Joseph Falzon

Testing the alpha-stable Lévy hypothesis on US stock market data

Evidence from a Fractal Process

School of Accounting and Finance
Master's Thesis in Finance
Master's Degree Programme in Finance

Vaasa 2024

UNIVERSITY OF VAASA
School of Accounting and Finance
Author: Joseph Falzon
Title of the Thesis: Testing the alpha-stable Lévy hypothesis on US stock market data. Evidence from a Fractal Process.
Degree: Master of Science in Economics and Business Administration
Programme: Finance
Supervisor: Associate Professor Klaus Grobys
Year: 2024 **Pages: 86**

ABSTRACT:

This study tests whether the alpha-stable Lévy hypothesis applies to US stock market data, namely, Dow Jones Industrial Average (DJIA). It makes use of daily data on the closing price of the DJIA, spanning from 28th May 1896 to 10th February 2023, a total of 32,820 trading days. Computation of Realised Variances (RVs) involved squaring daily returns of the asset over a specified time frame spanning from weekly RVs till semi-annual RVs were computed. 'Log-Log' Ordinary Least square (OLS) regressions were employed on each time-frame resulting in alpha exponents less than 2 for the 5-, 90, -125 resolutions confirming the main a priori hypothesis that DJIA has an infinite variance that is not defined. Moreover, by observing the 95% confidence interval using the uncertainty principle provided by the log-log regression, we find values ranging from 1.62 to 2.22 as the confidence interval for the power law exponent. It becomes evident that, regardless of the point estimate for other resolutions, all of them fall within this range of point estimate intervals. Therefore, from a statistical perspective, the invariance hypothesis is accepted. Sample split tests deliberately conducted to assess the constancy of power law exponents over time, confirmed that point estimators from the second subsample for time resolutions 5, 20, and 60 fell within the confidence interval range of the first sub sample in all cases. Hence, statistically, the power law exponents persisted across both sub-samples, implying the invariance quality of the power law. When compared against GARCH (1,1) model exponents, the power law null hypothesis prevailed in constancy whereas the former gave sample-specific point estimates.

KEYWORDS: Dow Jones Industrial Average, Power Law exponents, Realised variances, Time resolutions, OLS, invariance hypothesis, GARCH model

Acknowledgements

To my dearest mother Teresa Falzon;
Your love and sacrifices gave me everything I ever needed.

To my Master's thesis supervisor Associate Professor Klaus Grobys;
Your lectures in Quantitative Financial Data Analysis in MATLAB gave me an edge in the professional world, and I am grateful for the opportunity to learn from a genius like you.

To Associate Professor Tatiana King;
Your mastery in Financial Statement Analysis and Valuation and captivating teaching style have deeply motivated me, scoring a distinctive mark on my academic journey.

To Associate Professor Vanja Piljak;
Your belief in my potential has been a source of strength and inspiration.

To former Chief Risk Officer of Bank of Valletta Mr. Miguel Borg;
You were the first one to give me an opportunity in the professional work life.

To my Bachelor's thesis supervisor and Economist Mr. Owen Grech;
You encouraged me to pursue a Master's degree abroad, a decision I will forever cherish.

Contents

Figures

Tables

Abbreviations

ARCH	Autoregressive conditional heteroskedasticity
CAPM	Capital Asset Pricing Model
CDF	Cumulative Density Function
CLT	Central Limit Theorem
DJIA	Dow Jones Industrial Average
d.o.f	Degrees of Freedom

GARCH	Generalised Autoregressive conditional heteroskedasticity
GoF	Goodness of Fit
GMM	Generalised Method of Moments
KS	Kolmogorov Smirnov
LTCM	Long Term Capital Management
LLN	Law of Large Numbers
MLE	Maximum Likelihood Estimator
OLS	Ordinary Least Squares
RVs	Realised Variances
RMV	Realised momentum variance
S&P 500	Standard and Poor's 500

1 Introduction

1.1 Foreword

In academia, people often get excited about published research. But there is a serious problem hiding underneath: many studies cannot be successfully replicated.

Looking at the well-known case mentioned by Nassim Taleb in his book *The Black Swan* (2007), Long-Term Capital Management (LTCM) was a hedge fund located in Greenwich. They used trading strategies based on scientific research. These strategies involved taking considerable financial risks by borrowing a lot of money. LTCM started in 1994, and some of the people on its board were Myron S. Scholes and Robert C. Merton, who won the Nobel Memorial Prize in Economic Sciences in 1997 for their work on valuing derivatives. Backed by such notable figures, LTCM attracted significant initial investment. A total of 80 founding investors were persuaded to contribute a minimum of $10 million each; however, within four years of its establishment, LTCM found itself on the brink of collapse due to substantial losses that endangered investor capital (Grobys, 2023).

LTCM's reliance on models proved to be its downfall, as these models failed to predict unexpected shifts in the market. As a result, investors who placed their trust in these Nobel laureates suffered significant losses, highlighting the dangers of excessive reliance on academic quantitative models in financial markets. Investors should have a clear understanding of the risks associated with their portfolio and recognise that unforeseen 'black swan events' can happen more often than predicted by models. Taleb (2007) convincingly illustrates that events that appear highly improbable are actually more frequent than anticipated, often leading to significant consequences. These occurrences are frequently overlooked by experts because past data may not accurately predict future events.

Recent studies indicate that a significant portion of scientific studies published in prominent finance journals cannot be replicated (Harvey et al., 2018; Hou et al., 2020). Grobys (2022), proposes that factors such as p-hacking, pressure to publish, and bias in journal selection may

not fully account for this phenomenon. Instead, he suggests that the conventional methodologies commonly employed in finance research are inherently sample specific due to the statistical models and assumptions that do not capture the underlying characteristics of financial markets. Grobys's (2022) findings strongly suggest that the variance of variance is absent in the key financial markets under consideration. This raises doubts about the effectiveness of current research methodologies in finance.

1.2 Motivation

Scientific replication failure poses a significant challenge across various disciplines. Surprisingly, top-tier journals publish studies that are not replicable, indicating a systemic issue in scientific methodology.

Mandelbrot's (1963) influential study revealed that the theoretical variance of cotton price changes is either infinite or undefined. Subsequent research often overlooked this issue, relying on arguments such as Teichmoeller's (1971), which suggests that variance stabilises as sample size increases due to the convergence of return distributions to a normal distribution with time aggregation. However, Taleb (2020) contends that scholars cannot rely on Gaussian behaviour when kurtosis is infinite, even with lower moments existing. Even for power law exponents close to $\alpha \approx 3$, the Central Limit Theorem operates slowly and requires extensive data, which is often lacking in financial markets' history. Consequently, Gaussian methodologies may not be suitable for financial research.

Recent studies, including Renò and Rizza (2003) and Grobys (2021), have found evidence of power law behaviour in various financial contexts, such as unconditional volatility distribution in future markets and realised variances (RVs) in commodities, exchange rates, cryptocurrencies, and the Standard and Poor's (S&P) 500.

This study contributes to existing literature in several keyways. It complements a recent fresh approach taken by Grobys (2022) by modelling realised variances (RVs) as power laws and testing the invariance hypothesis on a stock market index. It also tests the invariance

hypothesis across different time frequencies and resolutions for a stock market index like the Dow Jones Index.

However, it also addresses a gap in the existing literature by comparing the consistency of point estimates derived from generalized autoregressive conditional heteroskedasticity (GARCH) models with those obtained from the power law null hypothesis across different time frequencies and resolutions. This comparison is essential for determining which model provides more robust and stable estimates across various data subsets.

1.3 Research question

This study delves into whether the alpha-stable Lévy hypothesis, which represents a power law process, accurately reflects the behaviour of US stock market data, specifically focusing on the Dow Jones Industrial Average (DJIA) index.

By closely examining the power law exponent, especially when it falls below 3, and considering the implications of undefined kurtosis—variance of the variance—for traditional statistical methods, the research aims to provide clearer insights into how well the alpha-stable Lévy hypothesis applies to financial data modelling. This investigation aims to offer valuable understanding into the underlying dynamics of the US stock market and potentially enhance the development of more reliable statistical methods tailored to financial market complexities.

1.4 Structure

This thesis comprises five main chapters. In Chapter 1, the topic is introduced along with the primary research question, alongside an explanation of the research motivation. Chapter 2 offers an examination of both the theoretical and empirical literature, while Chapter 3 outlines the methodology used in the research process. Chapter 4 presents the findings and discusses the robustness checks conducted on the methodology implemented in the previous

chapter. It also discusses the limitations and avenues for future research. Finally, Chapter 5 draws conclusions and outlines key policy implications derived from the study's findings.

2 Theoretical background

2.1 Replication failure in financial market research

Recently, Hou et al. (2018) undertook an investigation to examine whether cross-sectional asset pricing recorded in empirical finance literature conform to the presently accepted empirical finance criteria. The authors concluded that these studies cannot be reproduced in extended samples. They revealed that 82% of the 452 asset pricing anomalies, derived from 111 original papers published in prestigious finance journals, can be attributed to random noise. These anomalies failed to pass the initial test criterion of $|t| \geq 1.96$, and when a stricter multiple test criterion of 2.78 was applied, the failure rate remained drastically high.

Serra-Garcia and Gneezy (2021) document that scientific research findings that are potentially based on wrong models gain far more attention than robust reliable results. This finding, though subject to its own need for careful examination, has led the authors of the study to speculate that papers with more intriguing or attention-grabbing content may encounter less rigorous review processes by journal editors and reviewers. Once these papers are published, they tend to attract greater attention. It is worth noting that this selection bias is not entirely unexpected in leading finance journals. What is surprising, however, is that even top-tier journals are known for publishing studies that cannot be replicated in extended samples.

As a case in point, a world-wide known stream of documented literature, such as Jegadeesh and Titman's (1993) seminal paper on momentum strategies, attracted ample attention from finance scholars and the finance industry. Wiest (2022) postulates that 47 out of 60 momentum-related studies managed to get published in top-tier journals, such as the Journal of Financial Economics. Kelly et al. (2021) argue that scholars tried to come up with a span of explanations with the aim of rationalising 'momentum premium', which may be defined as the tendency for stocks with a recent strong performance—often termed 'winner stocks'—to yield statistically significant higher returns compared to stocks with a recent poor performance, known as 'loser stocks'. In fact, scholars adhere to the notion that this strategy

consistently yields a positive monthly return of 1%, supported by a representative t-statistic signifying statistical significance at the 5% level (Grobys, 2022).

In contrast to such a view, Daniel and Moskowitz (2016) establish the fact that positive pay offs and high Sharpe ratios stemming from momentum strategies are oftentimes disrupted by a stream of persistent negative returns termed as 'crashes'. Conforming to these authors, this strategy generated a staggering negative return of -40% in just 1 month (March of 2009) during the 2008-2009 financial crisis, followed by another exuberant dip of -46% in April of 2009 (Grobys, 2023).

As per Figure 1, Grobys (2022) plotted the compounded returns of the momentum strategy spanning the period between January 1965 and April 2022. The darker curve represents the full time series evolution, whereas the lighter curve is a representation of a time series evolution where the most exuberant observations were purposely excluded. As can be deduced from Figure 1, those deleted 1% of observations that deviated vigorously from the rest account for 90% of the total compounded returns.

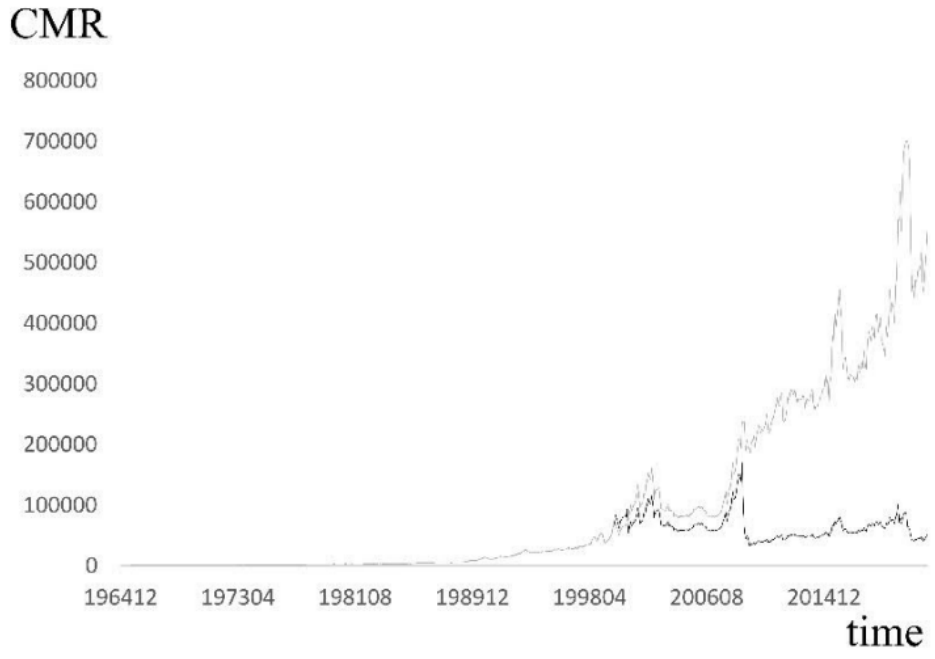


Figure 1: Time-series evolutions of the momentum strategy between January 1965 and April 2022.

While so many scholars and researchers cite the momentum strategy from Jegadeesh and Titman (1993), very rarely has there been a quantification study of the implied associated risk of following such a strategy in the real world. Hence, the question arises as to what contributes to scientific research failing to effectively replicate.

According to the findings of Hou et al. (2018), authors tend to perform what is known as a ‘specification search’. This involves repeatedly adjusting the sample criteria and testing methods until insignificant outcomes achieve statistical significance. This practice, often referred to as ‘p-hacking’ or ‘data mining’, may ultimately lead to an alarming increase in the number of false positive findings that cannot be reproduced in subsequent research efforts.

Schwert (2002) notes that, when cross-sectional asset pricing phenomena are initially observed and scrutinised in academic literature, these patterns often tend to dissipate, reverse, or weaken over time. Specifically, the diminishing t-statistics as sample sizes increase across various anomalies may be indicative of systematically inflated t-statistics. Schwert highlights the possibility that these asset pricing anomalies might have been mere statistical anomalies that initially captured the interest of scholars and professionals. This study contends that these 'statistical anomalies', as described by Schwert, stem from researchers relying on conventional research methods that may not be well-suited for financial research due to the unique nature of financial market data.

Similarly, Grobys (2021) argues that there could be an alternative explanation for the challenges in replicating studies. In particular, if there is an absence of the second moment or variance in the data generating process, it is inevitable that a t-statistic will be sample-specific. The theoretical mean would need one million observations until the theoretical mean converges to something reasonable. Given that the t-statistic is a function of the variance, one would observe a different variance in every sample.

Financial market research builds its underpinnings on the foundations of t-statistics to assess whether research results are, statistically, significantly different from zero or not. Normally, estimated sample t-statistics are computed as per Equation 1.

$$t = \frac{\bar{x} - \mu_x}{\sqrt{\frac{\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2}{\sqrt{T}}}}$$

Equation 1: Estimated sample t-statistic.

where:

$\bar{\mathbf{x}}$ = sample observation which is a normally distributed random variable with

E [$\bar{\mathrm{x}}$] = $\boldsymbol{\mu_x}$, VAR [$\bar{\mathrm{x}}$]= $\boldsymbol{\sigma_x^2}$;

$\boldsymbol{\mu_x}$ = standardised sample mean which is equal to 0;

$\boldsymbol{\sigma_x^2}$ = sample standard deviation

In statistics, any random variable following a normal distribution can be converted into the standard normal distribution. To achieve this, one must account for the following steps for each sample observation; viz., subtract it from the sample mean and then divide it by the sample standard deviation, as seen in Equation 1. This can be done for any distributed random variable.

When this transformation is applied, the standardised random variable will have an average (mean) of 0, and its distribution will correspond to the squared standard normal distribution. In statistical terms, this squared standard normal distribution is known as a 'chi-squared distribution' or (χ^2) with 1 degree of freedom (d.o.f). This is how it should be if Gaussian assumptions hold and if the standard financial theory holds. Such χ^2 distribution with 1 d.o.f. has a clearly defined mean, variance, and higher moments of statistics.

Accordingly, since the theoretical mean (first moment) and the theoretical variance (second moment) are clearly defined, the denominator in Equation 1 will converge towards the true population value, given that the sample is large enough. More importantly, statistical convergence depends on the fundamental existence of *variance* of the random variable. This means that the variance will not 'explode' and it will be operable because it will not be sample specific (Grobys, 2023).

This is crucial because x is a thin-tailed process requiring the presumption that the variance must be finite. This fundamental idea requires two additional pillar stones: i) the Central Limit Theorem (CLT), and ii) the Law of Large Numbers (LLN) must hold at all times. Hence, if x is heavily tailed, or a fat-tailed process, the variance would not converge. Taleb (2020) argues that in, the non-Gaussian world, the LLN would not even work, or it would work too slowly, as one would require too many sample observations for these two integral theories to run smoothly. Thus, even if the theoretical variance would still be computable, this variance would change with each different sample, leading to distinctive estimated t-statistics each time. This would largely pervert results and their true statistical significance.

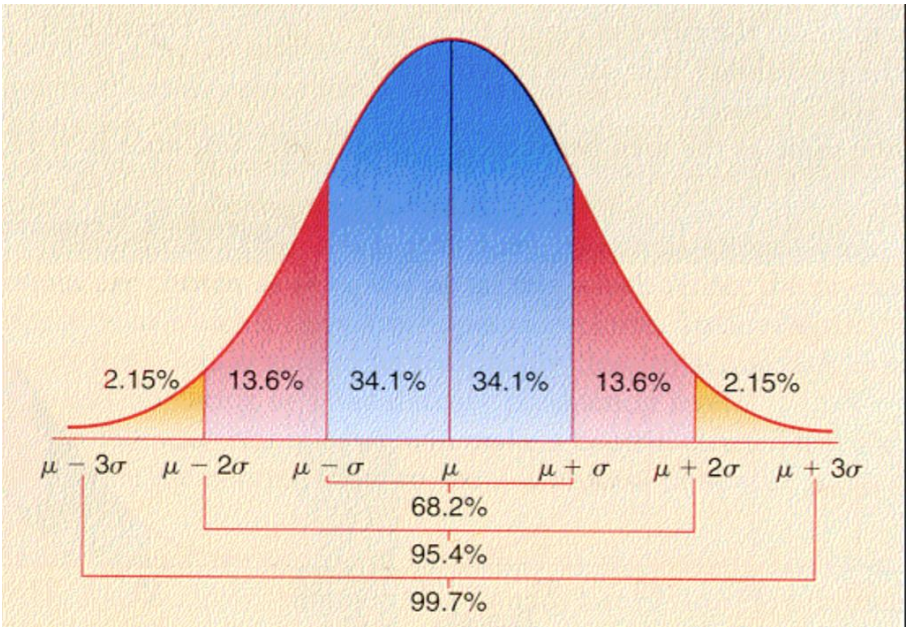


Figure 2: Standard normal distribution.

Figure 2 illustrates the standard normal bell-curve distribution, which, according to French Mathematician Bachelier (1900), is the underlying distribution of financial returns. As per Figure 2, around 68% of observations fall within one standard deviation from the mean, approximately 95% fall within two standard deviations, and nearly 98% fall within three standard deviations. Beyond the three standard deviations, the likelihood of observing data decreases significantly, and moving further away from the mean, such as from one to two standard deviations, results in only 27% of observations, and from two to three standard deviations, a mere 3% of observations.

In practical terms, this suggests that, according to the standard normal distribution, extremely rare events, like 6-sigma events, are virtually impossible to occur near the mean. In a simulation experiment based on 10,000,000 sample drawings from the standard normal distribution, Grobys (2022) inferred that the sum of the top 1% of observations only accounted for 3.6% of the cumulative total. This fact implies that the Gaussian models suggest that huge deviations from the mean have little impact on the cumulative total, and by extension, extremely fat-tailed events are practically impossible to observe in nature.

However, this contradicts that which was observed in real-life financial events, such as the stock market crash of October 1987 (Black Monday), where the S&P 500 index lost a substantial portion of its value. This event, as described by Mandelbrot and Hudson (2008), corresponded to a 22-sigma event, which has extraordinarily low odds, far beyond what is expected in natural occurrences. Complementary to this, in his lecture *Questionable Research in Finance: How Much Worth is a Nobel Prize in Economics?*, Grobys (2022a) explains that the norm in business school is for students to be taught about models using the continuous time 'Brownian motion' to derive their theoretical framework. Brownian motion incorporates three further assumptions; viz., that price changes are independent, stationary, and normally distributed. Yet, as seen in Figure 1—the momentum effect—and as per this documented historical 22-sigma event, 'discontinuity' is manifested in stock market behaviour. Mandelbrot and Hudson (2008) define discontinuity as events that have exceeded 6-sigma events.

Taleb (2020) nicknames the phenomenon of people relying on shaky Gaussian assumptions as the 'thick-tailed' or 'fat-tailed' problem. In his book, *Statistical Consequences of Fat Tails*, Taleb defines this "mother of all problems" as a small number of observations that rarely happen, but when they do, they commend much greater impact on the total statistical properties. So, in a given data set, these small number of observations represent the bulk of statistical properties.

According to both Taleb (2020) and Grobys (2021), there is a plethora of documented empirical evidence that shows that financial market returns do not behave according to the Gaussian-known stylized fact Despite this awareness, a significant portion of finance researchers, including well-known Nobel laureates like Eugene Fama—as discussed in section 2.2—persisted in using and promoting statistical models and techniques that assume Normality. These techniques include Ordinary Least Squares (OLS), Generalized Method of Moments (GMM), and GARCH-type models, all of which rely on the assumption that finite variance exists.

As shall be discussed in sub-section 2.3, there are many statistical consequences derived from the reliance on models and strategies published in popular academic journals that have instilled extravagant negative consequences in the real finance world.

2.2 Traditional finance teachings from Nobel laureates

Building upon the contribution and foundation of the work of Bachelier (1900), Markowitz (1952) published the reference paper, *Portfolio Selection*. Markowitz is known for his 'Modern Portfolio theory', as he studied the effect of asset risk, return, correlation, and diversification on probable investment portfolio returns. All such metrics were designed for normally distributed random variables.

Markowitz (1952) showed us that the core of the portfolio theory lies in the need for an objective criterion to assess and value risk, defined as the measure of dispersion from the central line of tendency, the mean, in a normal distribution space. The proxy for risk is the

standard deviation metric (σ^2). The portfolio theory shows us that, the more securities we add in a portfolio, the more we diversify away ´idiosyncratic risk´, which by definition, is that diversifiable risk pertaining to the individual security (Bodie et al., 2021).

Nonetheless, the lingering question persists: how can one effectively select stocks to optimise the benefits of diversification? The answer lies in the concept of the 'correlation coefficient'. According to Bodie et al. (2021), equities are classified as "perfectly positively correlated securities" with a correlation coefficient of +1, and as "perfectly negatively correlated securities" with a correlation coefficient of -1. As long as the correlation between securities is not absolutely positive, indicating a positive but not perfect correlation, incorporating these securities into a portfolio has the potential to result in gains. While a weaker correlation is more advantageous for risk reduction, even a slight improvement in correlation can be advantageous. It is, nevertheless, essential to reminisce that diversification is based on the assumption that the underlying distribution of individual security returns is the normal Gaussian distribution.

In the realm of portfolio return assessment, the approach involves combining the expected returns of multiple securities within the portfolio. The primary focus is on determining the anticipated return of the entire portfolio. This is achieved by applying a straightforward weighted average calculation, as per Equation 2. By multiplying each security's expected return by its respective weight—probability or allocation—one can derive the portfolio's expected return. The collective sum of these calculations results in the average return for the portfolio, which, despite fluctuations in individual security returns, provides an overall measure of anticipated return.

$$\mathrm{E(r)} = \sum w_i k_i$$

Equation 2: Portfolio returns.

The variance of a portfolio's return can be divided into different parts to understand its risk, as per Equation 3.

$$\sigma^2{}_p = (w_1{}^2\,\sigma_1{}^2) + (w_2{}^2\,\sigma_2{}^2) + 2\,(w_1)\,(w_2)(\rho_{1,2}\,\sigma_1\,\sigma_2)$$

Equation 3: Portfolio risk.

Where:

$\mathbf{w_1}$ = weight of security A

$\mathbf{w_2}$ = weight of security B

$\boldsymbol{\sigma_1^2}$ = portfolio variance for security A

$\boldsymbol{\sigma_2^2}$ = portfolio variance for security B

$\boldsymbol{\rho_{1,2}}$ = correlation coefficient between the returns of the two securities

To break down Equation 3, the variance of the portfolio's return is based on three main factors: the expected variance of Security A, the expected variance of Security B, and the covariance between Security A and Security B.

The variance of Security A measures how much its returns differ from the average, showing its inherent risk. The variance of Security B does the same for the second security in the portfolio. The covariance between Security A and Security B indicates how much their prices move together, reflecting their relationship and capturing systematic risk. This systematic risk, often called Beta, is a key factor for which investors anticipate compensation. It is not something that diversification can eliminate. It is important to analyse the variance of the portfolio to understand the impact of market-wide movements and the specific movements of the securities in the portfolio. It is also important to recollect that this model rests on the

assumption that returns are normally distributed and that covariances accurately capture the relationship between assets, which may not always be the case.

This Markowitz model can be extended to many assets and not just two securities, as per Equation 4. Markowitz's model corresponds to the initial step in portfolio management, which involves pinpointing the efficient array of portfolios, commonly known as the 'efficient frontier' of risky assets. Bodie et al. (2021) disseminate Markowitz's modern portfolio theory into three steps:

Step 1: The efficient frontier begins with the global minimum variance portfolio, offering the lowest risk and a balance between risk and return among investment options.

Step 2: To locate the optimal portfolio on the efficient frontier, we use the capital allocation line with its slope representing what is called the 'Sharpe ratio'. In finance, a fundamental principle asserts that, on average, a higher level of risk should be associated with a correspondingly higher level of return. In parallel, Sharpe (1964) is another Nobel laureate who effectively combined both risk and return into a single metric, known as the "risk-adjusted return measure", which is the renowned Sharpe ratio (see Equation 4). The higher the Sharpe ratio, the more favourable the investment opportunity. The array of potential portfolio combinations is distilled to one point, commonly known as the 'optimal risky portfolio'.

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$$

Equation 4: Sharpe ratio.

Where:

$\mathbf{R_p}$ = Return of portfolio

$\mathbf{R_f}$ = Risk-free rate

$\boldsymbol{\sigma_p}$ = Standard deviation of portfolio's excess return

This Sharpe ratio also contributed to the development of methods for the valuation of options and a returns-based style analysis for evaluating the performance of investment funds.

Step 3: Using Markowitz's portfolio theory (see Figure 3), Sharpe (1964) extended the optimal allocation line by introducing the market portfolio, which maximises the Sharpe ratio. He introduced the risk-free rate to ultimately yield the maximum reward to risk ratio, provided that the investors are rational and would only invest in the market portfolio. Creating the optimal complete portfolio involves combining the optimal risky portfolio with the risk-free rate while considering the investor's risk aversion. This is done by selecting an indifference curve tangent to the capital allocation line.

Once more, it is imperative to remember that the concept of the efficient frontier and the risk-free rate is based on assumptions regarding normally distributed returns and correlations, which may not consistently mirror real-world circumstances.

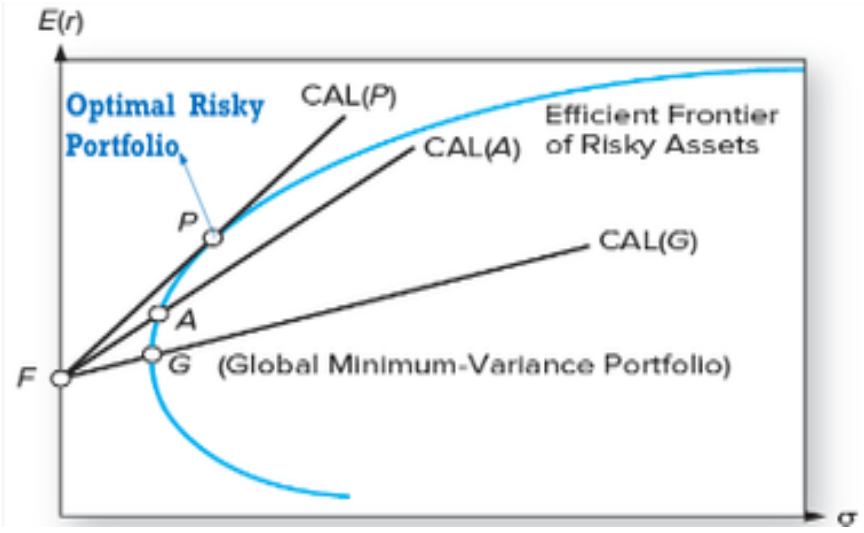


Figure 3: Markowitz's Modern Portfolio Theory.

Sharpe (1964) is also known for his contribution towards the asset pricing model, known as the 'capital asset pricing model', which is abbreviated as CAPM.

$$E(R_i) = r_f + \beta_i(E(R_m) - r_f$$

Equation 5: CAPM equation.

Where:

$\mathbf{E(R_i)} =$ Expected rate of return of security i

$\mathbf{r_f} =$ risk-free rate

$\boldsymbol{\beta_i} =$ Market-risk factor (sensitivity)

$\mathbf{E(R_m)} =$ expected return of the market

As per Equation 5, the CAPM is the expected rate of return of a security or an equity, over and above the risk-free rate, taking Beta into consideration.

Within the framework of the CAPM, the key innovation lies in the introduction of the concept of Beta. Instead of laboriously calculating correlation coefficients for every stock, the focus shifts to understanding how the returns of individual assets, denoted as 'A' and 'B', respond to market movements. Beta serves as a singular, more efficient measure that captures this relationship. It quantifies the extent to which the returns of individual stocks move in concert with the overall market, particularly when the market rises or falls.

It is crucial to emphasise once more that the CAPM model is predicated on the assumption that Beta precisely gauges the correlation between the security and the market, as well as the presumption that market returns follow a normal distribution, which may not invariably hold true.

The Fama and French three-factor model in Equation 6 provides a more comprehensive framework for understanding asset pricing when compared to the single-factor CAPM. It suggests that not only the market factor but also the size and value factors play significant roles in explaining stock returns.

$$E(R_i) = r_f + \beta_1(R_{Mt} - r_f) + \beta_2(SMB) + \beta_3(HML)$$

Equation 6: The Fama and French three-factor model.

With regards to the size factor, the SMB (Small Minus Big) factor represents the difference in returns between small-cap stocks and large-cap stocks. It is a measure of the size effect in the market. When it comes to the value factor, the HML (High Minus Low) factor captures the difference in returns between stocks, with high book-to-market (value stocks) and low book-to-market (growth stocks) ratios. It measures the value effect in the market.

The model assumes that returns follow a normal distribution. If actual returns deviate significantly from normality, it can lead to misestimation of risks and potential inaccuracies in expected returns.

Engle's (1982) ground-breaking work in the field of financial time series analysis has received widespread recognition and reshaped the way we model and analyse financial asset returns. The time-varying volatility, also known as autoregressive conditional heteroskedasticity (ARCH) models, address a core phenomenon observed in financial markets: the clustering of returns. In these models, we recognise that financial asset returns exhibit periods of both low and high volatility, creating a dynamic pattern in the data. Engle's pioneering insight led to the development of the ARCH model, which seeks to capture and quantify this volatility clustering, as shown in Figure 4.

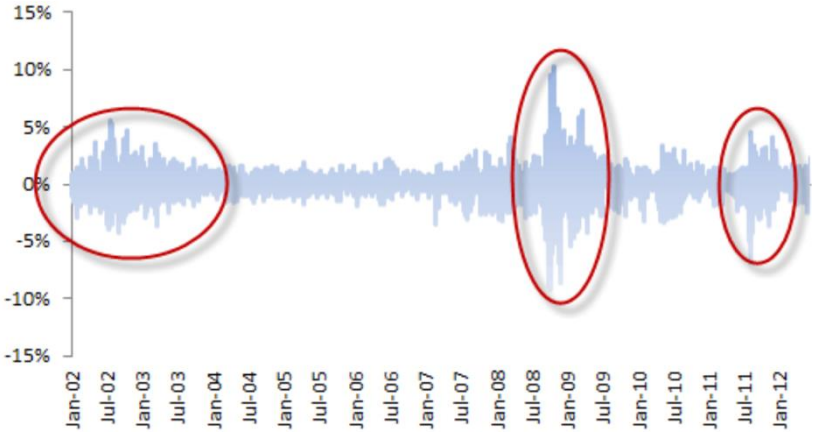


Figure 4: Volatility clustering.

At its core, an ARCH model consists of a mean term, μ, which represents the expected return, and a lagged value of the financial asset, capturing past returns. Additionally, there is an error term, ε, that introduces randomness into the model. This error term accounts for potential autocorrelation in financial asset returns, which is evident in market data.

The crucial component in modelling volatility clustering is the conditional standard deviation, denoted as σ_t, which represents the time-varying volatility of the financial asset. It is linked to the error term through the formula in Equation 7.

$$U_t = \sigma_t \, \varepsilon_t$$

Equation 7: Stochastic process.

Where: ε_t follows a normal distribution with mean (μ) zero and variance (σ^2_t) one.

Estimating the parameters in this model is typically done using the maximum likelihood estimation (MLE). However, because the error terms are not assumed to be normally distributed, this approach is referred to as quasi-maximum likelihood estimation.

The ARCH model extends to include a lag structure. The conditional standard deviation, σ_t, is determined by an intercept term (w_0) and a set of parameters (α_i), where i ranges from 1 to

p, with p representing the order of the model. These parameters capture the square root of the lagged squared residuals from previous time periods, allowing the model to adapt to changes in volatility over time. This can be expressed as shown in Equation 8.

$$\sigma_t = \sqrt{w_0 + \alpha_1 \varepsilon^2{}_{t-1} + \alpha_1 \varepsilon^2{}_{t-2} + \ldots + \alpha_p \varepsilon^2{}_{t-p}}$$

Equation 8: Conditional standard deviation.

One property of the ARCH model is that the conditional expectation of the error term, ε_t, must be zero. This property aligns with the efficient market hypothesis, suggesting that any predictable information has already been incorporated into asset prices.

Despite its elegance, the original ARCH model makes the most problematic assumption, once again, that ε_t follows normality. This has prompted the development of various extensions, such as the GARCH models, which account for the infinite past by summarising the lagged squared residuals and offer more flexibility in modelling the innovation process. Researchers also extended to t-GARCH models, for instance, to account for the fact that negative realisations in the innovation process have a higher impact than positive ones. Thus, they even added different parametrisations of the same original model. Researchers have also explored alternative distributions, like the t-distribution, or generalised error distributions to improve model performance. Therefore, literature sought to introduce various 'stylised fixes' in an attempt to refine and enhance the models.

2.3 Consequences of using the wrong models

Hou et al. (2018) emphasise that a multitude of academics and investment managers are actively pursuing significant anomalies. With trillions of dollars invested in factor-based, exchange-traded funds and quantitative hedge funds globally, there is an overwhelming financial interest, but it also comes with an equally substantial hidden risk.

Referring to the bankruptcy of the hedge fund LTCM, whose founding partners included Merton and Scholes—two Nobel laureates—Taleb (2020) highlights that relying on incorrect methods could potentially lead to very serious consequences. Taleb goes on to describe this project as people blindly trusting the models of Nobel laureates.

Scholes and Merton won the Nobel Prize for their work in the Black-Scholes-Merton option pricing theory, which relied heavily on the presumption of 'continuous-time finance'. The Black-Scholes (1973) model is used to calculate the value of a European call option. This value is determined by considering five key variables:

1. The current value of the underlying asset (S).
2. The strike price of the option (X).
3. The time to expiration of the option (T).
4. The risk-free interest rate (r).
5. The variance in the value of the underlying asset, often denoted as sigma squared (σ^2).

In the original Black-Scholes (1973) formula, the value of a call option is expressed in the ways shown in Equations 9, 10, and 11.

$$C_0 = S_0 N(d_1) - X e^{-rT} N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S}{X}\right) + (r + (\sigma^2)/2) * t}{\sigma\sqrt{t}}$$

$$d_2 = d_1 - \sigma\sqrt{t}$$

Equations 9, 10, and 11: The Black-Scholes-Merton option pricing model (1973) for a European call option.

The Black-Scholes (1973) model inherently incorporates the concept of a replicating portfolio. In this portfolio, one borrows money to purchase a specific number of shares. In the Black-Scholes formula, the term $\mathrm{X}e^{-\mathrm{rT}}$ represents the borrowed amount, while $N(d_1)$ is the quantity of underlying assets bought. This quantity is often referred to as the 'option delta'. In Equation 5, the term $e^{-\mathrm{rT}}$ represents a present value factor, indicating that the strike price X is adjusted to its present value because the payment is made at the option's expiration. The terms $N(d_1)$ and $N(d_2)$ involve cumulative normal distribution functions, which essentially represent probabilities, yielding values between -1 and +1.

To compute $N(d_1)$ and $N(d_2)$, one needs to input all the variables that affect the option value, including S, X, r, and σ, as seen in Equations 9 and 10. These variables are used to calculate d_1 and d_2, which are then transformed into $N(d_1)$ and $N(d_2)$ using the normal distribution. Thus, as evidenced in the above-mentioned formulae, the whole option pricing model is, once again, based on the assumption of the normal distribution of the pay offs.

In essence, the founders managed to pool a number of investors with top-notch resumes and convinced them of their sophisticated computations. Ultimately, events took a twisted turn in the summer of 1998, when a combination of events instilled by the ongoing Russian financial crisis was not captured by the so-called 'sophisticated models' designed by the owners of this hedge fund. Ultimately, LTCM went bust, and the exposures were so significant that they almost took down the whole financial system. As always, it was the investors who took the hit by falsely relying on the competence of Nobel Prize winners, and they lost a substantial portion of their own wealth (Grobys, 2023).

Grobys (2023) also highlights another recent incident involving a hedge fund facing considerable challenges. In March 2021, during the COVID-19 pandemic, a company called Greensill Capital, which had previously been a prominent player in the field of supply-chain finance, encountered a major crisis. Greensill's primary insurer declined to renew a substantial $4.6 billion contract, and the prominent bank Credit Suisse froze an extensive $10 billion in assets. Subsequently, Credit Suisse incurred a significant loss of $4.7 billion due to

the default of another hedge fund named Archegos Capital Management, based in the United States.

Grobys (2023) utilises the concept of a "master and disciples" archetype to describe the fact that the individuals managing the assets of both Greensill and Archegos were educated in the traditional normative theory, which aligns with the teachings of figures like Merton and Scholes. He posits that variance, a standard tool for assessing financial risk, becomes problematic in this context because the theoretical variance remains undefined due to the extreme price fluctuations that are inherent in financial markets. Consequently, the concept of risk becomes virtually infinite.

Grobys (2023) acknowledges that people do calculate the sample variance for datasets containing numerical values. However, he asserts that this variance loses its meaning when it is not defined within the theoretical underlying distribution. This issue circles back to the earlier problem of discontinuity in the return process of financial assets, which is characterised by a small number of significant deviations accounting for a substantial portion of the cumulative compounded returns. As a statistical consequence of this phenomenon, variance cannot be effectively defined for such a process, leading to research results that are specific to the dataset being analysed.

3 Literature review

3.1 Fractality: Modelling discontinuity and statistical self-similarity

> Novel ideas that evoked concepts of fractals, were viewed fascinating formal extensions of academic mathematics but were not generally perceived as playing any significant role in the real world. It fell to French mathematician Benoit Mandelbrot to make the crucial insight that, quite to the contrary, crinkliness, discontinuity, roughness and self-similarity – in a word, fractality – are, in fact, ubiquitous features of the complex world we live in. (West, 2017, p. 130)

Mandelbrot dedicated his life to exploring the intricacies of nature, emphasising its rugged surfaces and irregular forms—embracing gaps, sharp contours, edges, and spikes that epitomise the inherent chaos in everyday phenomena. Coined from the Latin term 'fractus', meaning 'broken', Mandelbrot introduced the term 'fractals' to capture nature's authentic manifestation. His goal was to formulate a comprehensive fractal theory, offering a framework for explanatory models. Recognising the profound importance of his work in unveiling nature's complexities, Mandelbrot anticipated a lasting impact on scientific development. This exploration into fractals spans diverse scientific domains, as highlighted by Olsen in his publication titled *Financial Markets: From Fractals to Power Laws* (2023).

Mandelbrot's influence resonates across disciplines, inspiring advancements in statistical physics, meteorology, anatomy, neurology, linguistics, engineering, computer graphics, and the social sciences. In his 1963 paper, *New Methods in Statistical Economics*, published in the Journal of Political Economy, Mandelbrot introduced power laws as a solution to the issue of extreme events.

Power laws essentially resemble ‘Pareto-type’ distributions, allowing for infrequent but significant deviations from the mean. The revelation of power laws marked a significant breakthrough, suggesting that, when examining a phenomenon solely at one scale, its irregularities obscure its inherent self-similarity. The gaps, rough edges, and sharp corners—although visually distinct—belie the fact that the statistical properties remain consistent across different scales. A prime illustration is the coastline, where regardless of the

observation scale, the shape exhibits analogous characteristics, making it challenging to discern whether it spans one mile, a hundred miles, or a thousand miles based solely on its form. In this context, Taleb (2020, p. 91) contends that "there are a lot of theories on why things should be power laws, as sort of exceptions to the way things work probabilistically. But it seems that the opposite idea is never presented: power laws should be the norm, and the Gaussian a special case."

Notably, Pareto (1897) had earlier formalised this concept. The distributional principle that governs the distribution of wealth, commonly known as the 'Pareto Law' or '80/20 rule', was introduced much earlier in 1897 in the same journal. The Pareto distribution in its original form postulates that 20% of the largest observations make up 80% of the overall wealth in the population. In a Gaussian world, 20% of the largest observations make up only 44% of the overall total wealth in the population.

Significantly, the Pareto distribution demonstrates scalability, in contrast to the normal distribution, which lacks scalability by definition. The scalability becomes evident when applying the 80-20 principle iteratively, resulting in a ratio of 50/1. For instance, if 20% of the population owns 80% of the wealth, then 20% of this 20% owns 80% of the 80%. In practical terms, this signifies that 4% of the population possesses 64% of the wealth. Through repeated iterations, another 20% of this 4% owns 80% of the 64%, indicating that 1% of the population owns 50% of the wealth. This characteristic is commonly termed fractality or statistical self-similarity, while the distribution itself is often labeled as a power law. These observations are shown in Table 1.

Table 1: Gaussian versus Pareto.

% of largest Observations	**% of the total Gaussian**	**% of the total Pareto**
20%	44%	80%
4%	12%	64%
1%	3.6%	50%

It is crucial to highlight that this distribution type displays distinct characteristics when compared to the normal distribution. In contrast to the normal distribution, where it can be demonstrated that the top 1% of observations account for only 3.6% of the cumulative total, this distribution type follows a markedly different pattern.

The importance of power laws becomes evident in chaotic systems, particularly in financial markets, which share a nature-like unpredictability. Establishing self-similarity within these systems unveils power laws that can lay the groundwork for a new economic theory. Mandelbrot, in his pioneering 1963 paper entitled *The Variation of Certain Speculative Prices*, delved into the analysis of financial markets, specifically identifying self-similarity in cotton prices. He was the first to show that cotton price changes are governed by a Paretian tail; that is, a power law. Mandelbrot showed that $\alpha < 3$ has serious implications.

When observing the four graphs in Figure 5, stemming from Grobys's (2023) lecture on *Modeling Discontinuity and Fractality For Financial Assets,* without disclosing their scales and focusing solely on their patterns, it becomes apparent that distinguishing between daily and monthly data is challenging. The patterns exhibit striking similarities, making it impossible to discern the resolution by mere visual examination. Despite the monthly data displaying a considerably broader range than the daily data, the visual patterns remain indistinguishable. Thus, the graphs manifest a form of self-similarity where, irrespective of the data resolution (frequency), the visual representation appears consistent.

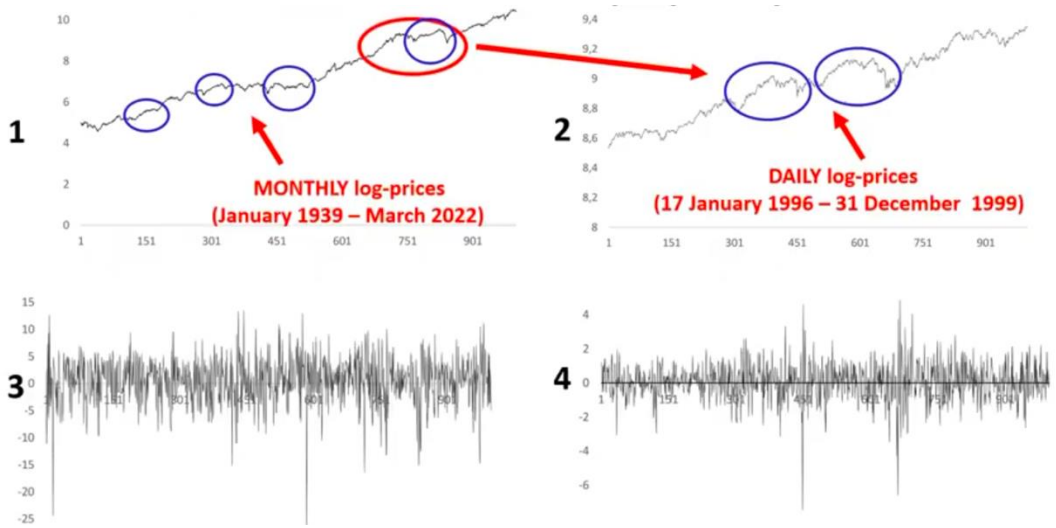


Figure 5: Self-similar patterns in financial data (Grobys, 2023).

Concurrently, in 1963, Mandelbrot presented a paper suggesting the application of stable Paretian distributions to depict the fluctuation of asset returns. Fama (1963) subsequently examined Mandelbrot's proposal and offered commentary, stating:

> The infinite variance assumption of the stable Paretian model has extreme implications. From a purely statistical standpoint, if the population variance of the distribution of first differences is infinite, the sample variance is probably a meaningless measure of dispersion. Moreover, if the variance is infinite, other statistical tools (e.g., least-squares regression) which are based on the assumption of finite variance will, at best, be considerably weakened and may in fact give very misleading answers. (p. 421)

Statistical inference in finance research, including methodologies like OLS regressions or GARCH models, is usually formulated based on Gaussian assumptions. Through a straightforward simulation experiment, it becomes evident that, in a Gaussian context, the cumulative impact of the largest 1% observations is merely 3.6% of the total. In this Gaussian world, significant deviations from the mean have a relatively minor impact, and extremely large deviations are nearly impossible to observe. For example, the likelihood of witnessing a 3-, 4-, or 5-sigma event is 1 in 740, 1 in 32,000, or 1 in 3,500,000, respectively. The odds of deviation decrease exponentially, as Taleb (2010) contended, justifying the dismissal of outliers due to the rapid decline in the probability of detecting a signal. Given the recurrent occurrence of extreme events in financial markets, the central research challenge is how to address and mitigate their impact in financial research (Grobys, 2022).

Paradoxically, despite being aware of the issues associated with financial analysis rooted in Gaussian frameworks, the majority of finance researchers—including Fama, who was supervised by Mandelbrot in his doctoral thesis—continued to utilise techniques like OLS in a significant portion of scientific studies published in leading finance journals. Another stream of literature aligned with Mandelbrot's concepts emerged, incorporating power laws into financial research. Frequently referenced studies in this domain include Gopikrishna, Plerou, Amaral, Meyer, and Stanley (1998), Jansen and de Vries (1991), Mantegna and Stanley (1995), and Lux (1996), among others.

One potential explanation for leading finance scholars, such as Fama, not adopting Mandelbrot's novel methods in financial research, despite being well-versed in the issue, could be the increasing acceptance of more readily applicable alternatives. For instance, Markowitz's (1952) modern portfolio theory, based on a Gaussian framework, appealed to finance scholars, even though its assumptions do not accurately reflect the reality of financial markets. It is noteworthy that Markowitz's study was published a decade before Mandelbrot's (1963) contributions. Incorporating Mandelbrot's ideas might have had potentially devastating consequences for the finance community by nullifying a decade of subsequent research built on Markowitz's (1952) foundation.

Could it be plausible that the high rate of replication failures in financial studies is a manifestation of researchers relying on Gaussian methodologies that, in the presence of sample-specific variances, yield highly misleading results, as noted in Fama's (1963) early paper?

3.2 Perspectives on Mandelbrot's fractal finance concepts

3.2.1 Overview of the current body of literature addressing power law dynamics in financial markets and its significance

To explain the staggering failure in scientific research to replicate, Grobys (2022) inferred that, if the variance of the data-generating process does not exist, a t-statistic will invariably be specific to the sample. In his ground-breaking paper, *What do we know about the second moment of financial markets?*, he investigated the stability of second moments for five key financial market variables. Such a study extended this novel stream of literature by (i) examining whether the variances of five key financial market variables follow power laws and (ii) identifying the existence of second moments of the variances. Utilising a research approach grounded in realised variance (RV), a financial market variable is deemed stable only if the variance of the variance exists. Given the heavy fat-tailed processes of RVs, this study aligns with recent literature by fitting power law distributions to the RVs of specified key financial markets, including equities, commodities, currencies, and cryptocurrencies. This methodology finds support in the work of Wang and Yang (2009), who observed that:

> Realised volatility based on intraday quotes is a consistent and highly efficient estimator of the underlying true volatility . . . a realised-volatility-based approach is able to uncover volatility features (asymmetric volatility in particular) that conventional GARCH type models fail to reveal. (p. 600)

To test the plausibility of the power law null hypothesis, hypothesis tests based on Kolmogorov-Smirnov distances were employed, following the approach proposed by Clauset et al. (2009). Various subsamples, different data frequencies, and simulation experiments were also considered. The findings revealed that the daily variances of all five key asset markets follow power law processes, with the power law null hypothesis statistically resistant to being rejected. Notably, the results were neither sample- nor method-specific. Remarkably, the research indicated that the variance of the variance statistically does not exist for any of these asset markets. Paradoxically, the foreign exchange market was found to be more susceptible to extreme events than the Bitcoin market. The implications of these findings are

fundamental, suggesting that, due to the non-existence of variances' variances, standard statistical analysis based on OLS or GMM may yield sample-specific results.

Furthermore, from a broader perspective, Grobys (2021) contributed to the literature on tail risks associated with human-engineered systems. Building upon the work of Clauset et al. (2009), which explored power law distributions in 24 real-world data sets across disciplines, this study supported Taleb's (2020) assertion that power law distributions govern many real-world phenomena. Gabaix (2009) further contributed by documenting power law processes in assorted variables, such as income, wealth, city and firm sizes, trading volume, international trade, and executive pay.

Recent financial studies, including Grobys (2021) on the RV of asset markets, and Grobys et al. (2021) on the volatility processes of cryptocurrencies, have applied power law functions. Grobys et al. (2021) contributed to this literature by adopting a fractal perspective on Bitcoin hackings. Empirical findings revealed that estimated losses from hackings significantly differ from basic statistics. Grobys et al. (2021) demonstrated that over 80% of the cumulative 20% of hacking incidents with the largest Bitcoin losses contribute to the overall distribution, strongly suggesting a non-Gaussian process. Using a recently proposed test based on Bayes' rule, the authors eliminated other commonly used distributions as potential candidates governing cyberattacks in the Bitcoin market. Given the distribution's close resemblance to the Pareto 80/20 distribution, they modelled cyberattacks as a fractal process.

In the same paper, the employed methodology of the MLE indicated a power law process similar to a fractal process with no defined theoretical mean, implying infinite loss (t-statistic 7.96). Given Bitcoin's limited supply, the authors posited that the mean of lost coins is finite, a hypothesis strongly supported by our statistical tests. Estimating the exponent of the power law process through the MLE, the authors computed a shadow mean. Combining the shadow mean with the sample mean for data not governed by a fractal process, they calculated the overall expected loss at 106,171.49 coins, nearly four times higher than the naive sample average of 29,050.18 coins.

Hence, their findings affirmed that the majority of observations are inconsequential in computing expected losses, as the most significant statistical information lies in the tail of the distribution. The inadequacy of naive risk management in estimating expected losses due to cyberattacks in the Bitcoin market points to the urgent need for cryptocurrency market regulations by governments and regulatory agencies to protect investors from potentially severe losses.

3.2.2 Challenging the infinite variance hypothesis

Recent literature in finance has extensively employed power laws to model both financial returns and volatility. Lux and Alfarano (2016) conducted a comprehensive literature review, summarising influential studies in financial economics that utilise power laws. Despite Mandelbrot's (1963) assertion in his seminal study on cotton price changes that the variance in such changes is infinite, Lux and Alfarano (2016) pointed out that subsequent research, prompted by Mandelbrot's influential paper, has cast doubt on the validity of the infinite variance hypothesis. This scepticism is rooted in challenges to the stability-under-aggregation property of the estimates. Lux and Alfarano attempted to challenge the infinite variance hypothesis. In fact, they asserted that the power law exponents for the return generating processes of most asset classes exceed 2 and approach 3; thereby, rejecting the infinite variance hypothesis. However, Mandelbrot's (1963) infinite variance hypothesis is affirmed for the returns on venture capital and R&D investments (Lux & Alfarano, 2016).

On a different note, Taleb (2020) posited that if the fourth moment is unknown, the stability of the second moment cannot be ensured. This implies that, if $\alpha < 5$, working with the variance is untenable, even if it theoretically exists in the distribution. While the variance is undefined for $\alpha < 3$, $\alpha < 5$ implies that kurtosis is undefined, rendering the variance unstable (Taleb, 2020, p. 187f).

In his paper entitled, *When "story telling" gets published in scientific journals: On the momentum illusion*, Grobys (2022c) conducted specific tests to assess invariances across different time frequencies. The examination of invariances across distinct time frequencies

demonstrated that the invariance hypothesis stands, indicating that scaling behaviour persists despite a decrease in resolution. This discovery is noteworthy, as per the impact of time aggregation, one would anticipate power law exponents to grow in economic magnitude with decreasing resolution. To delve deeper into this matter, the power law exponents were subjected to regression analysis against their corresponding resolutions. Surprisingly, the regression results indicated that, as time frequency decreases, the power law exponent diminishes. This starkly contrasts with existing literature on the effects of time aggregation, suggesting that momentum strategies pose even greater risks over longer time horizons.

Furthermore, through sample split tests, it was revealed that the invariance hypothesis holds true even across time periods. The estimated power law exponents for subsequent subsamples consistently fell within the 95% confidence intervals derived from the estimated power law exponents and their associated estimation uncertainties based on earlier subsamples. This lends support to the idea that the observed results remain robust irrespective of the examined resolution, implying a persistent scaling behaviour over time.

Similarly, in his paper, *A multifractal model of asset (in)variances*, Grobys (2023a) addresses the issue relating to the validity of the invariance hypothesis through the calculation of weekly and monthly RVs, summing the squared returns over 5 and 20 consecutive, non-overlapping trading days. Additionally, he examined the models' ability to produce multifractality across different time dimensions through a comparative analysis of the weekly and monthly models. The results indicate that RVs based on the multifractal model exhibit Paretian tails and long-term dependencies. More importantly, regarding the invariance hypothesis conundrum, this study also finds support for the invariance hypothesis across different time scales, challenging the perception that variances are less volatile at lower frequencies. The evidence suggests a scale invariance of S&P 500 RVs, emphasising the importance of testing the invariance hypothesis in financial markets.

4 Data and methodology

4.1 Data collection

This study makes use of U.S. daily data on the closing price of the DJIA, spanning from 28th May 1896 to 10th February 2023, a total of 32,820 trading days.

The DJIA, also known as 'the Dow', is a prominent stock market index tracking the daily movements of 30 leading U.S. publicly-traded companies listed on the NASDAQ or New York Stock Exchange (NYSE). These 30 companies are considered key players in the U.S. economy, making the Dow a widely recognised benchmark for assessing overall market performance. Unlike the S&P 500 index, which encompasses 500 of the largest public companies and uses a market capitalisation-weighted approach, the DJIA is price-weighted. This means that price movements of higher-priced stocks have a more significant impact on the index, potentially leading to increased short-term volatility. Consequently, the Dow may exhibit greater sensitivity to small percentage changes in share prices, particularly for companies with smaller market capitalisations but higher stock prices.

The source of the raw data pertaining to the study is Stooq.com. Stooq, a Polish website with English translation, provides data for global securities, ETFs, currency pairs, and cryptocurrencies. Additionally, the platform offers data on global indices, commodities, and bonds. Similar to Yahoo Finance, users access all available data through a web interface, with the option to download data in CSV format. For stocks listed on US exchanges, some fundamental data such as price-earnings (PE) ratio and market value is accessible, though historical downloads for these fundamentals are not available. This data was used to construct a database in Microsoft Excel, which provided flexibility with respect to the handling of the data.

Utilising the DJIA to test the Lévy alpha-stable hypothesis is justified given several key considerations. The index's historical data, often spanning several decades, allows for a

comprehensive examination of long-term market behaviour. Additionally, the DJIA's focus on large-cap stocks aligns with the heavy-tailed nature of the Lévy alpha-stable distribution, making it a pertinent candidate for this analysis. The DJIA's stability and resilience in the face of economic shifts and global events also contribute to its suitability. By applying the Lévy alpha-stable hypothesis to the DJIA, researchers can gain valuable insights into the distribution of extreme price movements within a well-established and influential index, enhancing our understanding of the broader financial market dynamics.

4.2 Realised variances: The power law modelling rationale

4.2.1 Realised variances

> By sampling intraday returns sufficiently frequently, the realised volatility can be made arbitrarily close to the underlying integrated volatility, the integral of instantaneous volatility over the interval of interest, which is a natural volatility measure. Hence for practical purposes, we may treat volatility as observed, which enables us to examine its properties directly, using much simpler techniques than the complicated econometric models required when volatility is latent. (Andersen et al., 2001, p. 42)

Realised variance (RV), in the context of the DJIA or any financial time series, is a measure of the actual observed variability or volatility in the daily returns of the asset. It is a metric that quantifies the extent of price fluctuations over a specific time period based on the observed historical data. RV provides a more accurate representation of market volatility compared to traditional measures that rely on closing prices alone.

The computation of RV involves the squared daily returns of the asset over a specified time frame. In the case of the DJIA, this would entail squaring the daily percentage changes in the index's value. The formula for computing RV for a given time period 't' is shown in Equation 12.

$$RV_t = \sum_{j=1}^{N} R_{j,t}^2$$

Equation 12: Realised variances (returns) computation.

Where:

- RV_t is the realised variance for time period 't',
- $R_{j,t}$ represents the daily return on day 'j' within time period 't', and
- 'N' is the total number of non-overlapping observations in the specified time period.

In practical terms, calculating the RV for the DJIA involves squaring the daily percentage changes in the index value over the selected time interval and then summing up these squared returns. This approach captures the actual intraday fluctuations and provides a more accurate reflection of volatility, making it a valuable tool for risk assessment, option pricing, and other financial analyses.

$R_{j,t}$ denotes the daily return of the DJIA on day 'j in time unit t, with N falling within the range of (5, 20, 60, 90, 125). For example, N = 5 corresponds to weekly RV, while N = 125 corresponds to semi-annual RV. The calculation of various RVs is based on non-overlapping observations, ranging from t = 1 to 263 for semi-annual data and t = 1 to 6564 for weekly data.

Chapter 5 will provide a thorough exploration of the characteristics of realised variances at different time resolutions, encompassing various time frequencies; thus, ensuring a more comprehensive exploration of the data set.

4.2.2 Assessing power law behaviour in the Dow Jones index price returns

In the case of financial time series data like DJIA price returns, comprehending the underlying distribution is of utmost importance. Statistical insights into the data's behaviour form a vital foundation for subsequent analyses, guiding methodological choices. Before employing advanced modelling techniques, it is essential to preliminarily explore the data's

characteristics, determining whether it aligns with Gaussian expectations or exhibits a power law distribution. This initial inquiry ensures that analytical approaches are tailored to the dataset's dynamics, enhancing the robustness and validity of findings.

When utilising 'log-log' regressions under a power law assumption (see section 4.3), a critical aspect is justifying why such an assumption is made, highlighting the need for a thoughtful rationale behind the chosen analytical approach. Here, we proceed under the assumption that our data exhibits heavy tails, indicating a potential power law influence. Recognising this, we expect that a small percentage of extreme observations may significantly contribute to the dataset, possibly constituting 75% or 80% of the total. For instance, when examining the RVs from 5 daily squared returns (see Figure 6), the time series evolution suggests a process with an alpha exponent close to 2, marked by substantial outliers and discontinuities that exert significant influence. This emphasises the plain differences between extreme values and the majority of observations.

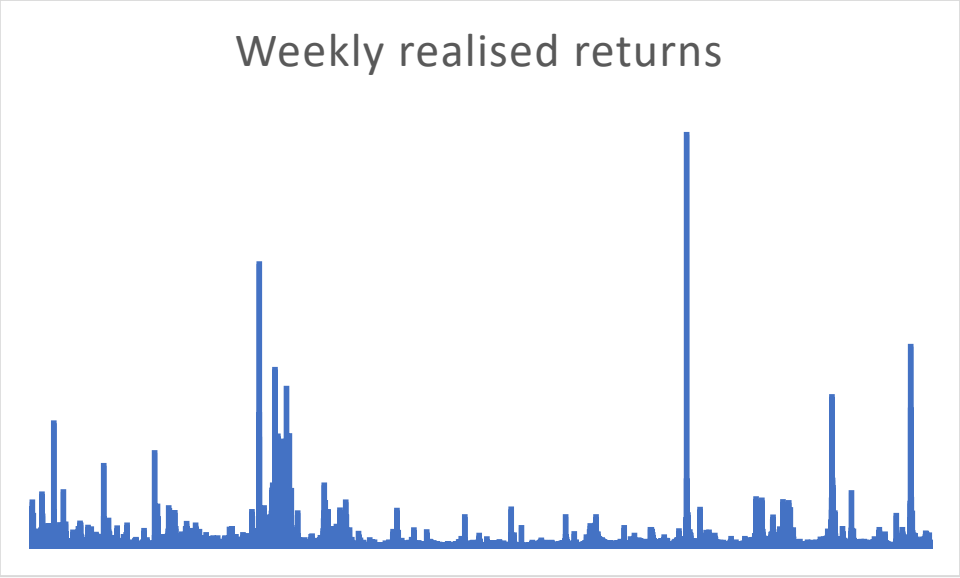


Figure 6: Weekly daily squared realised returns resolution.

In addition to examining the graphical representation in Figure 6, we can deepen our comprehension of the dataset by performing a quantitative analysis. This involves sorting the data vector in descending order, taking into account the presence of 6,564 observations. Consequently, the top 1% of these observations corresponds to 65.64 data points. Summing

the values of the largest 66 observations yields a total of 8,176. Expanding this evaluation to encompass the entire dataset of 41,398 summed-up observations, we calculate the sum of the largest 1%, amounting to 20% of the overall distribution. Notably, this 1% segment, making up a substantial proportion of the total, deviates significantly from what would be anticipated in a normal distribution, where only 3.6% falls within the top 1%. This departure from normal distribution expectations strongly indicates the presence of heavy tails in our dataset. The concentration of the distribution's mass within a limited percentile highlights the non-Gaussian nature of the dataset, suggesting a potentially fat-tailed underlying distribution.

4.3 Linear binning and estimation of the power law exponents

4.3.1 The 'log-log' regression technique

A widely employed technique for analysing empirical frequency distributions involves dividing observed data into bins of constant linear width, yielding a conventional histogram. The methodology of linear binning involves picking a bin with a consistent width, counting the observations in each bin, and then plotting this count against the average value of realisations in every bin.

As emphasised by White et al. (2008, p. 906), the selection of bin width is typically arbitrary in practical applications, representing a balance between the number of bins examined—i.e., the resolution of the frequency distribution—and the precision of each frequency value estimate, where fewer observations per bin result in a less accurate estimate. In the present investigation, the decision to employ uniform width of 10 bins ensures that the residuals of the corresponding log-log regressions, exhibit t-distributed residuals with a minimum of five degrees of freedom. The notation in this study adheres to Clauset et al. (2009).

The conventional method for estimating the power law exponent $\hat{\alpha}$ entails applying linear regression to logarithmically transformed values of both frequency and the mean realisation.

Furthermore, this method necessitates the exclusion of bins containing zero observations, as the natural logarithm of zero, ln (0), is undefined.

For each time frequency, the variance observations are grouped into 10 equally sized intervals. Subsequently, the average variance is computed within each interval, and the count of observations is determined. The representation of the natural logarithm of the observation count within each interval is indicated by denoting it as y_i. Similarly, the natural logarithm of the average RV within each interval is represented as x_i, where i ranges from 1 to N.

Appendix sections 1.1 to 5.1 represent log-log regressions for distinctive time resolutions, where a linear trend is incorporated within each graphical representation as per the visual inspections of these appendices. It becomes apparent that the dots exhibit a similar slope, aligning with Mandelbrot and Hudson's (2008, p. 163) observation that this is *prima facie* evidence of a power law behaviour. To explore the scaling patterns of RVs across different resolutions, the power law model and density function in Equation 13 were utilised.

$$p = Cx^{-\alpha}$$

Equation 13: Power law model.

where:

$p =$ probability density function

$$C = (\alpha - 1)\, x\, \frac{\alpha-1}{MIN}$$

x_{MIN} represents the minimum value that is governed by the power law process

α is the power law exponent, representing the magnitude of the tail exponent.

When we apply natural logarithms to both sides of our relative density function in order to convert the continuous equation into a discrete form, the resulting expression for the power law model appears as presented in Equation 14.

$$\ln(p(x)) = \ln(C) - \alpha \ln(x) + \varepsilon$$

Equation 14: Logarithmic expression of power law.

Regarding α, Taleb (2020, p. 34) noted that the tail exponent of a power law function, through extrapolation, encompasses the low-probability deviation not evident in the data. Nevertheless, this deviation significantly influences the mean. Employing this framework, it can be demonstrated that the expectation of RV is defined as $E[X]$, shown in Equations 15, 16, and 17.

$$E[X] = \int_{x\,MIN}^{\infty} xp(x)\,dx = \frac{(\alpha-1)}{(\alpha-2)} x_{MIN};$$

$$E[X^2] = \int_{x\,MIN}^{\infty} x^2 p(x)\,dx = \frac{(\alpha-1)}{(\alpha-3)} x_{MIN}^2;$$

$$E[X^k] = \frac{(\alpha-1)}{(\alpha-1-k)} x_{MIN}^k$$

Equations 15, 16, and 17: Formulae showing the expected values of the first, second, and k^{th} moment.

Derived from Equation 15, it is evident that the mean of the RV is only present when $\alpha > 2$, whereas the variance for RV is only evident when $\alpha > 3$. Representing the natural logarithm of the number of observations within each interval as y_i and the natural logarithm of the average RV within each bin as x_i (where i = 1, ..., N), the standard OLS regression model in Equation 18 is executed for each time frequency.

$$y_i = c + \alpha x_i + \varepsilon$$

Equation 18: OLS regression formula.

When working with matrix notation, the vector y_i which is also the main dependent variable, is of size *N*×1, where $y_i = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$. In addition, the regressor matrix on the right-hand side of the equation, is of size *N*×2, where $x_i = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$. Given the aforementioned input data, the regression model can be seen in Equation 19.

$$\mathrm{Y} = \mathrm{X}\beta + \mu$$

Equation 19: The regression model.

In the context of the simple linear regression model, we can therefore calculate the parameter β (point estimate), as shown in Equation 20.

$$\hat{\beta} = (\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathrm{Y};$$

$$\hat{\beta} = \begin{bmatrix} \hat{c} \\ \hat{\alpha} \end{bmatrix}$$

Equation 20: Beta hat estimator via MATLAB.

Additionally, the estimated regression residual, $\hat{\mu}$, can be computed by the formula in Equation 21.

$$\hat{\mu} = \mathrm{Y} - \mathrm{X}'\hat{\beta}$$

Equation 21: Estimated residual via MATLAB.

Consequently, once the estimated regression residual is computed, we can estimate the covariance matrix of $\hat{\beta}$ as presented in Equations 22 and 23.

$$\operatorname{cov}\hat{\beta} = \frac{\hat{\mu}'\hat{\mu}}{N-2} * (X'X)^{-1};$$

$$\operatorname{cov}\hat{\beta} = \begin{bmatrix} \operatorname{var}[\hat{c}] & \operatorname{cov}[\hat{\alpha},\hat{c}] \\ \operatorname{cov}[\hat{c},\hat{\alpha}] & \operatorname{var}[\hat{\alpha}] \end{bmatrix}$$

Equations 22 and 23: Covariance matrix of $\hat{\beta}$ via MATLAB.

Such computation is particularly relevant because it yields the corresponding variance of our estimated power law exponent: $var[\hat{\alpha}]$. To verify the statistical significance of our estimated power law exponent $\hat{\alpha}$, we extract the second element from eq. x, denoted as $\hat{\alpha}$, from our preceding computation of point estimate $\hat{\beta}$. Subsequently, we divide $\hat{\alpha}$ by the square root of element a22 in Equation 23, denoted as $var[\hat{\alpha}]$, from our prior calculation of $cov\ \hat{\beta}$ to obtain the t-statistics value. This can be illustrated in Equation 24.

$$t - \operatorname{stat}[\hat{\alpha}] = \frac{\hat{\alpha}}{\sqrt{\operatorname{var}[\hat{\alpha}]}}$$

Equation 24: The t-statistics value.

In this study, the estimation of the power law exponent through log-log regressions leads us to rely on t-statistics, specifically following a t-distribution. The degrees of freedom for this distribution are determined by the number of non-zero bins, denoted as N-2. The subtraction of 2 arises from the estimation of 2 parameters within the OLS framework.

Contrary to the normal distribution, the t-distribution exhibits heavier tails, necessitating different critical values for a given significance level. For instance, the commonly used critical values of -1.96 and +1.96 in the normal distribution shift to different values in the t-distribution.

The confidence interval calculation further emphasises the distinction between these distributions. If one were to incorrectly utilise the normal distribution, the 95% confidence interval would be based on the point estimate plus or minus 1.96 times the square root of the

estimated variance. In contrast, employing the correct t-distribution involves using t-distribution critical values, resulting in a wider confidence interval. The establishment of the power law model, as outlined in Equation 13, necessitates the determination of a suitable value for x_{MIN}. Consequently, a pivotal question emerges: what value should be selected for x_{MIN}?

The Hill estimator, aligning with the MLE, provides us with a solution (Grobys, 2023). For each chosen value of x designated as x_{MIN}, the Hill estimator yields the corresponding $\hat{\alpha}$, as per Equation 25.

$$\hat{\alpha} = 1 + T\left(\sum_{i=1}^{T} \ln\left(\frac{x_i}{x_{MIN}}\right)\right)^{-1}$$

Equation 25: Maximum likelihood estimation.

Where:

$\hat{\boldsymbol{\alpha}}$: Represents the maximum likelihood estimator (MLE) for a specified x_{MIN}

$\mathbf{T}$: represents the count of observations surpassing the designated x_{MIN} threshold.

The resultant alpha exponent from the log-log regressions provide a point estimate of the power law function. Hence, we proceed to identify the corresponding x-minimum value by examining the Hill plot. The Hill plot, presented as a scatter plots in Appendices sections, visually illustrates the interplay between x-values and their associated power law exponents. By identifying the alpha value corresponding to the RV—such as the weekly 5-day RV—on the plot, we determine the threshold indicating the commencement of the power law regime. The proportion of observations governed by a power law process is derived by dividing the count of such observations by the total dataset. This analytical approach unveils a clear distinction in the dataset, where observations surpassing the identified threshold exhibit characteristics aligning with a power law distribution, while those falling below manifest traits indicative of a thin-tailed distribution. Appendix sections 1.3 through to 5.3 present Hill plots for different time frequencies, while Table 3 provides the x_{MIN} value.

4.3.2 Clauset et al.'s (2009) Goodness of Fit test

In the analytical process of estimating optimal power law exponents, practitioners commonly employ a visual method involving the plotting of all alpha hat estimators derived from the MLE, aligned with the corresponding minimum values for x. Utilising the formulation presented in Equation 25 in conjunction with the Hill plots delineated in Appendix sections 1.3 through to 5.3, it is discernible that the estimation of $\hat{\alpha}$ is contingent upon the selected value for the cutoff parameter, denoted as x_{MIN}. The judicious selection of an appropriate cutoff is inherently complex. When the data behaves in a well-defined manner, this approach typically yields a function with a saddle point. The prevailing consensus is that the power law exponent deemed optimal aligns with this saddle point on the Hill plot.

It is imperative, however, to acknowledge the limitations of this visual approach, characterised as a rule of thumb lacking precision. The risk of misleading results looms considerably, particularly when the function is mis-specified. Deviating from the accurate power law exponent can lead to significant errors; an excessively high exponent may result in the underestimation of large realisations, while an excessively low exponent may erroneously convey a higher degree of roughness in the process. It is crucial to note that even marginal adjustments below an $\hat{\alpha}$ exponent of 3 exert substantial influence on the distribution, accentuating the need for precision in estimation. Furthermore, Lux (2000, p. 646) states that "in view of these problems of implementations, the recent development of methods for data-driven selection of the tail sample constitutes an important advance."

In accordance with Lux's (2000) rationale, the present section adopts a contemporary data-driven methodology to ascertain alpha hat, as advocated in the widely referenced work of Clauset et al. (2009). In particular, Clauset et al. advocate a Goodness of Fit (GoF) test that relies on minimising the distance, denoted as D, between the power law function and the observed empirical data. This analytical approach employs the Kolmogorov-Smirnov distance to quantify the maximum difference between the empirical and theoretical distributions. This process is iteratively applied to each maximum likelihood estimator within the set, producing a set of Kolmogorov-Smirnov distances. The optimal maximum likelihood estimator is then

identified as the one associated with the minimum value of the Kolmogorov-Smirnov distance, representing the model that best aligns with the empirical data distribution. Once armed with empirical data and the corresponding set of power law exponents and x_{MIN} values, an additional step involves estimating the theoretical probability for each exponent within the set. For each observation in the dataset, the power law model corresponding to the estimated exponent generates the cumulative probability density function.

Initially, the Kolmogorov–Smirnov (KS) distance is characterised as the maximum separation between the cumulative density functions (CDFs) of the data and the power law model fitted to it, formulated in Equation 26.

$$\mathrm{D} = \mathrm{MAX}|\mathrm{S(x)} - \mathrm{P(x)}|$$

Equation 26: Kolmogorov–Smirnov (KS) distance.

S(x) represents the CDF of the observed data for values equal to or greater than x_{MIN}, while P(x) denotes the CDF for the power law model that optimally captures the data in the range where x is greater than or equal to x_{MIN}. The estimated value x_{MIN} is subsequently determined as the value minimising D.

Moreover, when employing the parameter vector ($\hat{\alpha}$, x_{MIN}) that optimises D, the GoF test (Clauset et al., 2009) yields a p-value, serving as a quantitative measure of the credibility of the power law null hypothesis. Intuitively, this test checks how D compares to distances in similar theoretical datasets based on the assumed model. The p-value is then the proportion of those made-up distances that are greater than the actual distance observed in real data.

If one aims to apply a 5% significance level, the power law null hypothesis for p-values surpassing 5% should not be rejected. This is because any disparity between the actual data and the model may be ascribed solely to statistical fluctuations. The procedure for conducting this test is elaborated in the research by Clauset et al. (2009, p. 675–678).

In the conducted methodology, cutoffs in the dataset were selected using Clauset et al.'s (2009) approach. The estimation of power law exponents was carried out by applying Clauset et al.'s method to the available data. Previously conducted log-log regressions across various time frequencies provided a basis for comparing the consistency of power law exponents derived from Clauset et al.'s method with those obtained through traditional log-log regressions. The analysis focused on the verification of point estimates for alpha hat using Clauset et al.'s approach falling within the 95% confidence interval for alpha hat obtained from log-log regressions.

4.3.3 Can the hypothesis of infinite variance be statistically rejected?

The question emerges as to whether the point estimates derived from the log-log regressions fall statistically below the crucial threshold of 2. This is undeniably a pivotal concern. One must recall that Fama (1963) critically examined Mandelbrot's proposition and provided comments, stating that:

> From a purely statistical standpoint, if the population variance of the distribution of first differences is infinite, the sample variance is probably a meaningless measure of dispersion. Moreover, if the variance is infinite, other statistical tools (e.g., least-squares regression) which are based on the assumption of finite variance will, at best, be considerably weakened and may in fact give very misleading answers. (p. 421)

Therefore, to assess the hypothesis of infinite variance, a hypothesis test should be conducted for each resolution, examining the following hypothesis:

$$H_0: \alpha > 2$$
$$H_1: \alpha \leq 2$$

The null hypothesis (H_0) posits that the power law exponent α is greater than 2, while the alternative hypothesis (H_1) suggests that α is less than or equal to 2. Essentially, this setup allows us to statistically examine whether the observed data adheres to a distribution with an infinite variance, as indicated by an α value equal to or less than 2. The test outcome will provide insights into the nature of the distribution. If the null hypothesis is rejected, it implies

that the distribution of DJIA returns adheres to a power law model characterised by a theoretically undefined mean and variance by definition. In this scenario, extreme events or fluctuations in the Dow Jones returns exhibit a heavy-tailed nature beyond what would be expected in distributions with finite variance and well-defined means.

On the other hand, if H_0 is not rejected, it implies adherence to the law of large numbers, indicating that traditional statistical techniques—like OLS, GMM, GARCH modelling etc.—would work, as the variance would ultimately converge. The main *a priori* hypothesis of this study is that the null hypothesis (H_0) will be rejected, implying that DJIA does not have a defined mean and variance.

4.4 Testing for invariances

4.4.1 Testing for invariance across time frequencies

As highlighted by Mandelbrot and Hudson (2008), the congruence of slopes in log-log regression lines serves as indicative evidence for the manifestation of power law behaviour. It is imperative to acknowledge that, within the context of time series, a process attains fractal characteristics, characterised by statistical self-similarity, when the exponent remains invariant amid varying resolutions.

Consequently, a 95% confidence interval is computed for each time frequency. The construction of this interval involves utilising critical values from a t-distribution with degrees of freedom equivalent to the count of non-zero bins reduced by 2. For instance, when deriving the confidence interval for weekly variance data, critical values from the t-distribution with 5 degrees of freedom are employed, multiplied by the standard deviation of $\hat{\alpha}$ obtained from the regression model.

In this methodology, I started by determining the confidence interval for RVs at a chosen reference time frequency. I then looked at the calculated point estimates ($\hat{\alpha}$) for RVs at various

resolutions, considering different time frequencies. If the comparison of each point estimate with the established reference confidence interval consistently reveals that, regardless of the time frequency under consideration, the point estimates consistently fall within the reference interval, then this observation leads to the conclusion that the hypothesis of invariance cannot be rejected. This consistency signifies a stable relationship between RVs and time frequency. Furthermore, I extended my analysis by employing confidence intervals derived from other time frequencies as references, and the consistent pattern of point estimates within these intervals should persist across different resolutions.

4.4.2. Testing for invariance across time: Evidence from sample splits tests

While the previous section focused on testing the constancy of the power law exponent across different time frequencies, this section explores whether the estimated power law exponents remain consistent over time. Some fractal processes take a long time to reveal their characteristics (Taleb, 2010).

To ensure a sufficient sample size of at least 200 observations, I only considered RVs with 5 to 60 daily squared returns. In this part, I conducted a sub-sample analysis by partitioning the dataset into two distinct subsets. For each time frequency and subsample, I divided variance observations into 10 equal intervals/bins as per the methodology in section 4.3. I then calculated the average realised variance within each interval and counted the number of observations. Using natural logarithms for the observation count (y_i) and the average RV in each bin (x_i), where i = 1 to N, I ran the log-log regression model for each time frequency and subsample.

Within the first sub-sample, I obtained a point estimate closely aligned with the overall sub-sample's point estimator. Subsequently, I derived a distinct confidence interval specific to this point estimator. Notably, log-log regressions only consider bins with non-zero entries. Moreover, 95% of confidence intervals for alpha hat estimates were reported for the first subsample, along with t-statistics using critical values for a t-distribution with degrees of freedom equal to the number of non-zero bins reduced by 2.

In a rigorous examination, the goal is to ultimately assess whether the point estimator from the second sub-sample consistently falls within the confidence interval of the first sub-sample. Should this criterion be met across all resolutions, it substantiates that, statistically, the underlying process governing the power law exponent remains unaltered over time for each resolution. This persistence in the point estimator across both sub-samples would highlight the invariance quality inherent in the power law, affirming its stability throughout the temporal continuum.

4.4.3. Testing for invariance across time: Comparison with a GARCH (1,1) model

> Realised volatility based on intraday quotes is a consistent and highly efficient estimator of the underlying true volatility . . . a realised-volatility-based approach is able to uncover volatility features (asymmetric volatility in particular) that the conventional GARCH type models fail to reveal. (Wang & Yang, 2009, p. 600)

As depicted in section 2.3, ARCH and GARCH are both models used in econometrics and finance to capture the conditional heteroskedasticity, or time-varying volatility, of financial time series data. The primary difference between GARCH and ARCH lies in the model's structure and the variables included in the conditional variance equation. An ARCH model considers only past squared shocks in the conditional variance equation, while GARCH includes both past squared shocks and past conditional variances in the conditional variance equation, as per Equation 27.

$$\sigma^2{}_t = \omega + \alpha_1 \varepsilon^2{}_{t-1} + \beta_1 \sigma^2{}_{t-1}$$

Equation 27: GARCH (1,1) model.

Where:

$\boldsymbol{\sigma^2_t}$ = is the conditional variance of the time series at time t;

ω is the constant term, representing the long-run average of the conditional variance;

$\boldsymbol{\alpha}$ is the coefficient of the lagged squared error term which measures the impact of the past squared shocks on the current conditional variance;

$\boldsymbol{\beta}$ is the coefficient of the lagged conditional variance;

$\boldsymbol{\varepsilon^2_{t-1}}$ is the squared value of the past innovation or shock at time **t−1**. This term captures the information about the past volatility.

The normal distribution assumption is implicit in the model's primary purpose—forecasting volatility. In the context of GARCH, volatility is the conditional standard deviation of returns, and its accurate estimation relies on assuming a normal distribution for the underlying returns.

The GARCH model, under the assumption of normality, estimates volatility based on the conditional variance. This is a critical component for accurate forecasting. Departures from normality may impact the model's ability to precisely capture extreme events or outliers.

Hence, to further test the invariances hypothesis, in this section, I ran a GARCH (1,1) model to estimate conditional variances for DJIA weekly, monthly, and quarterly returns. The data was partitioned into two sub-samples of equal size for each temporal resolution. Subsequently, a GARCH (1,1) model was applied to each sub-sample to estimate conditional variances.

Employing the GARCH (1,1) model, which models the variance of the series, resulted in two crucial point estimates. One equation pertained to the mean, while the other addressed the time-varying variance at time 't'. Following this, point estimates were derived for two sub-samples, and subsequent confidence intervals were established by considering critical values. The crucial evaluation involved checking if the point estimates from the second sample fall within the confidence interval of the first sub-sample's GARCH (1,1) model.

If the point estimates do not align, this would imply that GARCH models generate point estimates that are inherently sample-specific, highlighting the distinctive nature of volatility estimation across various temporal resolutions. This would, by definition, contrast with the

power law model, which is expected to remain invariant and consistent across resolutions as per sections 4.4.1 and 4.4.2.

5 Results

5.1 Descriptive statistics

As outlined in section 4.1, a total of 32,820 daily observations were initially collected as raw data. However, in accordance with the methodology detailed in sections 4.2 through to 4.4, realised variances were calculated for various time intervals. Consequently, as indicated in Table 2, when computing weekly realised variances, the daily squared returns were aggregated, resulting in a decrease in the number of observations to 6,564. While this method aids in analysis and interpretation, it inherently decreases the level of detail in the observations. Indeed, as the data resolution shifted from weekly to semi-annual intervals, the size of the data sample naturally diminished, as illustrated in Table 2.

Table 2: Descriptive statistics of realised variances for different time resolutions.

Days	**5**	**20**	**60**	**90**	**125**
Mean	6.31	25.23	75.93	113.42	157.41
Median	2.72	12.92	42.32	66.37	89.17
Minimum	1.92	1.11	5.68	12.64	19.70
Maximum	663.54	896.87	974.91	1131.40	1391.58
Std. Dev.	17.05	51.23	112.72	157.45	204.26
Skewness	16.88	9.40	4.55	4.05	3.60
Kurtosis	475.74	129.07	28.18	21.74	17.55
Observations	6,564	1641	547	365	263
Sum of 100%	41,398	41,398	41,536	41,398	41,398
Sum of top 20%	28,265	26,107.82	24,468.17	23,757	23,472.79
Top 20%:100%	0.68	0.63	0.59	0.57	0.57

From Table 2, one can observe that, as time resolution decreases, the mean, median, minimum, maximum, and standard deviation increase. This is due to higher variability in returns over shorter periods. Conversely, as time resolution decreases, the kurtosis and skewness decrease as well. This indicates a shift towards a less peaked and more symmetric distribution of returns over longer time periods.

Notably, as per Table 2, the share of the top 20% for the realised variance—based on five daily squared returns—accounts for 68% of the total distribution, whereas, when based on 125 squared daily returns, it accounts for 57% of the total distribution. This is a significant matter, as it indicates that a minor portion of observations holds considerable influence on the overall distribution. This effect becomes even more apparent when contrasted with the expectations from a normal distribution, as outlined in section 3.1. In a normal distribution, the 20% of the cumulative total of distribution accounts for only 44% of the cumulative total. Hence, this is already a good indication that the distribution of the data sample appears to be behaving closely to the Pareto 80/20 rule.

To enhance the understanding of the dataset, as per Grobys's (2022) methodology—presented in Figure 1 in section 2.1—the top 0.5% and the bottom 0.5% of the sorted data set from largest to smallest returns were deleted. Figure 7 depicts an artificial time-series evolution drawn against the original dataset.

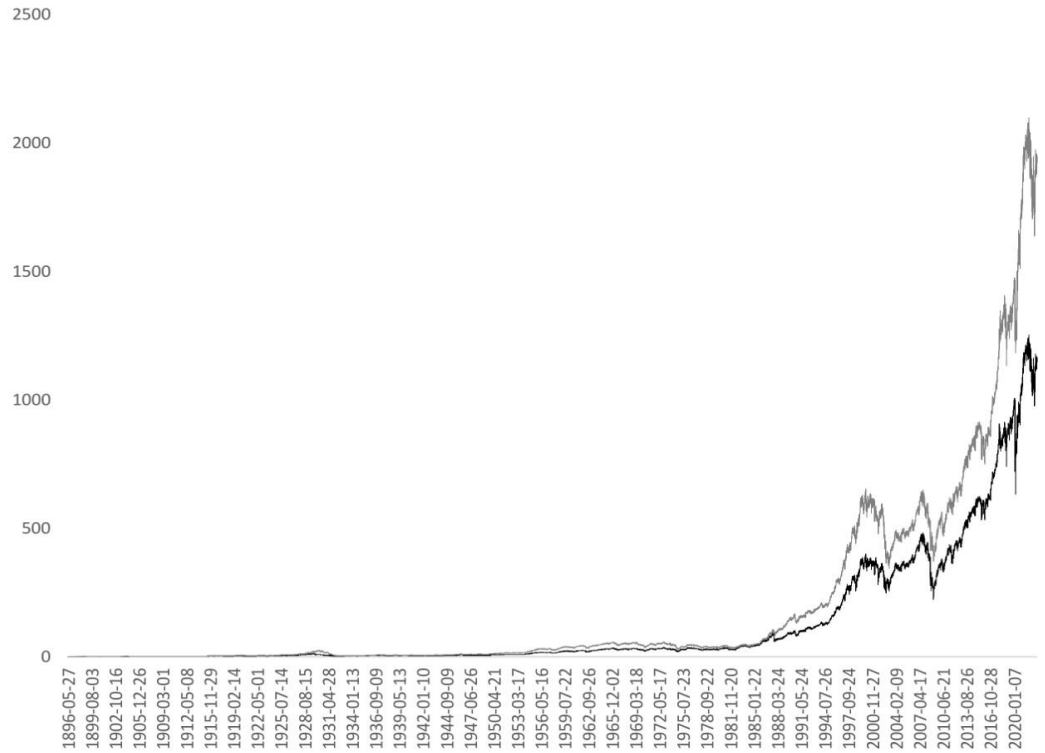


Figure 7: Compounded returns comparison for DJIA 30, May 1896 - February 2023.

The black line represents the compounded returns for the DJIA 30 over the sample, whereas the grey line shows compounded returns where the 1% of the observations are deleted, or set to 0. This means that 0.5% of the highest observations ranging from [0.038644, 0.154275] and 0.5% of the lowest observations ranging from [-0.04127, -0.22611] are deleted. The final returns are 1152.40796 versus 1931.30184, which implies that the deleted 1% of the observations comprise 67.59% of the total implied returns. The two sets of compounded returns—one including all observations and the other excluding the top 1%—exhibit such divergent patterns that they appear to represent entirely different time-series evolutions.

This observation likely stresses the significant impact that extreme observations—the top and bottom; i.e. 1% of the most exuberant realised returns—have on the overall behaviour and performance of the DJIA 30 index. The fact that excluding these extreme values leads to such a substantial difference in the visual evolution, as shown in the graph in Figure 7, indicates that these extreme observations are not just outliers but have a disproportionate influence on the index's performance.

5.2 Estimated power law exponents via OLS regressions and MLE in association with Kolmogorov Smirnov statistic

As detailed in section 4, the conventional methodology for finding the power law exponent involves regressing Y on x, where Y signifies the log-transformed values of frequency and x denotes the log-transformed values of realisations within each interval, with a bin width of 10. The steepness of the line offers an estimate of the $\hat{\alpha}$ exponent. In all Appendices, the plots display a viewable examination of log-log graphs for different time frequencies. A mere visual inspection reveals that the dots slope in a similar manner, aligning with the observation by Mandelbrot and Hudson (2008, p. 163) and providing initial evidence of power law behaviour.

The method of OLS regression provides a point estimate and an uncertainty surrounding that point estimate. In Table 3, the estimated intercept and $\hat{\alpha}$ exponents are presented alongside their t-statistics, provided that only non-zero bins are considered. Moreover, the 95%

confidence interval range for these exponents, as well as the R-squared values for each resolution, are reported.

Table 3: Point estimates from OLS regressions.

Days	**5**	**20**	**60**	**90**	**125**
$\hat{c}$	3.026	6.156	7.365	7.454	6.788
(*t*-statistic)	(5.078)	(7.944)	(8.94)	(5.22)	(6.64)
$\hat{\alpha}$	1.92	2.12	2.044	1.93	1.67
(*t*-statistic)	(16.51)	(14.53)	(14.44)	(8.13)	(10.24)
95% CI for α	[1.62, 2.22]	[1.84, 2.41]	[1.72, 2.37]	[1.37, 2.49]	[1.25, 2.09]
R-squared	0.98	0.99	0.96	0.90	0.95
Observations	6,564	1,641	547	365	263
x_{MIN}	1.87	10.61	26.122	29.33	25.29
$\hat{\lambda}$	0.08	-0.12	-0.044	0.07	0.33
(in %)	63%	58%	72%	87%	97%
(*t*-statistic)	19.89	21.10	18.31	10.24	15.73
Non-zero bins	7	7	10	9	7

As per the main *a priori* hypothesis mentioned in the methodology section 4.3.3, in most scenarios, particularly in 3 out of 5 instances (5-, 90, -125 resolutions), the estimated $\hat{\alpha}$ estimates fall below 2; hence, rejecting the null hypothesis. Additionally, they exhibit statistical significance, as indicated by the large t-statistic values beyond the range of -2 and 2. This aligns with the evidence of a power law behaviour, suggesting that the theoretical mean is undefined since the alpha exponent is less than 2. This is supported by an excellent Gof (R-squared) ranging between 0.90 and 0.99. In all cases, the alpha exponents are less than 3, implying that the variance is also undefined. The findings align with Groby's (2022) research, which indicate that the realised momentum variance (RMV) exhibits scaling behaviour. Specifically, the power law exponents for the variance processes consistently yielded an alpha exponent below 2 across various resolutions.

The undefined variance suggests that extreme events or outliers can occur more frequently than expected, leading to uncertainty about the behaviour of the distribution. The interesting question to consider is whether the point estimates are statistically equivalent. Upon closer examination at the highest resolution—namely the five-day resolution—by observing the 95%

confidence interval using the uncertainty principle provided by the log-log regression, we find values ranging from 1.62 to 2.22 as the confidence interval for the power law exponent. It becomes evident that, regardless of the point estimate for other resolutions, all of them fall within this range of point estimate intervals. Therefore, from a statistical perspective, we would be inclined to accept the invariance hypothesis. This suggests that, statistically, these estimates are equivalent regardless of the resolution under investigation.

Moreover, as depicted in Table 4, the methodology from Clauset et al. (see section 4.3.2) was also used to calculate the alpha exponent using the MLE technique. Scripts were retrieved from Clauset's website, and such calculations were conducted in MATLAB. The point estimates given indicate that the power law null hypothesis can be rejected because the GoF (p-value) is less than 0.05 in all cases. However, some of the point estimates from this approach—namely, the 20- and 90-day resolutions—are within the confidence interval range reported from the estimation of the OLS log-log regressions. Nonetheless, even if some of the exponents amounted to slightly more than 2, they are still less than 3, which confirms that the variance is infinite, or undefined, in all cases.

Table 4: Estimation of power law exponents using the MLE technique.

Days	**5**	**20**	**60**	**90**	**125**
$\hat{\alpha}$	2.311	2.35	2.59	2.43	2.57
x_{MIN}	6.33	23.73	106.76	77.92	171.06
GoF test	0.026	0.032	0.038	0.049	0.061

To test the issue of whether, statistically, the point estimates from the log-log regression are below the value of 2, a one-sided test statistic $\hat{\lambda}$ is also reported in Table 3. This was implemented through the equation: $\hat{\lambda} = (2 - \hat{\alpha})$. In all instances, the null hypothesis was rejected, implying that the computed realised variances appear to be governed by power laws with theoretically undefined mean and variance. This supports Grobys's (2022) findings and Mandelbrot's (1963) findings on the infinite variance of cotton price changes.

Besides the reporting of $\hat{\alpha}$ exponents, the corresponding x_{MIN} values are also reported. As mentioned in section 4.3.1, in graphical terms, the x_{MIN} value represents a cutoff point, beyond which the power law relationship becomes apparent. As per sections 4.3.1 and 4.3.2, the Hill estimator, which aligns with the MLE, provides us with the corresponding $\hat{\alpha}$ estimate for each selected value of x_{MIN} as per the equation 25 (see section 4.3.1). Table 3 reports these cutoffs derived from the visual inspection of the Hill plot. In addition, the proportion of observations governed by a power law process are also reported in percentages. Using this information, one can notice that part of the distributions governed by a Paretian tail ranges from 58%, the second highest resolution, to 97%, the lowest resolution. This finding complements with Grobys's (2022) findings that, as the resolution decreases, the percentage of observations governed by a power law increases. This contradicts the expectation that, as time resolution decreases, the percentage of observations governed by a power law should decrease as well.

5.3 Testing the invariance hypothesis across time frequencies and time

As depicted in Table 3, each $\hat{\alpha}$ point estimate falls within each confidence interval range regardless of the time resolution being considered. For instance, if one considers the weekly resolution estimated alpha exponent 1.92, this falls within the confidence interval range of all other resolutions. This implies that there is a stable relationship between RVs and different time frequencies. Hence, the hypothesis of invariance cannot be rejected, as all point estimates fall within the confidence intervals persisting across different time resolutions.

Table 5: Point estimates from OLS regressions for the first subsample.

Days	**5**	**20**	**60**
$\hat{c}$	3.79	6.21	5.96
(*t*-statistic)	5.96	6.39	8.78
$\hat{\alpha}$	2.06	2.03	1.69
(*t*-statistic)	15.48	11.07	13.74
R-squared	0.97	0.97	0.97
Non-zero bins	7	6	7
Observations	3,281	820	273
95% CI for $\hat{\alpha}$	[1.72,2.41]	[1.52,2.54]	[1.37, 2.01]

Table 6: Point estimates from OLS regressions for the second subsample.

Days	**5**	**20**	**60**
$\hat{c}$	2.72	4.70	6.38
(*t*-statistic)	2.66	2.98	10.08
$\hat{\alpha}$	1.93	1.84	1.82
(*t*-statistic)	9.09	6.22	16.48
R-squared	0.95	0.93	0.99
Non-zero bins	5	5	6
Observations	3,282	821	274

As outlined in the methodological section of 4.4.2, sample split tests were deliberately conducted to assess the constancy of power law exponents over time. This involved dividing the sampling data equally into first and second sub-samples of identical sizes for time resolutions of 5, 20, and 60. Subsequently, log-log regressions were performed on each sub-sample, resulting in new confidence intervals. As noted in Table 6, point estimators from the second subsample for time resolutions 5, 20, and 60 fell within the confidence interval range of the first sub sample in Table 5 in all cases.

Hence, statistically, the power law exponents persisted across both sub-samples, implying the invariance quality of the power law. The stability of the power law suggests that the scaling properties of the system remain consistent across different time scales. This implies that the

statistical regularities described by the power law persist regardless of the level of aggregation. In other words, the relationship between variables maintains its form even when observed over different time intervals.

These results complement Groby's (2022) novel finding that sample split tests reveal that the invariance hypothesis holds both for different time frequencies and across time.

5.3.1 GARCH model estimates over time

The purpose of employing a GARCH model is to construct a comparative narrative between a power law model that consistently confirms the invariance hypothesis and the attributes of GARCH modelling. Employing the same methodology involved conducting sample split tests across day resolutions of 5, 20, and 60, with the objective of investigating whether the GARCH model exponents persist over time or if new point estimates emerge for different samples.

Table 7: Point estimates from GARCH (1,1) model for the first subsample.

Days	**5**	**20**	**60**
ω	0.06	0.84	12.93
Resid $(-1)^2$	0.065	0.117	0.24
p-value	0.00	0.00	0.00
(*z*-statistic)	15.66	6.15	3.25
95% CI	[0.06,0.07]	[0.07,0.15]	[0.09,0.37]
GARCH(-1)	0.93	0.86	0.66
p-value	0.00	0.00	0.00
(z-statistic)	53.25	44.96	6.72
95% CI	[0.91,0.93]	[0.82,0.89]	[0.46, 0.85]
Observations	3,297	759	254

Table 8: Point estimates from GARCH (1,1) model for the second subsample.

Days	**5**	**20**	**60**
ω	0.27	1.00	13.75
Resid $(-1)^2$	0.16	0.095	0.13
p-value	0.00	0.00	0.1278
(*z*-statistic)	15.32	3.95	1.52
95% CI	[0.14,0.18]	[0.04,0.14]	[-0.03, 0.3]
GARCH(-1)	0.78	0.85	0.65
p-value	0.00	0.00	0.00
(z-statistic)	53.25	26.67	2.91
95% CI	[0.76,0.81]	[0.79,0.91]	[0.21,1.08]
Observations	3,299	759	254

According to Tables 7 and 8, and as described in section 4.4.3, the GARCH (1,1) model aims to capture the variance of the data series, yielding two distinct point estimates. At the weekly resolution, the two obtained point estimates Resid $(-1)^2$ and GARCH (-1) of the second sub sample (Table 8) fall outside the confidence interval of the preceding subsample 1 (Table 7). Furthermore, at the 60-day resolution (the quarterly resolution), the GARCH estimates indicate that the point estimate for Resid $(-1)^2$ in the second subsample (Table 8) lacks significance—as the p-value: 0.1278 is greater than 0.05—while it is significant in the first subsample.

These findings collectively suggest the presence of sample-specific characteristics within the data. Therefore, this strengthens the argument that, in situations where variance is considered infinite, conventional statistical approaches like GARCH modelling, which assume finite variance, may lose their effectiveness and could yield misleading results.

5.4 Limitations and avenues for future research

Given the Hill plots in the Appendices section, a practitioner might perceive that the exponent is approximately 2.5 or somewhere close to it. However, it is important to acknowledge the

possibility that, just by looking at the Hill plots, some may hold the perspective that the exponent is not less than 2. Nevertheless, even if the exponent is not less than 2, it still remains below 3. This implies that we do not witness the value of the realised variance in finite samples. To achieve this, we would need approximately 1,000,000,000 observations—a number far beyond what we currently possess.

Another limitation stems from the argument articulated in Cirillo and Taleb (2020), emphasising the behaviour of fat-tailed distributions. The authors assert that, in such distributions, the extremes carry more statistical significance than the more frequently occurring events, which may be considered as noise. This notion challenges the efficacy of Clauset et al.'s (2009) GoF test, which relies on optimised Kolmogorov–Smirnov distances with the power law model under the null hypothesis. Critics often argue that this test may be weak in capturing the differentiations of fat-tailed distributions, potentially impacting the reliability of the results derived from it.

To address this limitation and pave the way for future research, a promising avenue would be to investigate whether power laws offer a more accurate depiction of the behaviour of realised variances compared to the lognormal distribution. By conducting rigorous empirical analyses and statistical tests, researchers can assess the suitability of power laws in capturing the distributional properties of realised variances, particularly in the context of fat-tailed distributions. This exploration could provide valuable insights into the underlying dynamics of financial markets and contribute to the development of more robust modelling frameworks for financial data analysis.

Another limitation may be that some may argue that, when using the MLE technique, there is an assumption that the data are independently distributed. However, realised variances, which are commonly used, often exhibit autocorrelation, meaning they are not independent. To tackle this issue head-on, a method called blocks bootstraps, recently suggested by Grobys et al. (2021), can be employed. This method is designed to address dependencies in the data. Essentially, it involves simulating artificial time series of realised variances using blocks bootstraps. Each time frequency undergoes a simulated 1,000 artificial time series, using

randomly drawn block lengths. These block lengths are specifically chosen to mitigate the effects of autocorrelation, with an expected block length of 10% of the sample length. The distribution of these block lengths is governed by a geometric distribution, ensuring randomness. By computing statistical values for each simulated time series and testing against a hypothesis, one can assess the robustness of the results. This approach could help address the limitations associated with autocorrelation in realised variances.

5.5 Conclusion and policy implications

The methodology adopted to investigate the alpha-stable Lévy hypothesis within U.S. stock market data was characterised by a rigorous approach. An extensive collection of daily data, spanning from the inception of the DJIA index in 1896 to 2023, ensured a rich and comprehensive dataset for the analysis. Subsequently, the computation of realised variances quantified the observed volatility in daily DJIA returns, employing log-log regression techniques to analyse the empirical frequency distributions. To validate such findings, Clauset et al.'s (2009) GoF test, a sophisticated data-driven methodology aimed at minimising the Kolmogorov-Smirnov distance between the empirical and theoretical distributions, was also applied.

Next, hypotheses tests were employed to evaluate whether the point estimates obtained from log-log regressions significantly deviated from the critical threshold of 2; thereby, critically examining the hypothesis of infinite variance. Formulating the null hypothesis (H_0) as $\alpha > 2$ and the alternative hypothesis (H_1) as $\alpha \leq 2$, the analysis delved into the statistical adherence of the observed data to a distribution with infinite variance. Rejection of H_0 would signify adherence to a power law model characterised by a theoretically undefined mean and variance, suggesting a heavy-tailed nature of extreme events or fluctuations in DJIA returns beyond expectations in distributions with finite variance. Conversely, non-rejection of H_0 would imply adherence to the law of large numbers, indicating the potential convergence of variance. This procedure substantiated the primary *a priori* hypothesis that the DJIA lacked a defined mean and variance; thus, contributing significantly to the understanding of the nature of financial market dynamics.

The results posit that, in three out of five cases of resolutions 5-day, 90-day and 125-day, estimated alpha exponents fall below 2, leading to the rejection of the null hypothesis and, by extension, aligning with the primary *a priori* hypothesis. This was supported by high t-statistic values indicating statistical significance and high R-squared values signifying a strong GoF. These findings align with Grobys's research on scaling behaviour in realised momentum variance (RMV), suggesting a power law pattern with an undefined theoretical mean due to alpha being less than 2 and variance also being undefined. The examination of confidence intervals at the highest resolution supported the acceptance of the invariance hypothesis, with Clauset et al.'s (2009) methodology yielding similar results. Despite some estimates slightly exceeding 2, they remained below 3, confirming infinite variance. A consistent rejection of the null hypothesis through a one-sided test statistic implied that realised variances are governed by power laws with undefined mean and variance, consistent with the previous research by Grobys (2022) and Mandelbrot (1963).

Furthermore, a detailed examination of the stability of power law exponents across different time resolutions was employed to reveal the enduring nature of power law phenomena across various time scales. Additionally, obtained exponents from GARCH (1,1) models were contrasted with those obtained from power law models. This comprehensive analysis was done to confirm the presence of power law dynamics in the DJIA index and to clarify its consistent stability and behaviour over time, market efficiency, and risk assessment.

The analysis revealed that estimated alpha values consistently fall within confidence intervals across various time resolutions, indicating stability in the relationship between RVs and different frequencies and supporting the hypothesis of invariance. Sample split tests confirmed the constancy of power law exponents over time by showing that estimators from different sub-samples align within confidence intervals. This implied a persistent power law and consistent scaling properties across time scales, suggesting that statistical regularities endure regardless of observation intervals. These findings echo Groby's research, reinforcing the invariance hypothesis across different time frequencies and over time.

The results showed that, in the GARCH (1,1) model, point estimates for variance parameters differ between subsamples, with some falling outside confidence intervals and others lacking significance in the second subsample. This points to sample-specific characteristics within the data. Such findings imply that conventional statistical approaches like GARCH modelling may be less effective in situations where variance is infinite, potentially yielding misleading results.

A policy implication arising from the findings is the need for a paradigm shift in statistical modelling practices. The consistent confirmation of the invariance hypothesis through power law models suggested that power laws should be considered the norm rather than the exception in financial modelling and analysis. This implies a departure from traditional statistical models like OLS, GMM, and others, which rely heavily on the assumption of normal distribution. However, the reality depicted by the research indicates that such models may not fully capture the complexity and dynamics of financial systems. Therefore, policymakers should acknowledge the limitations of traditional statistical approaches and explore alternative modelling frameworks that better accommodate the observed non-normal behaviours of financial data. Additionally, there is a need for enhanced risk management strategies that account for the potential for extreme events and outliers, which are more prevalent in systems governed by power laws. Integrating these insights into policymaking processes can lead to more robust and resilient financial systems capable of withstanding unexpected shocks and ensuring long-term stability.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics, 61*(1), 43-76. https://doi.org/10.1016/S0304-405X(01)00055-1

Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de L'Ecole Normale Supérieure*, 17, 21--88.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637-654. http://www.jstor.org/stable/1831029

Bodie, Z., Kane, A., & Marcus, A. J. (2021). *Investments* (12th ed.). Mcgraw-Hill Education.

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661-703. http://www.jstor.org/stable/25662336

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987-1007. https://doi.org/10.2307/1912773

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics, 122*(2), 221-247. https://doi.org/10.1016/j.jfineco.2015.12.00

Fama, E. F. (1963). Mandelbrot and the stable Paretian hypothesis. *The Journal of Business*, *36*(4), 420-429. http://www.jstor.org/stable/2350971

Gabaix, X. (2009). *Power laws in economics and finance*. Annual Reviews. https://doi.org/10.1146/annurev.economics.050708.142940

Grobys, K. (2021). What do we know about the second moment of financial markets? *International Review of Financial Analysis*, *78*, 101891. https://doi.org/10.1016/j.irfa.2021.101891

Grobys, K. (2022a). *Fractal finance L1: Questionable research in finance: How much worth is a Nobel Prize in economics?* YouTube. https://www.youtube.com/watch?v=K6wrINLgNRc&t=1789s

Grobys, K. (2022b). *Rationality: The antidote to being fooled by the industry*. BoD - Books on Demand.

Grobys, K. (2022c). *When "story telling" gets published in scientific journals: On the momentum illusion*.

Grobys, K. (2023a). A multifractal model of asset (in)variances. *Journal of International Financial Markets, Institutions and Money*, *85*, 101767. https://doi.org/10.1016/j.intfin.2023.101767

Grobys, K. (2023b). *2023 RFM lecture 03: On the second moment of financial assets*. YouTube. https://www.youtube.com/watch?v=qbtlF2p4Q6I

Grobys, K. (2023c). A fractal and comparative view of the memory of Bitcoin and S&P 500 returns. *Research in International Business and Finance*, *66*, 102021. https://doi.org/10.1016/j.ribaf.2023.102021

Grobys, K., Junttila, J., Kolari, J. W., & Sapkota, N. (2021). On the stability of stablecoins. *Journal of Empirical Finance*, *64*, 207-223. https://doi.org/10.1016/j.jempfin.2021.09.002

Hou, K., Xue, C., & Zhang, L. (2018). Replicating anomalies. *The Review of Financial Studies*, *33*(5), 2019-2133. https://doi.org/10.1093/rfs/hhy131

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65-91. https://doi.org/10.2307/2328882

Kelly, B. T., Moskowitz, T. J., & Pruitt, S. (2021). Understanding momentum and reversal. *Journal of Financial Economics*, *140*(3), 726-743. https://doi.org/10.1016/j.jfineco.2020.06.024

Lux, T. (2000). On moment condition failure in German stock returns: An application of recent advances in extreme value statistics. *Empirical Economics*, *25*, 641-652.

Lux, T., & Alfarano, S. (2016). Financial power laws: Empirical evidence, models, and mechanisms. *Chaos, Solitons & Fractals*, *88*, 3-18. https://doi.org/10.1016/j.chaos.2016.01.020

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business, 36*(4), 394-419. http://www.jstor.org/stable/2350970

Mandelbrot, B. B., & Hudson, R. L. (2008). *The (mis)behaviour of markets: A fractal view of risk, ruin and reward*. Profile.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77-91. https://doi.org/10.2307/2975974

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, *4*(1), 141-183. https://doi.org/10.2307/3003143

Olsen, R. B. (2023). Financial markets: From fractals to power laws. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4344887

Renò, R., & Rizza, R. (2003). Is volatility lognormal? Evidence from Italian futures. *Physica A: Statistical Mechanics and its Applications*, *322*, 620-628. https://doi.org/10.1016/S0378-4371(02)02023-X

Schwert, G. W., et al. (2002). Anomalies and market efficiency.

Serra-Garcia, M., & Gneezy, U. (2021). Nonreplicable publications are cited more than replicable ones. *Science Advances*, *7*(21), eabd1705. https://doi.org/10.1126/sciadv.abd1705

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425-442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable.* New York, NY: Random House.

Taleb, N. N. (2020). *Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications.*

Teichmoeller, J. (1971). A note on the distribution of stock price changes. *Journal of the American Statistical Association*, *66*(334), 282-284. https://doi.org/10.2307/2283922

Wang, J., & Yang, M. (2009). Asymmetric volatility in the foreign exchange markets. *Journal of International Financial Markets, Institutions and Money*, *19*(4), 597-615. https://doi.org/10.1016/j.intfin.2008.10.001

West, G. (2017). *Scale: The universal laws of life, growth, and death in organisms, cities, and companies*. Penguin.

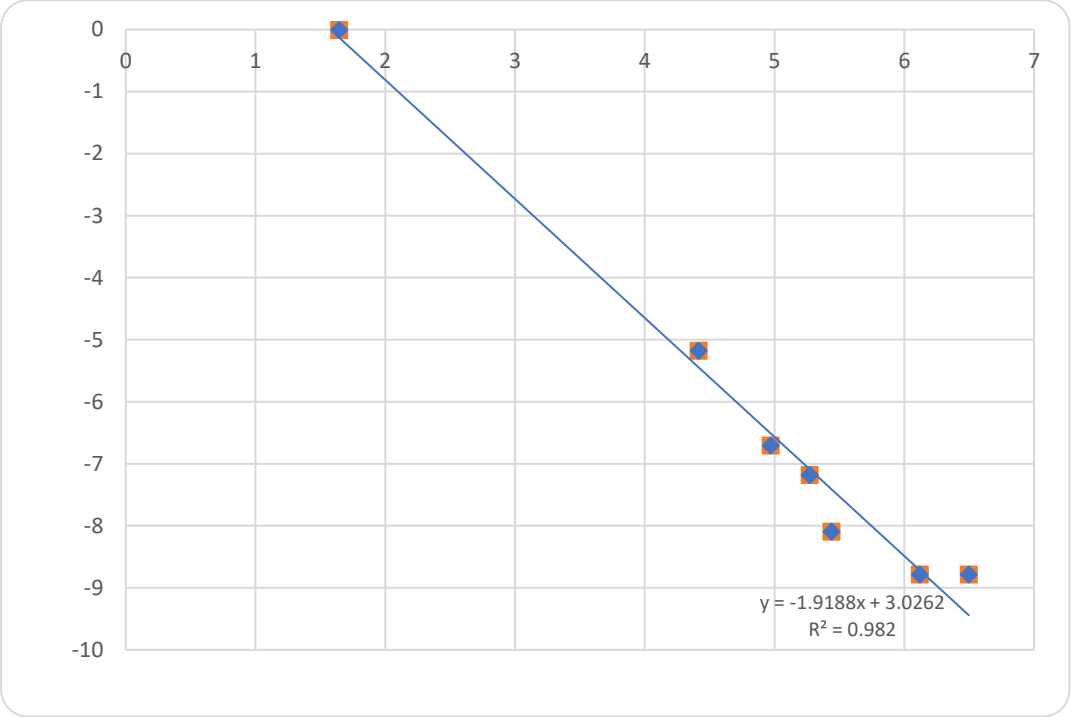
White, E. P., Enquist, B. J., & Green, J. L. (2008). On estimating the exponent of power-law frequency distributions. *Ecology*, *89*(4), 905-912. http://www.jstor.org/stable/27651627

Wiest, T. (2022). *Momentum: What do we know 30 years after Jegadeesh and Titman's seminal paper?* Springer Science and Business Media LLC. https://doi.org/10.1007/s11408-022-00417-8

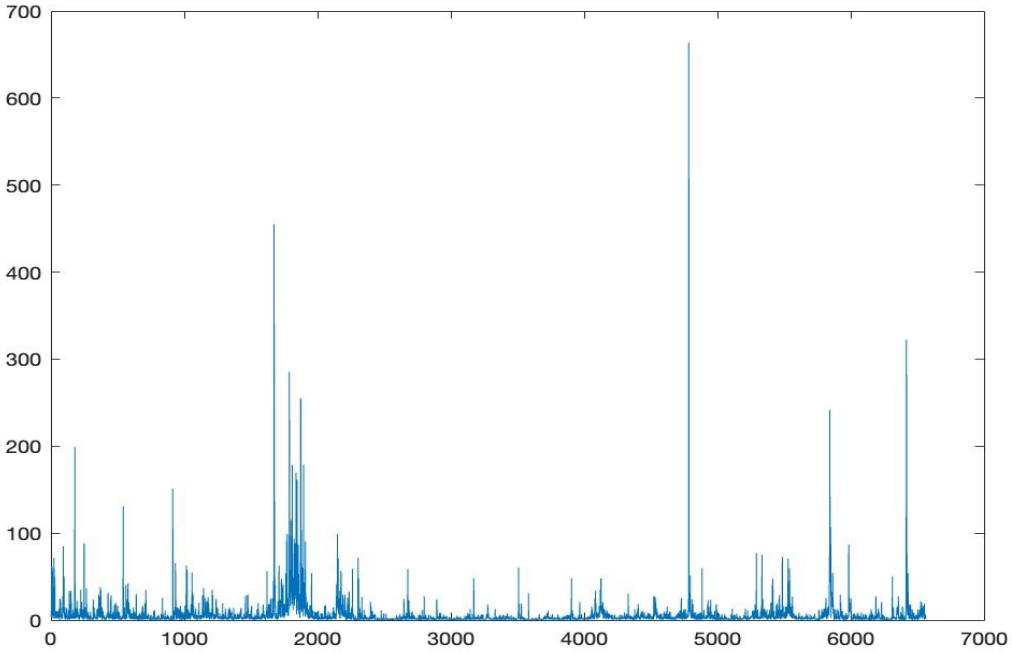
Appendices

Appendix 1: 5 squared daily realised returns time resolution

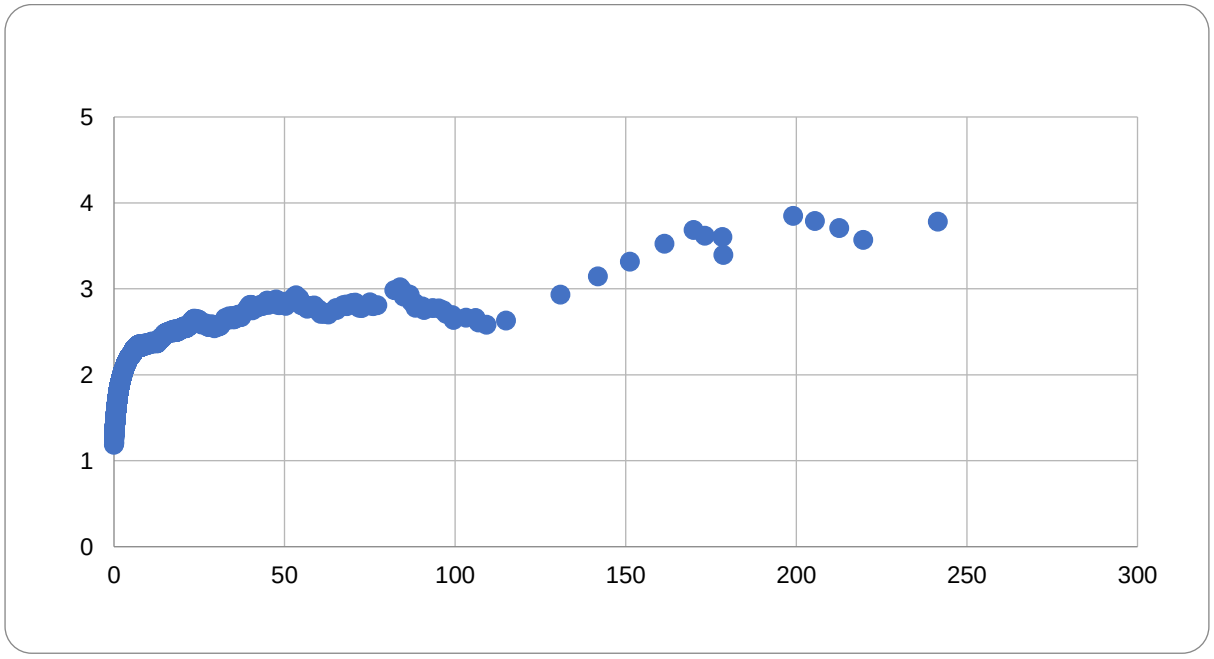
Appendix 1.1: Log-log OLS regression technique for 5 squared daily realised returns



Appendix 1.2: Visual inspection of 5 daily squared returns resolution

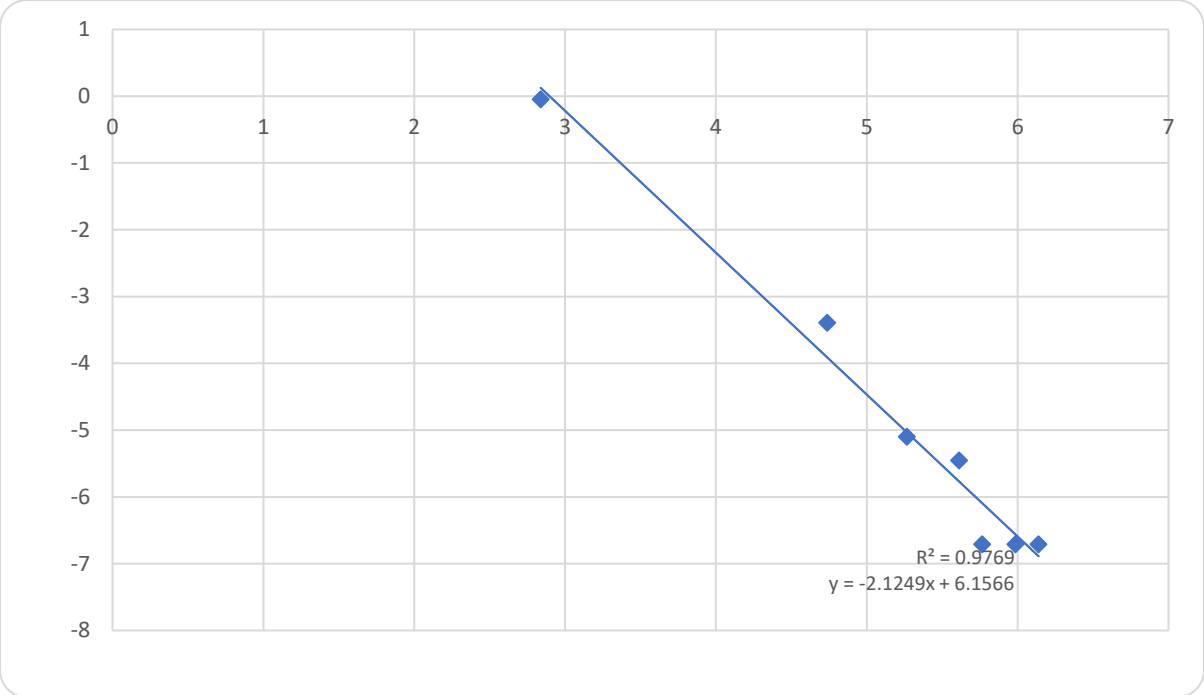


Appendix 1.3: Hill plot for realised variance based on 5 squared daily returns

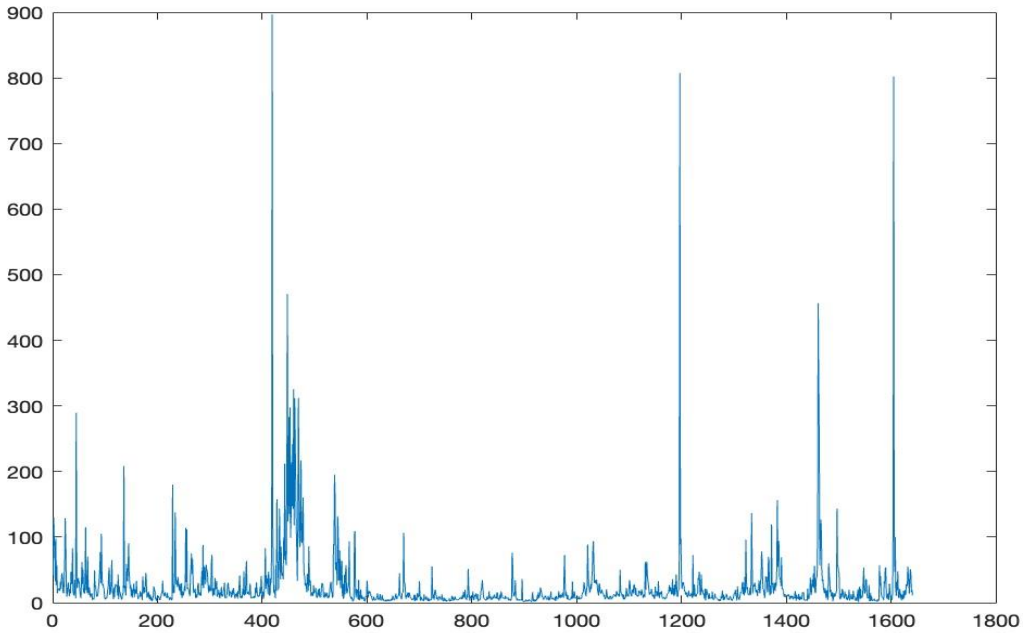


Appendix 2: 20 squared daily realised returns time resolution

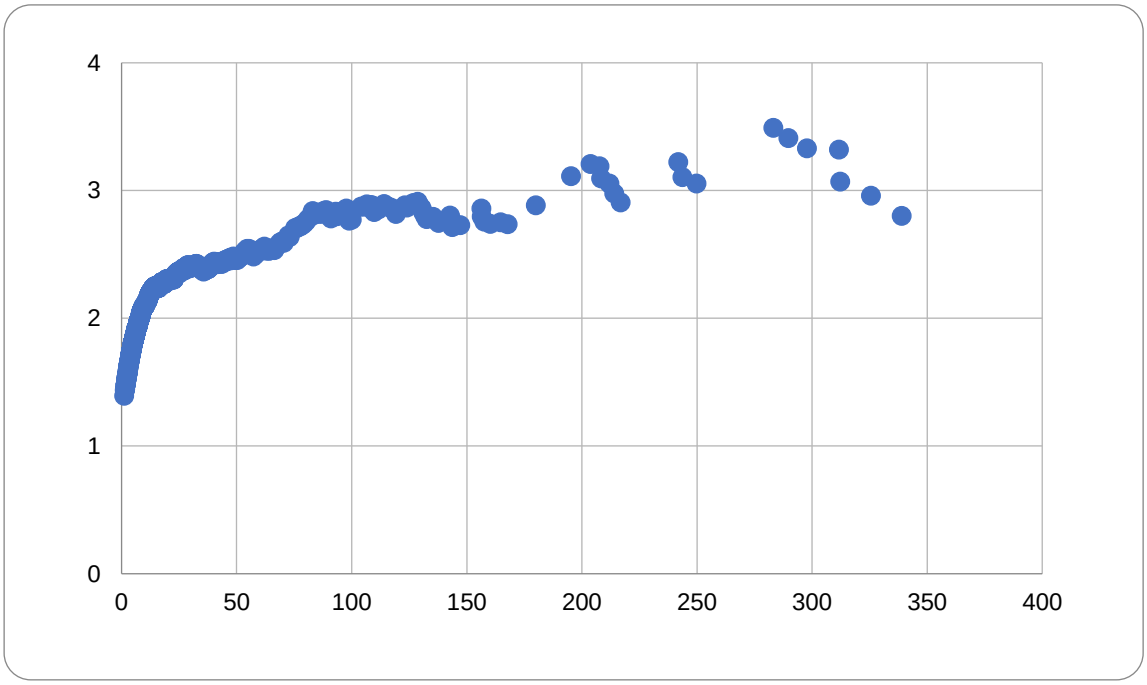
Appendix 2.1: Log-log OLS regression technique for 20 squared daily realised returns



Appendix 2.2: Visual inspection of 20 daily squared returns resolution

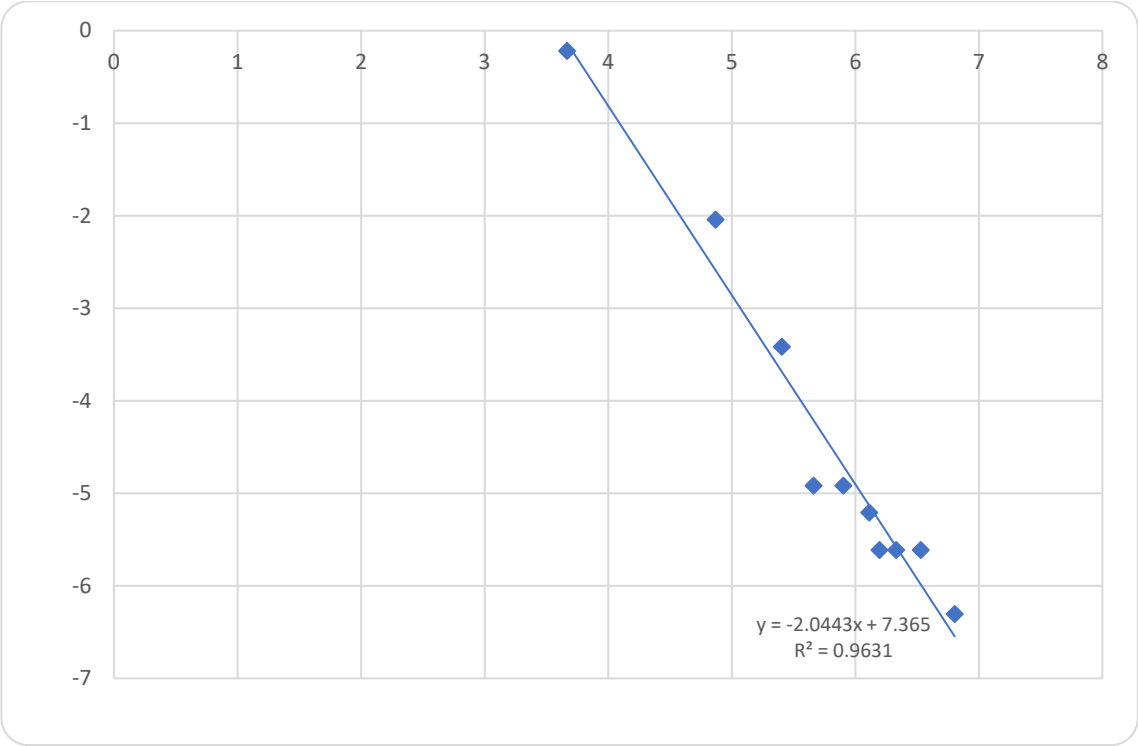


Appendix 2.3: Hill plot for realised variance based on 20 squared daily returns

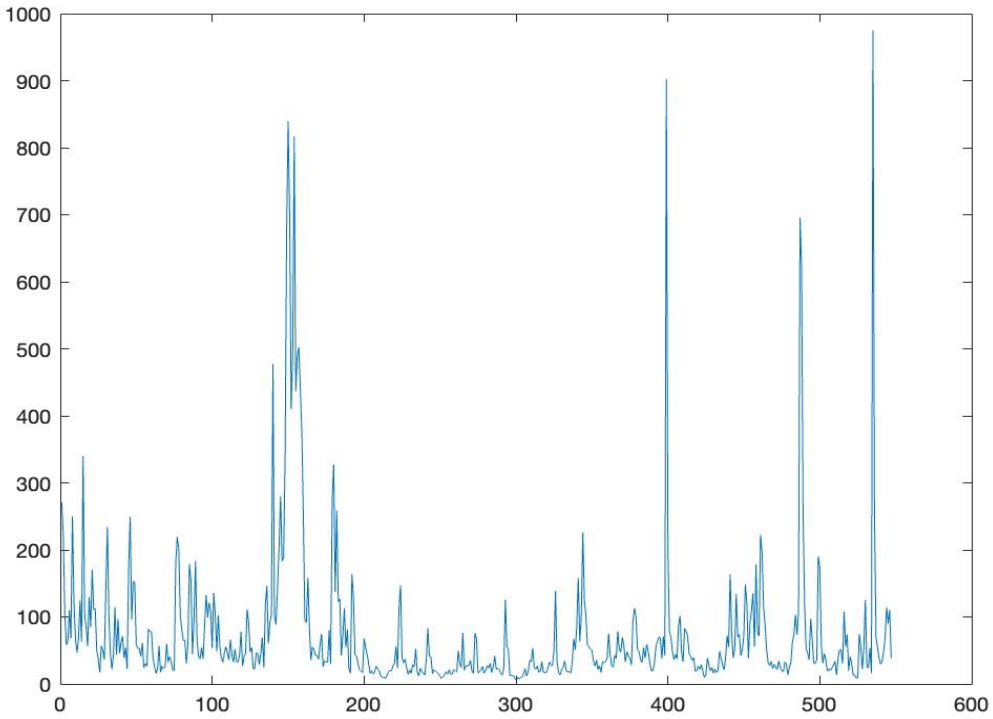


Appendix 3: 60 squared daily realised returns resolution

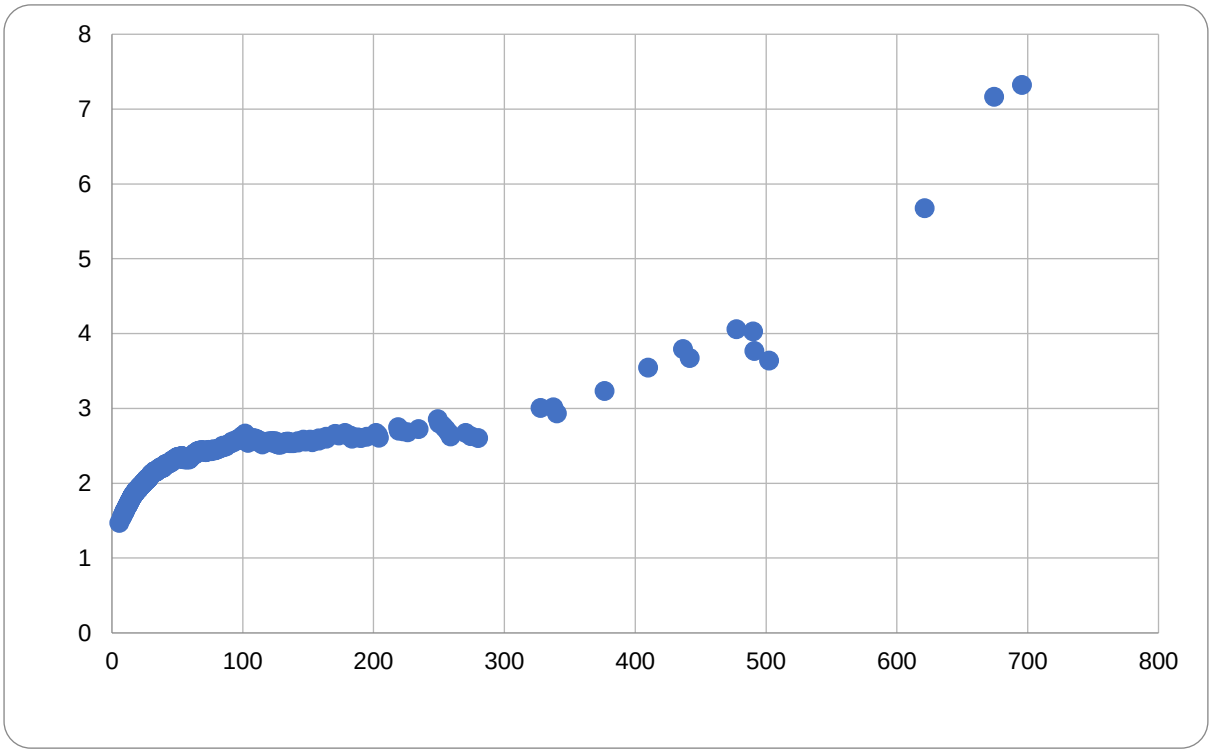
Appendix 3.1: Log-log OLS regression technique for 60 squared daily realised returns



Appendix 3.2: Visual inspection of 60 daily squared returns resolution

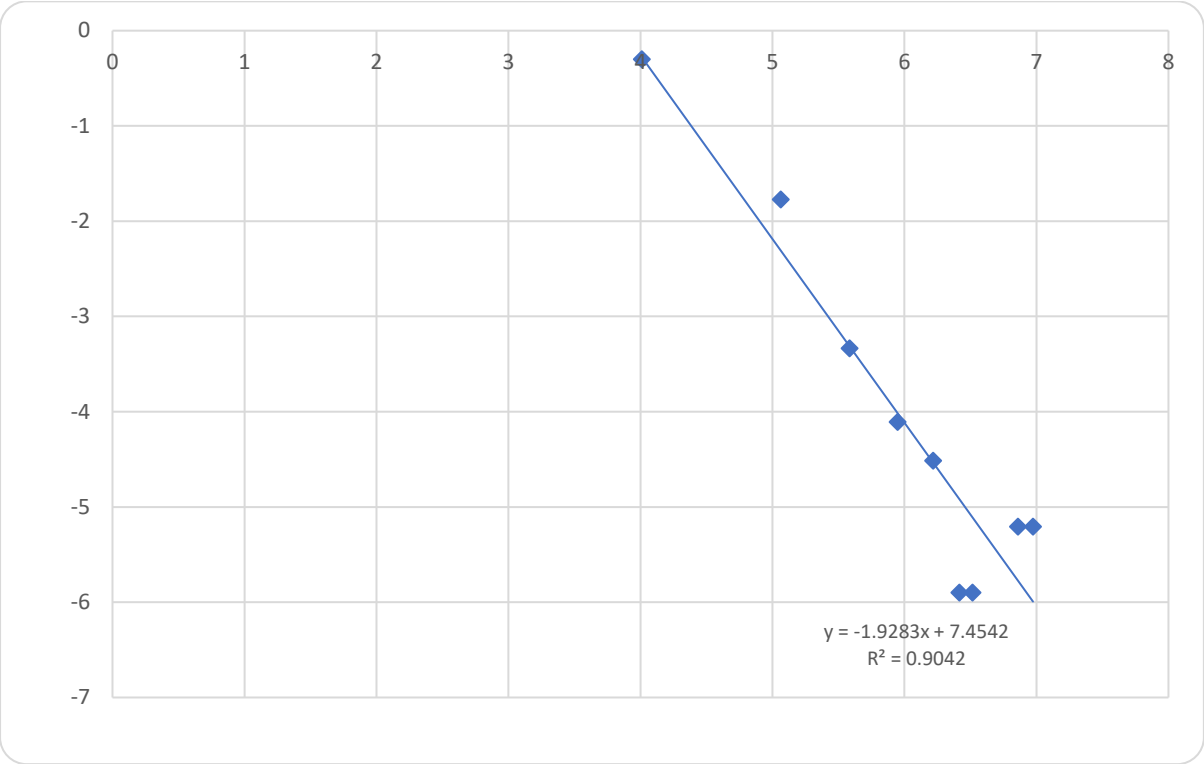


Appendix 3.3: Hill plot for realised variance based on 60 squared daily returns

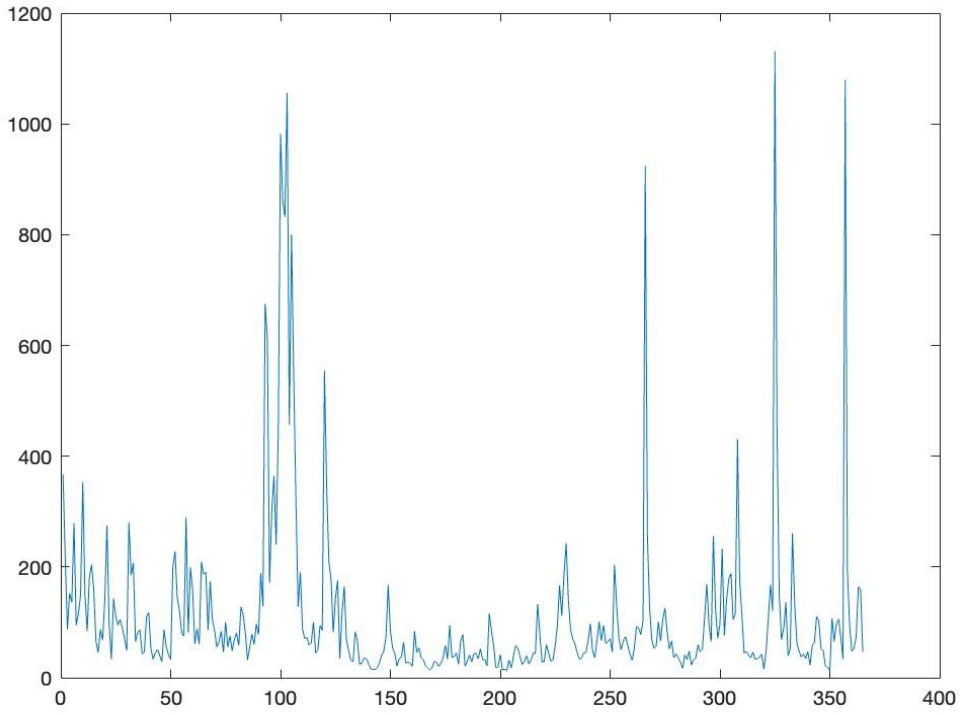


Appendix 4: 90 squared daily realised returns resolution

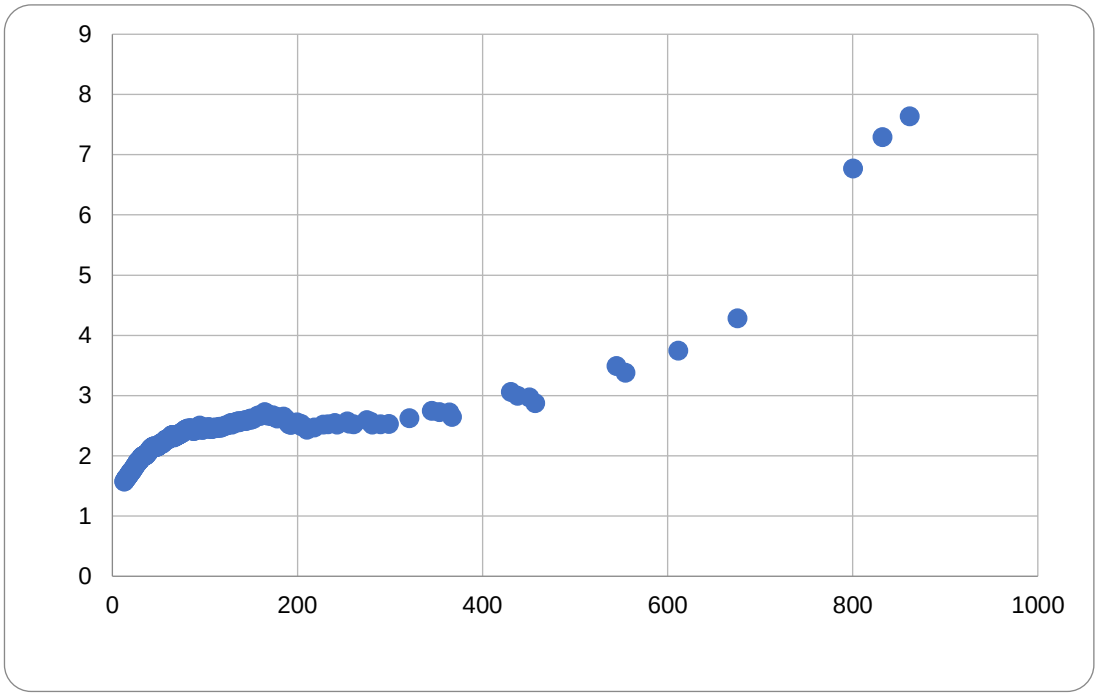
Appendix 4.1: Log-log OLS regression technique for 90 squared daily realised returns



Appendix 4.2: Visual inspection of 90 daily squared returns resolution

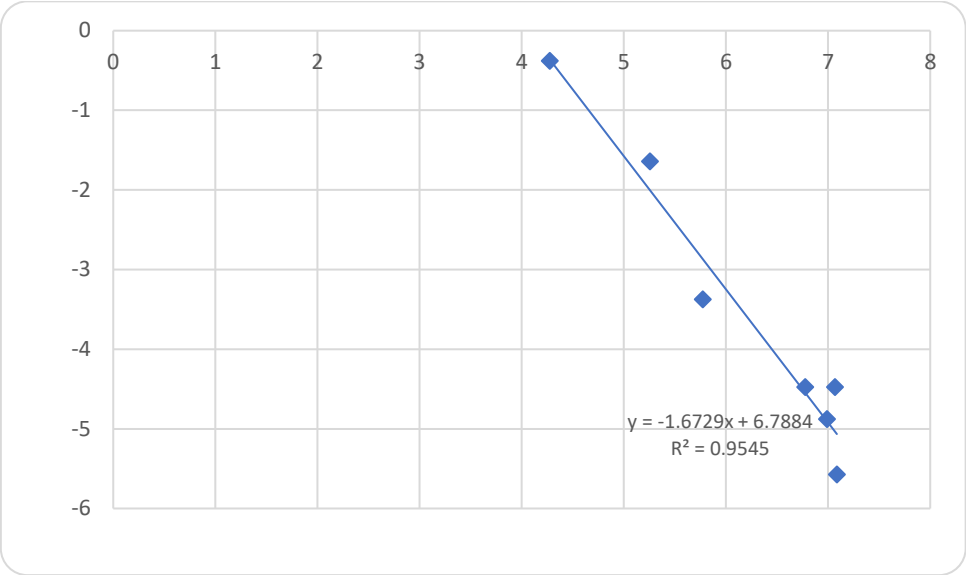


Appendix 4.3: Hill plot for realised variance based on 90 squared daily returns

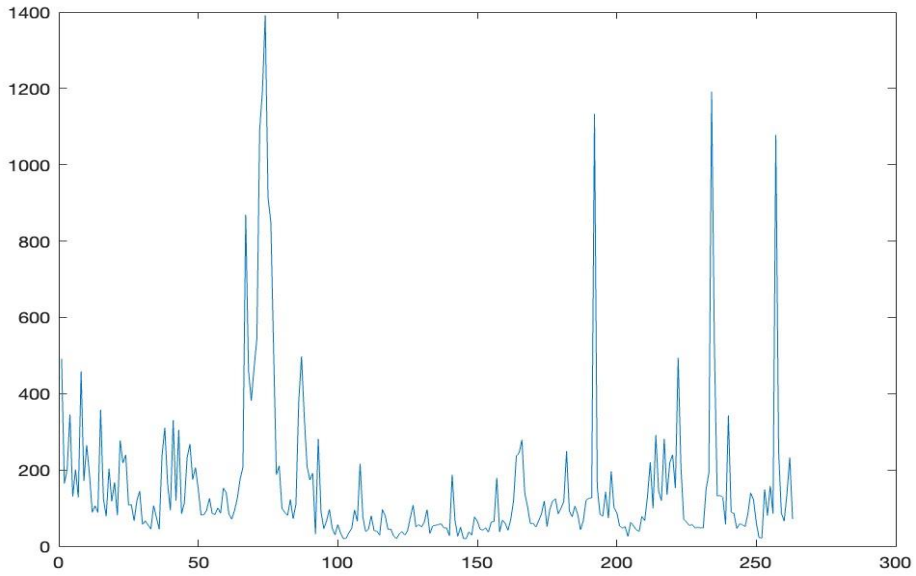


Appendix 5: 125 squared daily realised returns resolution

Appendix 5.1: Log-log OLS regression technique for 125 squared daily realised return



Appendix 5.2: Visual inspection of 125 daily squared returns resolution



Appendix 5.3 Hill plot for realised variance based on 125 squared daily returns