidence from Australian equity portfolios. *Australian Economic Papers*, *38*(4), 393-406.

Horowitz, J. L., Loughran, T., & Savin, N. E. (2000). Three analyses of the firm size premium. *Journal of Empirical Finance*, *7*(2), 143-153.

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review/Revue Internationale de Statistique*, 163-172.

Li, Y., & Giles, D. E. (2015). Modelling volatility spillover effects between developed stock markets and Asian emerging stock markets. *International Journal of Finance & Economics*, *20*(2), 155-177.

Liew, V. K. S. (2004). Which lag length selection criteria should we employ?. *Economics bulletin*, 3(33), 1-9.

Liow, K. H. (2015). Volatility spillover dynamics and relationship across G7 financial markets. *The North American Journal of Economics and Finance*, *33*, 328-365.

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303.

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction?. *The review of financial studies*, *3*(2), 175-205.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, *36*(4), 394–419.

Mech, T. S. (1993). Portfolio return autocorrelation. *Journal of Financial Economics*, *34*(3), 307-344.

Medeiros, M. C., & Veiga, A. (2009). Modeling multiple regimes in financial volatility with a flexible coefficient GARCH (1, 1) model. *Econometric Theory*, 25(1), 117-161.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246-256.

Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of financial Economics*, *9*(1), 19-46.

Reinganum, M. R. (1982). A direct test of Roll's conjecture on the firm size effect. *The Journal of Finance*, *37*(1), 27-35.

Roll, R. (1981). A possible explanation of the small firm effect. *The Journal of Finance*, *36*(4), 879-888.

Ross, S. A. (1989). Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy. *The Journal of Finance*, *44*(1), 1-17.

Stoll, H. R., & Whaley, R. E. (1983). Transaction costs and the small firm effect. *Journal of Financial Economics*, *12*(1), 57-79.

Taylor, S. J. (2011). *Asset price dynamics, volatility, and prediction*. Princeton university press.

Yu, C. H., & Wu, C. (2001). Economic sources of asymmetric cross-correlation among stock returns. *International Review of Economics & Finance*, *10*(1), 19-40.

Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, *18*(5), 931-955.

Annexes

Annex A - Summary and Diagnostic Statistics of Index Sectors' Returns

Russell 2000 Index					
	RGUSDS	**RGUSSS**	**RGUSES**	**RGUSFS**	**RGUSHS**
Mean	0,00019	0,00035	-0,00024	0,00007	0,00036
Median	0,00091	0,00084	0,00071	0,00041	0,00104
Maximum	0,08352	0,06846	0,18807	0,10913	0,09195
Minimum	-0,13104	-0,08595	-0,18606	-0,16571	-0,08744
Std. Dev.	0,01778	0,01284	0,02736	0,01859	0,01614
Skewness	-0,336	-0,273	-0,486	-0,200	-0,360
Kurtosis	7,787	6,537	8,362	11,994	5,793
Jarque-Bera	2452*	1344*	3115*	8504*	873*
ADF (intercept)	-36,132*	-39,684*	-35,725*	-40,453*	-36,778*
Ljung-Box	13,028	7,056	15,481	3,059	15,785
ARCH-LM	269*	336*	294*	398*	280*

Russell 2000 Index				
	RGUSPS	**RGUSTS**	**RGUSMS**	**RGUSUS**
Mean	0,00018	0,00036	0,00033	0,00015
Median	0,00093	0,00137	0,00117	0,00051
Maximum	0,09152	0,08947	0,10606	0,09678
Minimum	-0,12432	-0,11010	-0,14373	-0,09664
Std. Dev.	0,01802	0,01721	0,01956	0,01301
Skewness	-0,315	-0,251	-0,407	-0,237
Kurtosis	7,499	6,260	7,708	10,269
Jarque-Bera	2166*	1142*	2395*	5568*
ADF (intercept)	-37,040*	-36,909*	-36,664*	-39,492*
Ljung-Box	7,299	14,622	12,418	10,940
ARCH-LM	282*	268*	269*	287*

S&P500 Index					
	S5COND	**S5CONS**	**S5ENRS**	**S5FINL**	**S5HLTH**
Mean	0,00030	0,00027	0,00008	-0,00010	0,00029
Median	0,00090	0,00053	0,00048	0,00050	0,00070
Maximum	0,12313	0,08835	0,16960	0,17201	0,11713
Minimum	-0,10099	-0,06648	-0,16884	-0,18639	-0,07415
Std. Dev.	0,01450	0,00922	0,01831	0,02319	0,01126
Skewness	-0,106	-0,004	-0,369	-0,121	-0,081
Kurtosis	10,709	13,012	14,535	16,266	12,451
Jarque-Bera	6240*	10516*	14018*	18470*	9373
ADF (intercept)	-38,139*	-40,978*	-40,794*	-38,363*	-39,829*
Ljung-Box	12,343	5,923	4,756	39,514*	12,221
ARCH-LM	276*	287*	311*	466*	331*

S&P500 Index				
	S5INDU	**S5INFT**	**S5MATR**	**S5UTIL**
Mean	0,00020	0,00033	0,00015	0,00011
Median	0,00070	0,00093	0,00085	0,00086
Maximum	0,09516	0,11461	0,12473	0,12684
Minimum	-0,09215	-0,09670	-0,12934	-0,08530
Std. Dev.	0,01439	0,01407	0,01692	0,01213
Skewness	-0,401	-0,078	-0,415	0,240
Kurtosis	8,716	9,716	9,987	14,628
Jarque-Bera	3496*	4735*	5194*	14210*
ADF (intercept)	-37,311*	-38,559*	-37,560*	-39,829*
Ljung-Box	7,726	14,283	16,252	9,100
ARCH-LM	283*	298*	275*	278*

Annex B – Index Sectors' In-Sample Estimation and LR Test Results

	RGUSDS			RGUSSS			RGUSES		
	APARCH	GARCH	GJR	APARCH	GARCH	GJR	APARCH	GARCH	GJR
c	0,000*	0,001*	0,000	0,000#	0,001*	0,001*	0,000	0,001#	0,000
θ_1				-0,071	-0,080*	-0,074*			
θ_2				-0,034	-0,046#	-0,043#			
θ_3				0,009	0,001	0,001			
θ_4				-0,013	-0,027	-0,019			
θ_5				-0,021	-0,032	-0,029			
ω	0,000*	0,000	0,000	0,000*	0,000	0,000	0,000	0,000#	0,000
α_1	0,071*	0,085*	0,017*	0,064*	0,069*	0,000	0,066*	0,085*	0,016
β_1	0,927*	0,906*	0,914*	0,916*	0,914*	0,912*	0,933*	0,907*	0,918*
γ_1	0,689*		0,113*	0,999*		0,125*	0,688*		0,103*
δ_1	1,108*			1,113*			0,979*		
LR		55,42*	11,98*		69,29*	17,17*		52,37*	17,92*

	RGUSFS			RGUSHS			RGUSPS		
	APARCH	GARCH	GJR	APARCH	GARCH	GJR	APARCH	GARCH	GJR
c	0,000	0,001*	0,000*	0,001	0,001*	0,001#	0,000	0,001*	0,000
φ_1	-0,896*	-0,913*	-0,903*	-0,021	-0,031	-0,023	-0,037	-0,037	-0,034
φ_2							0,011	0,000	0,003
φ_3							0,016	0,003	0,015
φ_4							-0,023	-0,032	-0,027
φ_5							-0,044#	-0,047#	-0,047#
θ_1	0,811*	0,818*	0,812*						
θ_2	-0,076*	-0,093*	-0,088*						
θ_3	0,003	-0,011	-0,004						
θ_4	-0,031	-0,042	-0,032						
θ_5	-0,072	-0,079*	-0,075*						
θ_6	-0,030	-0,049	-0,037						
θ_7	0,027	0,009	0,021						
θ_8	0,038	0,035	0,038						
ω	0,000	0,000	0,000	0,001	0,000	0,000*	0,000	0,000	0,000
α_1	0,088*	0,101*	0,025	0,072*	0,096*	0,000	0,063*	0,087*	0,000
β_1	0,917*	0,891*	0,899*	0,909*	0,882*	0,882*	0,930*	0,901*	0,916*
γ_1	0,699*		0,130*	0,999*		0,158*	0,964*		0,132#
δ_1	0,949*			0,908*			1,048*		
LR		70,90*	25,65*		86,63*	22,71*		78,19*	18,44*

	RGUSTS			RGUSMS			RGUSU		
	APARCH	GARCH	GJR	APARCH	GARCH	GJR	APARCH	GARCH	GJR
c	0,001#	0,001*	0,001#	0,000	0,001*	0,000	0,000	0,000*	0,000
φ_1	-0,037	-0,035	-0,038	-0,493*	0,392*	0,376*			
φ_2				-0,996*	-0,886*	-0,889#			
θ_1				0,491*	-0,402*	-0,389*	-0,090*	-0,092*	-0,091*
θ_2				0,999*	0,910*	0,916#	-0,009	-0,016	-0,015
θ_3							0,007	-0,001	0,004
θ_4							-0,014	-0,016	-0,016
θ_5							-0,023	-0,027	-0,026
ω	0,000	0,000#	0,000*	0,000	0,000	0,000	0,000	0,000	0,000
α_1	0,067*	0,086*	0,000	0,051*	0,085*	0,000	0,058*	0,085#	0,002
β_1	0,922*	0,896*	0,894*	0,942*	0,905*	0,930*	0,925*	0,898*	0,922*
γ_1	0,991*		0,156*	0,999*		0,116*	0,760*		0,113*
δ_1	0,953*			1,180*			1,366*		
LR		78,86*	23,60*		66,32*	12,06*		44,09*	6,50#

	S5COND			S5CONS			S5ENRS		
	APARCH	GARCH	GJR	APARCH	GARCH	GJR	APARCH	GARCH	GJR
c	0,000#	0,001*	0,001*	0,000*	0,001*	0,000*	0,000	0,001*	0,000#
φ_1	-0,007	-0,013	-0,008	0,253*	-1,819*	-0,276*	-0,177*	0,008	-0,077
φ_2	-0,021	-0,034	-0,029	0,724*	-0,964*	0,720*	0,580*	0,744*	0,675
θ_1				-0,329*	1,740*	0,202*	0,122#	-0,072	0,016
θ_2				-0,727*	0,807*	-0,769*	-0,580*	-0,741*	-0,672#
θ_3				0,064*	-0,089*	0,022*	0,010	0,014	0,015
ω	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
α_1	0,083*	0,110*	0,010	0,074*	0,136*	0,017	0,072*	0,085*	0,014
β_1	0,910*	0,881*	0,888*	0,877*	0,837*	0,850*	0,927*	0,908*	0,920*
γ_1	0,872*		0,171*	0,998*		0,187*	0,661*		0,103*
δ_1	1,129*			1,379*			1,112*		
LR		78,16*	18,59*		70,92*	21,62*		47,36*	12,95*

	S5FINL			S5HLTH			S5INDU		
	APARCH	GARCH	GJR	APARCH	GARCH	GJR	APARCH	GARCH	GJR
c	0,000#	0,001*	0,000#	0,000*	0,001*	0,001*	0,000#	0,001*	0,000#
φ_1				-0,654*	-0,566	-0,520	-0,865*	-0,888*	-0,866*
φ_2				-0,185*	-0,099	-0,095			
θ_1	-0,077*	-0,085*	-0,079	0,606*	0,506	0,467	0,846*	0,866*	0,847*
θ_2				0,123*	0,029	0,034	-0,021	-0,033	-0,025
θ_3							0,010	-0,013	0,002
θ_4							0,026	0,009	0,022
ω	0,000	0,000	0,000	0,001	0,000	0,000*	0,000	0,000	0,000
α_1	0,113*	0,139*	0,051*	0,086*	0,116*	0,000	0,068*	0,112*	0,000
β_1	0,897*	0,860*	0,867*	0,900*	0,859*	0,870*	0,916*	0,880*	0,899*
γ_1	0,568*		0,157*	0,999*		0,183*	0,999*		0,170*
δ_1	1,093*			0,895*			1,286*		
LR		55,48*	13,38*		86,90*	25,81*		95,40*	13,45*

	S5INFT			S5MATR			S5UTIL		
	APARCH	GARCH	GJR	APARCH	GARCH	GJR	APARCH	GARCH	GJR
c	0,001*	0,001*	0,001*	0,000	0,001*	0,000#	0,000*	0,001*	0,000*
φ_1	0,093#	0,818*	0,483	-0,031#	-0,030	-0,027	-0,611	-0,566	-0,535
θ_1	-0,124*	-0,848*	-0,514				0,566	0,519	0,487
θ_2							-0,054	-0,054	-0,053
θ_3							-0,034	-0,030	-0,033
θ_4							-0,024	-0,023	-0,025
θ_5							-0,035	-0,041	-0,040
ω	0,000	0,000	0,000*	0,000	0,000	0,000	0,000	0,000	0,000
α_1	0,088*	0,108*	0,000	0,070*	0,102*	0,007	0,067*	0,092*	0,038
β_1	0,898*	0,879*	0,871*	0,937*	0,893*	0,912*	0,909*	0,894*	0,907*
γ_1	0,999*		0,206*	0,892*		0,138*	0,270#		0,069*
δ_1	1,075*			0,910*			1,948*		
LR		96,06*	24,57*		73,02*	28,41*		11,78*	0,72

Annex C – Full-Sample Summary and Diagnostic Statistics of Index Returns

	Russell 2000 Index	**S&P500 Index**
Mean	0,00026	0,00027
Median	0,00101	0,00070
Maximum	0,08976	0,10957
Minimum	-0,15344	-0,12765
Std. Dev.	0,01651	0,01313
Skewness	-0,675	-0,553
Kurtosis	11,55	15,917
Jarque-Bera	11004*	24684*
ADF (intercept)	-42,716*	-45,308*
Ljung-Box	4,76	5,868
ARCH-LM	198*	205*

Annex D – APARCH Parameter Estimation

	Russell 2000 Index	S&P500 Index
c	0,000*	0,001*
φ_1	-0,058*	-0,679*
φ_2	0,009	-0,042*
φ_3	0,011	-0,003
φ_4	-0,013	-0,006*
φ_5	-0,009	-0,021*
φ_6	0,007	-0,025*
φ_7	0,020	0
θ_1	0	0,616*
ω	0,000	0,000*
α_1	0,081*	0,114*
β_1	0,920*	0,892*
γ_1	0,933*	0,963*
δ_1	0,938*	0,934*