



## Full Length Article

## What do we know about the second moment of financial markets?☆

Klaus Grobys

School of Accounting and Finance, University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland

## ARTICLE INFO

## JEL classification:

G12  
G14

## Keywords:

Finance  
Pareto distributions  
Power laws  
Second moment  
Variance  
Variance of variance

## ABSTRACT

Recent research shows that the vast majority of scientific studies published in leading finance journals fails scientific replication (Hou, Xue, and Zhang, 2020; Harvey, Liu, and Zhu; 2016). This study argues that p-hacking, publication pressure and the selection bias from leading finance journals are perhaps not the underlying root cause for this issue. This study shows that standard methodologies often used in finance research are inevitably sample-specific due to the very nature of financial markets. While the consensus of earlier research postulates a rejection of the time-honored Levy hypothesis, the results of this study strongly indicate that the variance of variance does not exist in any of the financial key markets that are considered. An unexpected finding of this study is that the variance process governing the U.S. dollar foreign exchange rate market is generating more extreme events than the Bitcoin market. The results cast doubts on the validity of methodologies currently used in finance research.

*“Truth – or more precisely, an accurate understanding of reality – is the essential foundation of any good outcome. Most people fight seeing what’s true when it’s not what they want it to be.”*

(Ray Dalio, Founder of Bridgewater Associates, and Author of ‘Principles’)

## 1. Introduction

In a recent study, Hou, Xue, and Zhang (2020) investigate whether cross-sectional asset pricing phenomena documented in the finance literature hold up to currently acceptable standards for empirical finance. In their study, the authors implement a scientific replication of 452 asset pricing anomalies which are based on 111 original research papers.<sup>1</sup> Imposing the higher multiple test hurdle of 2.78, as proposed by Harvey, Liu, and Zhu (2016), they find that 82% of those anomalies fail scientific replication. This is shocking news.

Referring to a recent study of Serra-Garcia and Gneezy (2021), *The*

*Guardian* documents that “scientific research findings that are probably wrong gain far more attention than robust results, according to academics who suspect that the bar for publication may be lower for papers with grabbier conclusions. Studies in top science, psychology and economics journals that fail to hold up when others repeat them are cited, on average, more than 100 times as often in follow-up papers than work that stands the test of time. The finding – which is itself not exempt from the need for scrutiny – has led the authors to suspect that more interesting papers are waved through more easily by reviewers and journal editors and, once published, attract more attention.”<sup>2</sup> The selection bias from leading (finance) journals is perhaps not surprising. The surprising issue is that top-notch journals are also the leading journals in publishing studies that are not replicable.

The important question arises: Why do studies fail to replicate? Considering the background of the agency problem, as elaborated in Harvey’s (2017) study, Hou et al. (2020) argue that authors sometimes engage in specification search, selecting sample criteria and test procedures until insignificant results become significant (p-hacking) which could, in turn, result in an embarrassingly large number of false positives

☆ I would like to thank the editor Brian Lucey and two anonymous reviewers for insightful comments.

E-mail address: [klaus.grobys@uva.fi](mailto:klaus.grobys@uva.fi).

<sup>1</sup> Table 3 in association with the reference list provided in Hou et al. (2020) reveals that 58% of those studies have been published in the leading finance journals (*Journal of Finance*, *Review of Financial Studies*, *Journal of Financial Economics*). Here, it is referred to leading finance journals to any journal that is first (i) categorized as a ‘finance journal’, and second (ii) exhibits a journal ranking of ABS-4\* in line with the Academic Journal Guide of the Chartered Association of Business Schools (ABS).

<sup>2</sup> See <https://www.theguardian.com/science/2021/may/21/research-findings-that-are-probably-wrong-cited-far-more-than-robust-ones-study-finds?fbclid=IwAR3m7yHKFOW8RMzx1La38mlR8fi9lnkM7lkWjOOilM3nVUuGBYeHP9-zeTU>.

that cannot be replicated in the future.<sup>3</sup> This study argues that there could be another reason for why studies fail to replicate. Specifically, if the variance of the data generating process's variance does not exist, a  $t$ -statistic will be inevitably sample-specific. Implicitly, this supposition is in line with Schwert (2003), who observes that once cross-sectional asset pricing phenomena are documented and analyzed in the academic literature, these cross-sectional patterns often seem to disappear, reverse, or attenuate. Specifically, decreasing  $t$ -statistics with increasing sample sizes across anomalies could be a manifestation of systematically inflated  $t$ -statistics. In this regard, Schwert (2003) emphasizes that one possibility could be that those asset pricing anomalies were simply statistical aberrations that attracted the attention of academics and practitioners. This study argues that those 'statistical aberrations', in the parlance of Schwert (2003), are manifestations of researchers relying on traditional research methodologies that unfortunately are not applicable in financial research contexts due to the very nature of financial market data.<sup>4</sup>

The ultimate purpose of this study is to test whether research methodologies often used in traditional financial research, such as Ordinary Least Squares (OLS) or Generalized Method of Moments (GMM), are applicable to financial market data. The rationale for setting-up the research design of this study is straight forward: First (i), it is noted that most research in empirical finance typically relies on  $t$ -statistics derived from, for instance, OLS or GMM, for evaluating the validity of results. Second (ii), irrespective of (a) which type of research method is used, or (b) which type of  $t$ -statistics are used, any  $t$ -statistic can only be used for drawing statistical conclusions *if and only if* the variances of the model variables do exist.<sup>5</sup> This is definitely not a trivial issue because if the variance of variance is either infinite or does not even exist, we are not in a research environment allowing us to draw conclusions based on (any)  $t$ -statistics because this metric would be inevitably sample-specific. Hence, to explore this research question, the current study analyzes whether or not the second moments of five key financial market variables are stable. Using a research approach based on realized variances, a financial market variable is defined as stable *if and only if* the variance of the variance does exist. Since realized variances are heavily fat-tailed processes, this study follows a recent stream of literature and fits power law distributions to the realized variances of the following key financial markets:

<sup>3</sup> Note that cherry picking may be committed intentionally or unintentionally. Moreover, in practice, it is perhaps difficult (or virtually impossible) to differentiate between those two groups. In fact, one can only observe the consequences and the consequences are unfortunately the same.

<sup>4</sup> Correctly using incorrect methodologies would be alarming news because Hou et al. (2020) stress out that armies of academics and investment managers actively engage in searching for significant anomalies and with trillions of dollars invested in factors-based exchange-traded funds and quantitative hedge funds worldwide, the financial interest is overwhelming and in this study it is argued, so is the hidden risk. Referring to the bankruptcy of the hedgefunds Long-Term Capital Management (LTCM), in which Robert Merton Jr. and Myron Scholes were founding partners, Taleb (2010, p.288) points out that the consequences of relying on wrong methods can be destructive: "...during the summer of 1998, a combination of large events, triggered by a Russian financial crisis, took place that lay outside their models. It was a Black Swan. LTCM went bust and almost took down the entire financial system with it, as the exposures were massive."

<sup>5</sup> The most often used method is perhaps the Ordinary Least Squares (OLS) technique used in different settings (time series regressions, cross-sectional regressions, panel regressions). Since OLS estimation requires some strict assumptions, Hansen (1982) derived the so-called Generalized Method of Moments (GMM) estimator that relaxes many of the OLS assumptions. However, both the OLS and GMM estimators require that the kurtoses of the input variables exist. Moreover, in attempts to address dependency structures in the first and/or second moment of the (financial) variables used in the estimation procedures, various adjustments such as Heteroscedasticity Consistent Covariance Matrix Estimator (HCCME), Heteroscedasticity and Autocorrelation Consistent Covariance Estimator (HAC), or some type of bootstrapped  $t$ -statistics have been discussed in the literature (see White, 1980; Newey-Newey & West, 1987; Godfrey, 2009).

equities, commodities, currencies and cryptocurrencies. To test whether or not the power law null hypothesis is plausible, hypothesis tests based on Kolmogorov-Smirnov distances are employed, as proposed in the seminal paper from Clauset, Shalizi, and Newman (2009). Moreover, also various subsamples are considered, different types of data frequencies and simulation experiments.

This study has some clear and fundamentally important contributions. First, while Hou et al. (2020) argue that 'cherry-picking' manifested in specification search or p-hacking could be a reason for the high rate of replication failures, this study adds to Schwert's (2003) 'statistical aberration' hypothesis by exploring whether data processing could be a root cause for the high rate of replication failures. Second, taking a broader perspective, this paper contributes to the literature on tail risks that appear to be a trademark of human-engineered systems. In this regard, the study of Clauset et al. (2009) is often-cited work exploring whether 24 real-world data sets from a range of different disciplines follow power law distributions. The evidence documented in Clauset et al. (2009) supports Taleb's (2010) view that power law distributions govern many real world phenomena and help to better understand man-made phenomena. Another popular study in this stream of literature is the one of Gabaix (2009) documenting that a variety of variables such as income and wealth, the size of cities and firms, trading volume, international trade or executive pay, for instance, are governed by different power law processes. The current study contributes to this stream of literature by first (i) exploring whether the variances of five key financial market variables are governed by power laws, and second (ii) by identifying whether the second moments of the variances exist.

Third, taking the perspective of more finance-specific research, power law distributions are used to model the return variation of financial assets. Since power laws are one-sided distributions, it may be not surprising that most research uses the absolute amount of an asset return, that is,  $|ret|$  for modeling power law functions, as pointed out in Lux and Alfarano (2016). Based on the seminal paper from Mandelbrot (1963), early contributions in this stream of literature are Gopikrishnan, Plerou, Amaral, Meyer, and Stanley (1999), Jansen and de Vries (1991), Mantegna and Stanley (1995) and Lux (1996). The studies from Gabaix (2009) and Lux and Alfarano (2016) provide detailed overviews on that literature. The current study extends this strand of literature first by modeling the variation of asset returns using realized variances computed using daily high and low prices which incorporates more information than two arbitrary points in the data series (the closing prices). It will be seen later in this study that this is not a trivial issue. Additionally, the current research makes use of a realized volatility measure based on daily data to compute monthly realized variances. Another novel feature of this paper is that it also investigates the variance of the largest cryptocurrency market, that is, Bitcoin exhibiting a market capitalization in excess of \$1 trillion as of April 29, 2021. In this regard, Fry and Cheah (2016, p.350) highlight that "from an economic perspective the sums of money involved [in cryptocurrency markets] are substantial". Obviously, studying the variation of cryptocurrency prices is both an important and timely issue.

The results of the current research indicate that the daily variances of all five key asset markets are governed by power law processes. Statistically, the power law null hypothesis cannot be rejected. It is shown that the results are neither sample- nor method-specific. Notably, the findings of the current research strongly suggest that the variance of the variance does statistically not exist for any of those asset markets. Paradoxically, the foreign exchange market is more prone to extreme events than the Bitcoin market. The findings of this study have fundamentally important implications that cannot be swept under the carpet: First, due to the non-existence of the variances' variances, standard statistical analysis based on standard OLS or GMM inevitably leads to sample-specific results. Second, simulation experiments show that in 66% of synthetic samples, the sample variances are underestimated which results in inflated  $t$ -statistics, provided data samples are finite. Seemingly significant asset pricing phenomena due to inflated  $t$ -statistics could be one manifestation

of what Schwert (2003) refers to as ‘statistical aberration’. As a consequence, the results of this study cast doubts on the validity of methodologies often-used in financial market research.

And here is how the story of this paper unfolds: In the next section the background is described. The third section describes the processing of the data, whereas the fourth section outlines the statistical model, and provides robustness checks. The last section concludes.

## 2. Background

Traditional finance research typically relies on *t*-statistics to assess the statistical significance of research results. The *t*-distribution is defined as,

$$t = \frac{x}{\sqrt{\frac{z}{v}}} \tag{1}$$

where *x* is a normally distributed random variable with  $E[x] = \mu_x$  and  $VAR[x] = \sigma_x^2$ , and *z* is a  $\chi^2(v)$ -distributed random variable with  $E[z] = \mu_z = v$  and  $VAR[z] = \sigma_z^2 = 2v$ . The crucial point here is to note that convergence requires that the variance of the random variable *z* must exist. In empirical finance, the estimated *t*-statistic for a specific sample is then given by,

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

where it is typically tested  $\mu_x = 0$ . From the previous definitions, it follows that Eq. (2) implies that,

$$t = \frac{\bar{x} - \mu_x}{\sqrt{\frac{\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2}{\sqrt{T}}}} = \frac{\bar{x} - \mu_x}{\frac{\sqrt{\sigma_x^2}}{\sqrt{T}}} \tag{3}$$

From Eq. (1) in association with  $z \sim (\mu_z, \sigma_z^2)$ , it follows from Eq. (3) that,

$$\frac{\bar{x} - \mu_x}{\frac{\sqrt{\sigma_x^2}}{\sqrt{T}}} \Rightarrow \frac{\bar{x} - \mu_x}{\frac{\sqrt{\mu_z}}{\sqrt{T}}}$$

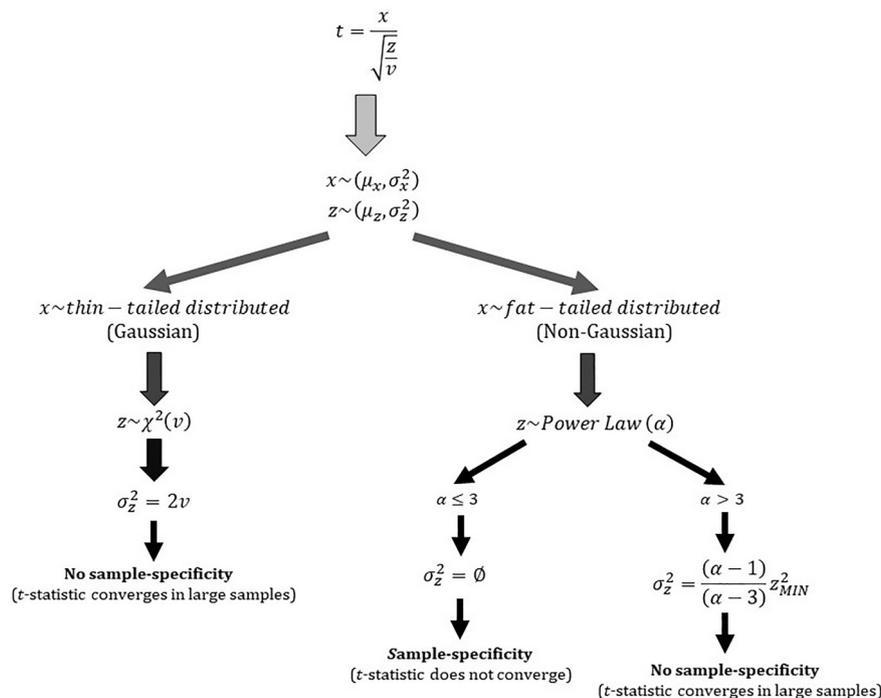
The empirical sample’s variance of variance is then given by,

$$\hat{\sigma}_z^2 = \frac{1}{T} \sum_{t=1}^T ((x_t - \bar{x})^2 - \sigma_x^2)^2. \tag{4}$$

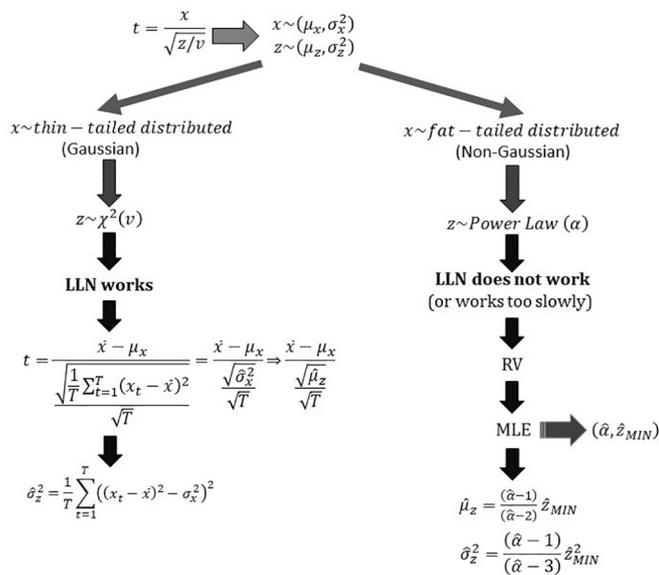
Given that *x* is a thin-tailed (Gaussian) process,  $\hat{\sigma}_z^2$  must consequently be finite. It is important to note here that this often-used methodology requires first (i) that the *Central Limit Theorem* (CLT) holds, and second (ii) that the *Law of Large Numbers* (LLN) works. The idea is illustrated on the left-hand side of Figs. 1 and 2.

However, it is a well-known and stylized fact that financial returns are typically not normally-distributed. Accounting for fat tails, a wide stream of literature models the absolute amount of asset returns using power law functions (Mandelbrot, 1963; Fama, 1963; Gopikrishnan, Plerou, Amaral, Meyer, and Stanley, 1999; Jansen & de Vries, 1991; Mantegna & Stanley, 1995; Lux, 1996; Warusawitharana, 2019). In this regard, Taleb (2020, p.91) argues that “[t]here 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.” Concerning the application of power laws to financial market data, Lux and Alfaro (2016, p. 4) emphasize that “power laws in returns and in volatility seem to be intimately related: none of them was ever observed without the other and it, therefore, seems warranted to interpret them as the joint essential characteristics of financial data.”

If *x* is a (heavily) fat-tailed process (i.e., non-Gaussian), we cannot use  $\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2$  because Taleb (2020, p.27–28) stresses out that the LLN does either not work or works too slowly. Since maximum likelihood estimation (MLE) is a valid estimation method for fat-tailed data, Taleb advocates to use the so-called ‘plug-in estimation’ technique, where in the first step (i) the power law exponent is estimated via MLE, and in the second step (ii) the theoretical moments for the corresponding power law distribution are estimated. Hence, using realized variances (RV) one can use MLE to reliably estimate the power law process governing the distribution of *z*. Unlike employing the term  $\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2$  to estimate the sample return variance, employing RV computed via the price range of intraday asset prices has the advantage that substantially more information is incorporated in the data than using only two arbitrary points in this series (the closing prices), as



**Fig. 1.** When is the *t*-statistic sample-specific? This flow chart illustrates under which conditions the *t*-statistic is sample-specific.



**Fig. 2.** When does the Law of Large Numbers not work? This flow chart illustrates how to estimate the variance of the variance depending on whether or not the Law of Large Numbers (LLN) works. In this flow chart, the abbreviations RV and MLE denote realized variance and maximum likelihood estimation, respectively.

pointed out in Chou, Chou, and Liu (2010).

Deriving the moments for power law distributions (see the next sections for details), it follows that the variance for  $z$  is defined if and only if the power law processes governing  $z$ 's distribution exhibits a power law exponent larger than three. An undefined variance for  $z$  implies, in turn, that one does not observe the true mean for  $z$  in finite sample, which results in a sample-specific  $t$ -statistic, as illustrated on the right-hand side of Figs. 1 and 2. Note that the same logic can be applied to regression model frameworks often used in finance research.

For instance, in asset pricing research, the statistical significance of cross-sectional asset pricing phenomena is typically assessed using regression models fulfilling the purpose of adjusting asset returns for potential risk factor exposures. In this regard, the Fama and French factor models have received enormous attention in academic work and are often used as benchmark models. Recently, Fama and French (2018) proposed a six-factor model, that accounts for the momentum factor, given by,

$$R_{i,t}^{ex} = a_i + b_i Mkt_t^{ex} + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + m_i MOM_t + u_{i,t}, \quad (5)$$

where  $R_{i,t}^{ex}$  typically denotes the excess return of an equity portfolio  $i$  at time  $t$ ,  $Mkt_t^{ex}$  denotes the return of the excess market factor at time  $t$ ,  $SMB_t$  and  $HML_t$  denote the returns of the size and value factor at time  $t$  (Fama & French, 1992, 1993),  $RMW_t$  and  $CMA_t$  denote the returns of the profitability and investment factor at time  $t$  (Fama & French, 2015, 2017), and  $MOM_t$  denotes the return of the momentum factor at time  $t$  (Fama & French, 2018). It is important to understand that similar types of factors models are also widely used for assessing the significance of risk-adjusted returns in other financial assets markets, such as traditional foreign exchange markets or cryptocurrency markets (Lustig, Roussanov, & Verdelhan, 2011; Shen, Urquhart, & Wang, 2020). Hence, the same issues discussed here for the equity market apply to any other financial asset market also.

Following standard econometric modeling, one can stack the risk factors, including a Tx1 vector of ones denoted as  $\mathbf{1}$ , in a regressor matrix  $\mathbf{X}$  defined as,

$$\mathbf{X} = (\mathbf{1}, \mathbf{x}_{Mkt}^{ex}, \mathbf{x}_{SMB}, \mathbf{x}_{HML}, \mathbf{x}_{RMW}, \mathbf{x}_{CMA}, \mathbf{x}_{MOM}),$$

that has the dimension  $T \times 7$ . Denoting the  $T \times 1$  vector of portfolio excess returns as  $\mathbf{R}^{ex}$  and the  $T \times 1$  residual vector as  $\mathbf{u}$ , one knows from standard econometrics classes that using Ordinary Least Squares (OLS), the

estimated covariance matrix of the point estimator  $\hat{\beta}$  is given by

$$COV(\hat{\beta}) = \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}, \quad (6)$$

where  $\hat{\sigma}^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}}/T$ ,  $\hat{\mathbf{u}} = \mathbf{R}^{ex} - \mathbf{X}\hat{\beta}$  with  $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}^{ex}$  where  $\hat{\beta} = (a_i, b_i, s_i, h_i, r_i, c_i, m_i)$ . Furthermore, in standard econometrics it is typically assumed that  $u_{i,t} \sim IIDN(0, \sigma^2)$  and, hence,  $COV(\hat{\beta})$  is normally distributed also. A severe problem in this model framework arises if  $\sigma^2$  is sample-specific because if  $\sigma^2$  is sample-specific then  $COV(\hat{\beta})$  will be sample-specific too, and critical values derived from the normal distribution will be meaningless. For instance, Fergusson and Platen's (2006) study provides evidence on that the unconditional distribution of daily returns appears to be remarkably close to the Student  $t$  distribution with  $v = 3$  degrees of freedom. Assuming that  $u_{i,t} \sim t(3)$ , it follows that  $E_T[u_{i,t}^2] = 3$ , whereas  $E_T[u_{i,t}^3] = \infty$ .

In turn, an infinite kurtosis implies that the variance is not stable. In this regard, Taleb (2020, p.50) vehemently stresses out that if the kurtosis does not converge, "the sample error is huge; or it may not exist so the measurement is heavily sample dependent. If we don't know anything about the fourth moment, we know nothing about the stability of the second moment. It means we are not in a class of distribution that allows us to work with the variance, even if it exists." Note that the same statement can be made for the variance of variance, that is, if we don't know anything about the variance of the variance, we know nothing about the stability of the second moment. Using a simulation experiment, later will be learnt more on the implications of this issue.

Furthermore, Hansen (1982) proposes the Generalized Method of Moments (GMM) estimation technique that relaxes some of the standard OLS assumptions. In this regard, the key point in GMM estimation is that the employed variables exhibit ergodic stationarity which, again, implies that the fourth moment of the variables must be finite. If the variance of the variance does not exist, GMM is, in turn, as sample-specific as OLS.

While the Student's  $t$  distribution could be an interesting approximation for the unconditional distribution of daily returns in equity market settings, it is probably not able to capture the extremely fat tails of variance distributions, especially the variances in cryptocurrency markets which appear to exhibit extremely high levels of uncertainty (Baur, Hong, & Lee, 2018). Hence, in what follows, the variance processes of five key asset markets is modeled as power laws and it is tested whether these models reasonably describe the data generating processes.

### 3. Data

Publicly available daily data on the S&P 500, gold, crude oil, the exchange rate U.S. dollar against British pounds, and Bitcoin are downloaded from [finance.yahoo.com](http://finance.yahoo.com). Due to data availability, the data sample ranges between September 17, 2014 to March 31, 2021 for Bitcoin, and April 20, 1982 to March 31, 2021 for the S&P 500. Especially, the data include the highest daily prices, lowest daily prices, and closing prices for each trading day of the samples.

### 4. Methodology

#### 4.1. Realized variance

Realized variances are computed for each asset market  $i$ , where  $= \{equity, gold, crude\ oil, U. S. dollar, cryptocurrency\}$ . Specifically, realized annualized daily variances are compounded in line with Parkinson (1980), that is,

$$\sigma_{i,t}^2 = T \frac{1}{4 \ln(2)} (\ln(H_{i,t}) - \ln(L_{i,t})) \quad (7)$$

where  $H_{i,t}$  and  $L_{i,t}$  denote the highest and lowest price for asset market  $i$

**Table 1**  
Descriptive statistics.

Metric	S&P 500	Gold	Crude oil	U.S. dollar	Bitcoin
Mean	2.15	0.16	14.92	2.39	48.89
Median	0.89	0.05	6.95	0.41	15.53
Maximum	473.02	10.10	1,871.01	7,356.17	3,153.47
Minimum	0.00	0.00	0.00	0.00	0.09
Std. Dev.	6.99	0.40	49.72	109.83	120.75
Skewness	35.46	9.74	20.27	66.95	10.87
Kurtosis	2145.24	153.30	560.45	4483.45	211.39
Observations	9,821	5,131	5,139	4,486	2,384

This table reports the descriptive statistics for the annualized daily realized variance for the S&P 500, gold, crude oil, the exchange rate of the U.S. dollar against the British pound and Bitcoin. The annualized daily realized variances for each asset market  $i$ , where  $i = \{equity, gold, crude\ oil, U. S. dollar, cryptocurrency\}$  are in line with [Parkinson \(1980\)](#) computed as

$$\sigma_{i,t}^2 = T \frac{1}{4 \ln(2)} (\ln(H_{i,t}) - \ln(L_{i,t})),$$

where  $H_{i,t}$  and  $L_{i,t}$  denote the highest and lowest price for asset market  $i$  on day  $t$ ,  $\sigma_{i,t}^2$  denotes asset market  $i$ 's corresponding realized annualized variance, and  $T = 365$  for the cryptocurrency market because this market allows for trading 24/7, whereas  $T = 250$  for any other asset market. Publicly available daily data on the S&P 500, gold, crude oil, the exchange rate U.S. dollar against British pounds, and Bitcoin were retrieved from [finance.yahoo.com](https://finance.yahoo.com).

on day  $t$ ,  $\sigma_{i,t}^2$  denotes asset market  $i$ 's corresponding realized annualized variance at time  $t$ , and  $T = 365$  for the cryptocurrency market because this market allows for trading 24/7, whereas  $T = 250$  for any other asset market. Since the [Parkinson \(1980\)](#) estimator uses the price range of intraday asset prices, [Chou et al. \(2010\)](#) emphasize that it incorporates substantially more information than two arbitrary points in this series (the closing prices).<sup>6</sup> When computing the realized variance for crude oil, the observations on April 20, 2020 and April 21, 2020 are excluded from the sample because the lowest prices were negative on those days, and hence, the realized variances in line with Eq. (7) are not defined.<sup>7</sup>

In [Table 1](#), the descriptive statistics are reported and [Table 2](#) reports the share of the top 1% and the top 20% of the cumulative total of the distribution. From [Table 1](#) in association with [Table 2](#) it can be observed that the realized variance processes for all asset markets are heavily fat-tailed, which is not a surprising feature per se.

However, a surprising observation from these tables is that the variance process governing the U.S. dollar foreign exchange rate market – which obviously is the exchange rate market of the most important national currency – is considerably heavier fat-tailed than the cryptocurrency market. This is surprising because [Baur et al. \(2018\)](#) argue that Bitcoin returns exhibit an extremely high kurtosis with relatively more tail events compared to other assets, and therefore Bitcoin serves rather as speculative asset than as a medium of exchange. On January 27, 2012, the exchange rate U.S.\$/U.K.£ ranged between the highest value corresponding to 1.57 and the lowest value corresponding to 0.64; that is, on the same day, the exchange rate dropped by 41%. For comparison, the largest daily drop in the history of the S&P 500 occurred on October 19, 1987 where the S&P 500 ranged between the highest index value corresponding to 282.70 and the lowest index value corresponding to 224.83 - a relative drop by 20%. One can see that rare events have a considerably stronger impact in the U.S. dollar foreign exchange market than in the U.S. equity market.

Next, from [Table 2](#) it can be observed that 1% of the largest observations in the realized variance process of the U.S. dollar foreign

<sup>6</sup> I acknowledge that other price range estimators have been discussed in the literature. [Shu and Zhang \(2006\)](#) investigated the relative performance of the various range-based volatility estimators and conclude that the performance of all of them is very well.

<sup>7</sup> The high and low price for crude oil were \$17.85 and -\$40.32 (\$13.86 and -\$16.74) on April 20, 2020 (April 21, 2020).

**Table 2**  
Share of the top 1% and top 20%.

% of largest observations	% of the cumulative total of the distribution				
	S&P 500	Gold	Crude Oil	U.S. dollar	Bitcoin
1%	21.01% (98)	19.31% (51)	22.96% (51)	73.50% (45)	18.34% (24)
20%	69.05% (1,964)	74.28% (1,026)	65.71% (1,028)	88.47% (897)	75.13% (477)

This table reports the share of the top 1% and the top 20% of the cumulative total of the distribution for the annualized daily realized variance for the S&P 500, gold, crude oil, the exchange rate of the U.S. dollar against the British pound, and Bitcoin. The number of observations are reported in parentheses.

exchange rate market correspond to 73.50% of the cumulative total of observations, whereas the share of the top 1% of the realized variance processes of the other asset markets comprises between 18.34% (Bitcoin market) and 22.96% (Crude oil market) of the cumulative total of observations. Comparing these figures with [Table 3](#) in [Taleb \(2010, p.265\)](#) strongly suggests a Paretian tail with power law exponents close to 2.5.

In fact, the variance of the U.S. dollar foreign exchange rate market considered here is the most extreme process in terms of its Paretian tail. Specifically, the traditional 80/20 Pareto distribution – which is the archetype of a power law process – suggests that 20% of the largest observations comprise 80% of the cumulative total of observations. As pointed out in [Taleb \(2010, p.235\)](#) according to its scalability, this suggests in turn, that 1% of the largest observations comprise about 50% of the cumulative total of observations. With respect to the foreign exchange market's variance considered here, this feature is even more pronounced. There is no other distribution class than power laws that allows for this type of extremeness in fat tails as observed here. A fundamental follow-up question that arises is then: If Bitcoin does not fulfill the requirements as being a medium for exchange or store of value due to its high uncertainty, how can the U.S. dollar be considered stable? Since the 80/20 Pareto distribution does not have a variance, any

**Table 3**  
Estimates for power law models.

Metric	S&P 500	Gold	Crude Oil	U.S. dollar	Bitcoin
$\hat{\alpha}$	2.58	2.66	2.48	2.60	3.02
$\hat{x}_{MIN}$	6.36	0.51	22.42	2.76	224.33
KS test (p-value)	0.5830	0.3180	0.5110	0.5090	0.1660
$N$	9,821	5,131	5,139	4,486	2,384
$N_{PL}$	5.96%	6.72%	12.22%	3.66%	4.78%
Period (MM/DD/YYYY)	4/20/ 3/31/ 2021	8/30/ 3/31/ 2021	8/23/ 3/31/ 2021	12/1/ 3/31/ 2021	9/17/ 3/31/ 2021

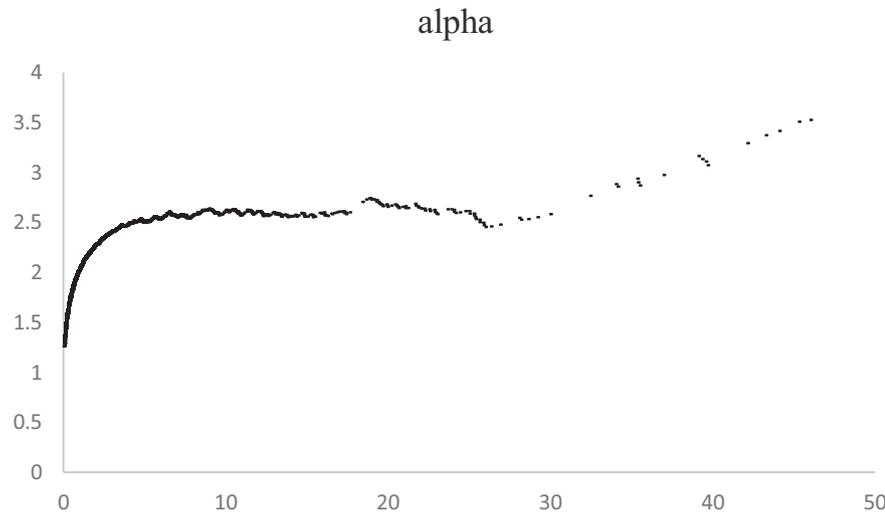
This table reports the estimates for power law models  $p(x) = (\alpha - 1)x_{MIN}^{\alpha-1}x^{-\alpha}$  using maximum likelihood estimation (MLE). The tail exponent  $\alpha$  is estimated as,

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

where  $\hat{\alpha}$  denotes the MLE estimator and  $N$  denotes the number of observations, provided  $x_i \geq x_{MIN}$ . In this model, the estimate  $\hat{x}_{MIN}$  is assessed via the Kolmogorov-Smirnov or KS statistic, which is the maximum distance between the CDFs of the data and the fitted model:

$$D = \text{MAX}_{x \geq x_{MIN}} |S(x) - P(x)|,$$

where  $S(x)$  is the CDF of the data for the observation with value at least  $x_{MIN}$ , and  $P(x)$  is the CDF for the power law model that best fits the data in the region  $x \geq x_{MIN}$ . The estimate  $\hat{x}_{MIN}$  is the value of  $x_{MIN}$  that minimizes  $D$ .  $N_{PL}$  denotes the percentage of sample observations that are governed by a power law process.



**Fig. 3.** Hill plot for the S&P 500 variance.

This figure shows the Hill plot for the S&P 500 variance. The Hill plot shows the estimated  $\hat{\alpha}$  as a function of  $x_{MIN}$  defining the minimum value of the variance that is governed by the power law, given by the maximum likelihood estimator (MLE),

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

where  $\hat{\alpha}$  denotes the MLE estimator,  $x_i$  is the annualized daily realized variance of the S&P 500, provided  $x_i \geq x_{MIN}$ , and  $N$  denotes the number of observations for which  $x_i \geq x_{MIN}$  is satisfied. (Note for improved visualization, the graph is cut off at observation  $x_{0,787}$  which is common practice.)

estimated  $t$ -statistic is consequently sample-specific. Next, power laws are fitted to the variance processes of the five key asset markets and then the power law null hypothesis is tested.

#### 4.2. Statistical model

##### 4.2.1. Moments of power laws

To investigate the stability of the realized variance processes, the realized variances are modeled using the following power laws:

$$P(X > x) = p(x) = Cx^{-\alpha}, \tag{8}$$

where  $C = (\alpha - 1)x_{MIN}^{\alpha-1}$  with  $\alpha \in \{\mathbb{R}_+ | \alpha > 1\}$ ,  $x \in \{\mathbb{R}_+ | x_{MIN} \leq x < \infty\}$ ,  $x_{MIN}$  is the minimum value of realized variance that is governed by the power law process, and  $\alpha$  is the magnitude of the specific tail exponent.<sup>8</sup> Regarding  $\alpha$ , Taleb (2020, p. 34) observed that the tail exponent of a power law function captures via extrapolation the low-probability deviation not seen in the data, which plays a disproportionately large share in determining the mean. Using the model framework of the current research, it can be shown that the expectation of the variance defined as  $E[X]$  is given by

$$E[X] = \int_{x_{MIN}}^{\infty} xp(x)dx = \frac{(\alpha - 1)}{(\alpha - 2)}x_{MIN}, \tag{9}$$

whereas the second moment  $E[X^2]$ , or the variance of the variance, is defined as:

$$E[X^2] = \int_{x_{MIN}}^{\infty} x^2p(x)dx = \frac{(\alpha - 1)^2}{(\alpha - 3)}x_{MIN}^2. \tag{10}$$

Higher moments of order  $k$  are analogously defined as:

$$E[X^k] = \frac{(\alpha - 1)}{(\alpha - 1 - k)}x_{MIN}^k. \tag{11}$$

From Eq. (9), one knows that the mean only exists for  $\alpha > 2$ , whereas the variance only exists for  $\alpha > 3$ .

##### 4.2.2. Maximum-likelihood estimation

In line with White, Enquist, and Green (2008) and Clauset et al. (2009), who found that maximum likelihood estimation (MLE) performs best for

estimating power law exponents, the tail exponent is estimated as:

$$\hat{\alpha} = 1 + N \left( \sum_{i=1}^N \ln \left( \frac{x_i}{x_{MIN}} \right) \right)^{-1}, \tag{12}$$

where  $\hat{\alpha}$  denotes the MLE estimator,  $N$  is the number of observations exceeding  $x_{MIN}$ , and other notation is as before. Figs. 3–7 plot the estimated parameters for  $\hat{\alpha}$  depending on the value for  $x_{MIN}$  for all five asset market variances.<sup>9</sup> A crucial issue is how to determine the corresponding values for  $\hat{\alpha}$  and  $x_{MIN}$  to accurately estimate the probability density functions. Clauset et al. document that it is common practice to choose the value for  $x_{MIN}$ , where beyond which  $\hat{\alpha}$  is stable. From Fig. 5 in Clauset et al. (2009, p. 670), it can be observed that this value corresponds to the saddle point in a  $\hat{\alpha}/x_{MIN}$ -graph. From the Hill plots (see Figs. 3–7) it is evident that for most asset markets variances  $\hat{\alpha}$  appears to be stable below a value of three. However, it is not clear which  $x_{MIN}$  is optimal. (See Fig. 8

##### 4.2.3. Kolmogorov–Smirnov test statistics

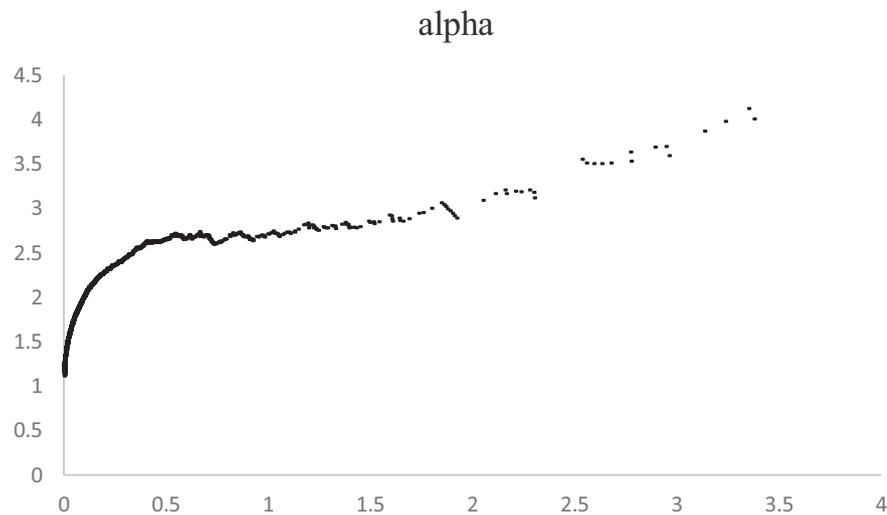
Determining the exact value for  $x_{MIN}$  is, however, not a trivial issue. Clauset et al. emphasize that if one chose too low a value for  $x_{MIN}$ , one would get a biased estimate of  $\hat{\alpha}$  since one will be attempting to fit a power law model to non-power-law data. On the other hand, if one chose too high a value for  $x_{MIN}$ , one would effectively remove legitimate data points  $x_i < \hat{x}_{MIN}$ , which increased both the statistical error on  $\hat{\alpha}$  and the bias from finite sample size effects. To address this issue, Clauset et al. propose an approach that chooses the value for the estimate  $\hat{x}_{MIN}$  that makes the probability distributions of the measured data and the best-fit power law model as similar as possible above  $\hat{x}_{MIN}$ . Since the analyzed data are non-normal, the authors make use of the Kolmogorov–Smirnov or KS statistic which is the maximum distance between the CDFs of the data and the fitted model:

$$D = \text{MAX}_{x \geq x_{MIN}} |S(x) - P(x)|, \tag{13}$$

where  $S(x)$  is the CDF of the data for the observation with value at least  $x_{MIN}$ , and  $P(x)$  is the CDF for the power law model that best fits the data in the region  $x \geq x_{MIN}$ . The estimate  $\hat{x}_{MIN}$  is then the value of  $x_{MIN}$  that minimizes  $D$ . Clauset et al. show that their proposed method gives excellent results in practice and outperforms other methods. Hence, this study uses Clauset et al.'s approach and reports the corresponding estimates  $\hat{x}_{MIN}$  and  $\hat{\alpha}$  for the

<sup>8</sup> This study follows the notation in Clauset et al. (2009). To keep the notations clear, the index  $i$  denoting the respective realized variance of the individual asset market is dropped. Note that in this section and in the following sections, the realized variances are denoted as  $x$  as opposed to  $z$  in section 2.

<sup>9</sup> These graphs are often referred to as Hill plots.

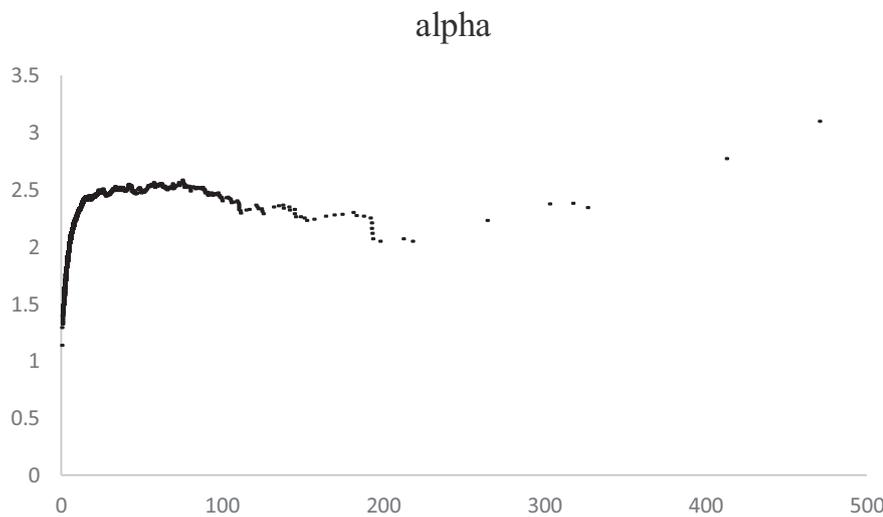


**Fig. 4.** Hill plot for the gold variance.

This figure shows the Hill plot for the gold variance. The Hill plot shows the estimated  $\hat{\alpha}$  as a function of  $x_{MIN}$  defining the minimum value of the variance that is governed by the power law, given by the maximum likelihood estimator (MLE),

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

where  $\hat{\alpha}$  denotes the MLE estimator,  $x_i$  is the annualized daily realized variance of gold, provided  $x_i \geq x_{MIN}$ , and  $N$  denotes the number of observations for which  $x_i \geq x_{MIN}$  is satisfied. (Note for improved visualization, the graph is cut off at observation  $x_{5,116}$  which is common practice.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Hill plot for the crude oil variance.

This figure shows the Hill plot for the crude oil variance. The Hill plot shows the estimated  $\hat{\alpha}$  as a function of  $x_{MIN}$  defining the minimum value of the variance that is governed by the power law, given by the maximum likelihood estimator (MLE),

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

where  $\hat{\alpha}$  denotes the MLE estimator,  $x_i$  is the annualized daily realized variance of crude oil, provided  $x_i \geq x_{MIN}$ , and  $N$  denotes the number of observations for which  $x_i \geq x_{MIN}$  is satisfied. (Note for improved visualization, the graph is cut off at observation  $x_{5,127}$  which is common practice.)

variances of the five asset markets in Table 3.<sup>10</sup>

#### 4.2.4. Estimated power law exponents

From Table 3 it is evident that the power law exponent for the variance processes of the S&P 500, gold, crude oil, and the U.S. dollar is below 3. From Eqs. (9), (10) it can be inferred that the variances of the variances do not exist for those asset markets. This means, in turn, that  $t$ -statistics based on the estimator in Eq. (6) and its derivatives will be, as a consequence, sample-specific. Interestingly, the point estimator for Bitcoin's variance is  $\hat{\alpha} = 3.02$  implying that Bitcoin's variance of variance does exist. Clauset et al. show that the standard deviation of  $\hat{\alpha}$  is given by  $\sigma = \frac{(\hat{\alpha}-1)}{\sqrt{N}} + O(1/n)$  which enables us to test

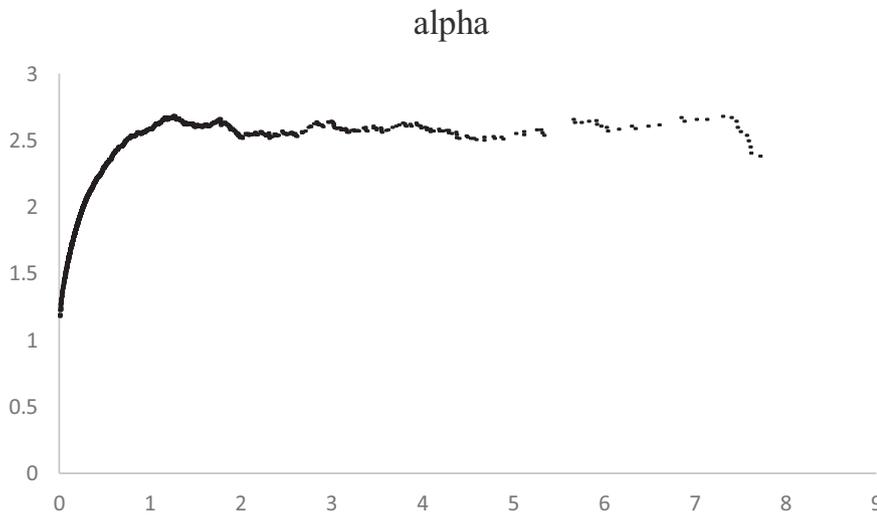
$$H_0: \hat{\alpha}_{Bitcoin} \leq 3.00 \text{ versus } H_1: \hat{\alpha}_{Bitcoin} > 3.00.$$

<sup>10</sup> I use the code *plfit* written by Aaron Clauset to estimate the  $\hat{\alpha}$  for each asset market variance. Since the code does not provide the corresponding  $\hat{x}_{MIN}$  as additional output, I assess the corresponding  $\hat{x}_{MIN}$  directly from the Hill plots. The code is available at <http://www.santafe.edu/~aaronc/powerlaws/>. I would like to thank Professor Clauset for making this code available.

Given that  $\alpha \in \{\mathbb{R}_+ | \alpha > 1\}$ , the 95% confidence interval for the relevant one-sided test is (1; 3.0883].<sup>11</sup> Hence, it can be inferred that the null hypothesis cannot be rejected implying that the variances of variances for all asset markets do statistically not exist.<sup>12</sup> This result may somewhat come as a surprise, given that Gabaix (2009) and Lux and Alfarano (2016) argue that the consensus in the literature is that the absolute amount of an asset return denoted as  $|ret|$  and modeled as  $P(|ret| > x) = x^{-\alpha}$  exhibits a power law exponent of  $\alpha \cong 3$ . Interpreting the  $|ret|$  as measure of an asset's price fluctuation, and hence, as a measure for the asset variation,  $\alpha \cong 3$  would imply that the variation of the variation exists meaning that asset returns are not Lévy distributed. Using range volatility models, which according to Chou, Chou, and Liu (2010, pp.1273-1281) incorporate substantially more information than two arbitrary points in this series (the closing prices), the evidence in the current study does not support  $\alpha \cong 3$ .

<sup>11</sup> The 95% confidence interval for a two-sided test is [2.9389; 3.1011] and 95% confidence intervals for one-sided tests are [2.9517;  $\infty$ ) or (1; 3.0883], respectively.

<sup>12</sup> Using daily data means that one has a large number of observations available making the hypothesis test very powerful. Also, the point estimates based on MLE exhibit a high level of accuracy.



**Fig. 6.** Hill plot for the U.S.\$/U.K.£ exchange rate variance.

This figure shows the Hill plot for the U.S.\$/U.K.£ exchange rate variance. The Hill plot shows the estimated  $\hat{\alpha}$  as a function of  $x_{MIN}$  defining the minimum value of the variance that is governed by the power law, given by the maximum likelihood estimator (MLE),

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

where  $\hat{\alpha}$  denotes the MLE estimator,  $x_i$  is the annualized daily realized variance of the U.S.\$/U.K.£ exchange rate, provided  $x_i \geq x_{MIN}$ , and  $N$  denotes the number of observations for which  $x_i \geq x_{MIN}$  is satisfied. (Note for improved visualization, the graph is cut off at observation  $x_{4,459}$  which is common practice.)

**4.2.5. Implications of non-existing variances of variances of asset markets**

The question arises what are the implications of non-existing variances of variances? Taleb (2020, p.50) emphasizes that if the kurtosis of a random variable does not exist, the second moment will be unstable. Likewise, if the variance of the variance does not exist, the variance will be unstable. Instable second moments imply, in turn, that  $t$ -statistics will be heavily sample-dependent. Furthermore, Taleb argues that if the kurtosis does not exist, we are not in a class of distribution that allows us to work with the variance, even if it exists. Likewise, working directly with the variances using realized variances, the current research argues that if the variance of the variance does not exist, we are not in a class of distribution that allows us to work with the variance, even if it exists. The non-existence of the variances of our asset markets' variances is manifested in sample-specific  $t$ -statistics.

This study argues that this issue could explain the enormous failure rate in replicating academic studies in financial economics. It is interesting to note that sample-specificity of financial data is obviously nothing new to some practitioners. In this regard, in a Bloomberg seminar covering the topic 'Safe Havens', Mark Spitznagel, hedge fund manager at the company *Universa*, which gained an incredible return of 3,600% in March 2020, stressed out that "stock markets are non-ergodic".<sup>13</sup> Since non-ergodicity implies sample-specificity, the findings of the current research indicate that this appears to be the case for any other financial asset market also.<sup>14</sup>

To harden this argument, the following simulation experiment is implemented: Using Eq. (8) it follows that,

$$(1 - p(x)) = \int_{x_{MIN}}^{x_{p(x)}} (\alpha - 1)x_{MIN}^{\alpha-1}x^{-\alpha}, \tag{14}$$

and one can easily show that,

$$x_{p(x)} = \left[ (1 - \alpha) \left[ \frac{(1 - p(x))}{(\alpha - 1)x_{MIN}^{\alpha-1}} + \frac{1}{(1 - \alpha)x_{MIN}^{(1-\alpha)}} \right] \right]^{\frac{1}{(1-\alpha)}} \tag{15}$$

In Eq. (15),  $x_{p(x)}$  denotes the corresponding value of the power law function that is associated with the probability  $p(x)$ . Employing the estimates  $\hat{\alpha} = 2.58$  and  $\hat{x}_{MIN} = 6.36$  for the S&P 500 and a random number generator giving values between 0 and 1, one can use Eq. (11) to simulate  $B = 100,000$  synthetic samples of the distribution  $p(x) = 29.38x^{-2.58}$ . Each sample has  $T = 500$  data observations. Interpreting

<sup>13</sup> Marks Spitznagel's presentation is available here: <https://www.youtube.com/watch?v=gGpt8VNpCkw>.

<sup>14</sup> Note that Mark Spitznagel's hedge fund strategy is built upon the non-ergodicity feature of financial markets.

the data as time series observations means about 42 years of data is analyzed. Fig. 8 shows the sample variance of each simulated sample. Clearly, one observes large outliers which is a typical characteristic of the distribution. Specifically, in 66.19% of the synthetic samples, the theoretical means are underestimated. Since the process  $p(x) = 29.38x^{-2.58}$  governs the variance, underestimation of the variance results in inflated  $t$ -statistics. More precisely, from Eq. (6) it follows that if

$$\hat{\sigma}^2 < \sigma^2 \implies \text{COV}(\hat{\beta}) < \text{COV}(\beta)$$

Hence, inflated  $t$ -statistics could be one of the potential reasons for why the majority of academic studies fails to replicate.<sup>15</sup> Dealing with power laws where  $\alpha < 3$ , the popular *Law of Large Numbers* works too slow to work, and given we are dealing with finite samples, we do not observe the mean of the distribution (Taleb, 2020).

**4.3. Robustness checks**

**4.3.1. Are the results sample-specific? A statistical replication**

Due to data availability, the data that are used in the MLEs vary between 2,384 and 9,821 daily observations for Bitcoin's realized variance and the S&P 500's realized variance, respectively. To explore the stability of the estimated power law exponents, all samples are restricted to include only the last 2,384 observations so that the estimates across asset markets line up with the sample comprising the least number of observations.<sup>16</sup> The results reported in Table 4 show that the power law exponents are very close the figures reported in Table 3. Using the hypothesis test as discussed in section 4.2.4, all power law exponents are statistically significantly below 3, implying that none of the asset market variance processes exhibits a defined second moment. The results of these robustness checks strongly support the previous evidence.

**4.3.2. Are the results method- or sample-specific? A scientific replication**

While the robustness check in section 4.3.1 satisfies a statistical replication, Hou et al. (2020) express the urge for implementing scientific replications of reported results. Hence, in the following, Hou et al. (2020) is followed

<sup>15</sup> Note that this result supports Schwert's (2003) hypothesis that asset pricing anomalies could be simply statistical aberrations that attracted the attention of academics and practitioners.

<sup>16</sup> It is noteworthy, that according to Hamermesh (2007), this type of replication is not a 'scientific replication' but a 'statistical replication' as the same statistical model is used to estimate the realized variances as given by Equation (3) in association with a different sample period.

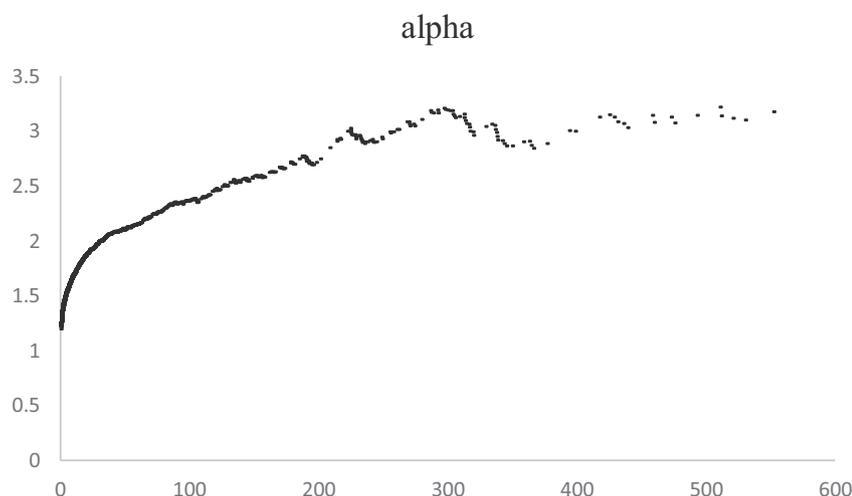


Fig. 7. Hill plot for the Bitcoin variance.

This figure shows the Hill plot for the Bitcoin variance. The Hill plot shows the estimated  $\hat{\alpha}$  as a function of  $x_{MIN}$  defining the minimum value of the variance that is governed by the power law, given by the maximum likelihood estimator (MLE),

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

where  $\hat{\alpha}$  denotes the MLE estimator,  $x_i$  is the annualized daily realized variance of Bitcoin, provided  $x_i \geq x_{MIN}$ , and  $N$  denotes the number of observations for which  $x_i \geq x_{MIN}$  is satisfied. (Note for improved visualization, the graph is cut off at observation  $x_{2,365}$  which is common practice.)

and a scientific replication in the spirit of Hamermesh (2007) is implemented, by first (i) using a similar but not identical methodology to estimate the realized variance, and second (ii), this methodology is implemented for a different population. Furthermore, one can argue that finance research typically operates with monthly as opposed to daily data.<sup>17</sup>

To address these concerns, daily data for the S&P 500 are retrieved covering the period from March 4, 1957 when the original S&P 500 companies were added to the index until March 31, 2021.<sup>18</sup> Realized monthly variances are computed as  $\sum_{j=1}^{22} R_{j,t}^2$ , where  $R_{j,t}$  denotes the daily return of the S&P 500 on day  $j$  in month  $t$ . Specifically, assuming 22 trading days per month, the realized monthly variances are computed using non-overlapping squared daily observations. For instance, the realized variance for the first month in this sample is the sum of squared daily S&P 500 returns from March 5, 1957 until April 3, 1957, whereas the realized variance for the second month is the sum of squared daily S&P 500 returns from April 4, 1957 until May 6, 1957, and so on. This approach results in 732 consecutive (non-overlapping) realized monthly variances covering the March 1957 to March 2021 period and is plotted in Fig. A.1. in the appendix.<sup>19</sup> Descriptive statistics are reported in Table A.1. in the appendix.

Using these data, again the MLE as outlined in section 4.2.2 is implemented. Strikingly, it can be found that  $\hat{\alpha} = 2.56$  which is virtually the same estimate as reported in Table 3. Further, the estimated  $\hat{x}_{MIN} = 26.83$  suggests that 20% of the sample are governed by the power law process.<sup>20</sup> The KS-test results in a  $p$ -value of 0.6680 suggesting that one cannot reject the power law null hypothesis. This study interprets this result as strong evidence supporting the key results.<sup>21</sup>

<sup>17</sup> The reason for choosing daily data in the first place was to have sufficient statistical power to estimate the power law exponent and given that rare events are rare in their very nature, the usage of daily makes it more likely to reveal power law processes.

<sup>18</sup> This data set comprises 16,131 daily observations.

<sup>19</sup> From Fig. A.1. it is evident that the spikes in observations 351 and 722 correspond to October 1987 and March 2020.

<sup>20</sup> For 147 out of 732 observations  $x_i \geq x_{MIN}$  is satisfied.

<sup>21</sup> As an additional robustness check one might think of implementing Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models for the return series where one makes use of a  $t$ -distribution for modeling the innovation process in an attempt to account for fat tails. If the optimal degrees of freedom were less than 5, the results would be supported as it would imply an infinite kurtosis. Unreported results show that such a model implemented for the S&P 500 suggests instability because the sum of the point estimates for the variance equation is larger than one. Hence, this study does not make use of any GARCH-type models.

#### 4.4. Conclusion

Recent research documented that the vast majority of studies fails scientific replication. Why is that? While earlier research argued that ‘cherry-picking’ could be one possible explanation, this study argues that ‘cherry picking’ is probably not the underlying root cause for this issue. In contrast, it is argued that it is possible to identify whether or not research methodologies employed in a specific research environment are valid. A hypothetical root cause for the high rate of replication failures in financial economics could be that many researchers correctly use incorrect methods, that is, these methods do not work well, given the very nature of financial markets.

Furthermore, Lux and Alfarano (2016, p.5) argue that the literature found a consensus in rejecting the Levy hypothesis as a power law exponent of 3 means that the decay of the outer part of the distribution is faster than allowed by this family of distributions. Modeling the variation of financial asset returns using realized variances compounded via daily high and low prices which incorporates more information than two arbitrary points in the data series (the closing prices), and using MLE, the results of this study strongly indicate that the power law exponent is statistically significantly less than 3 across different financial asset markets. It is also shown that the results are neither sample-specific nor method-specific.

The results of this study may have some important implications for the education system. It is interesting to note that the well-known psychologist Jordan Petersen in his publicly available lecture on ‘openness, intelligence and creativity’, expressed his concerns for why students in psychology are not sufficiently educated in using power laws, respectively, Pareto distributions, despite of the fact that creative production in any given domain are governed by power laws.<sup>22</sup> Peterson also highlights that power laws are “the inevitable consequence of multiple trades that are conducted randomly.”<sup>23</sup> Assuming that financial markets, which serve the fundamental purpose of trading, would not be governed by power laws seems irrational. In this regard, Taleb (2020, p.91) stresses out 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.” Given the evidence documented in Hou et al. (2020) in association with the results

<sup>22</sup> Petersen emphasizes that the natural law governing this process a Pareto distribution and was studied in detail in the domain of scientific productivity by De Solla Price (1965). The lecture is available on Petersen’s Youtube channel: <https://www.youtube.com/watch?v=fjtBDa4aSGM&t=0s>.

<sup>23</sup> ebenda.

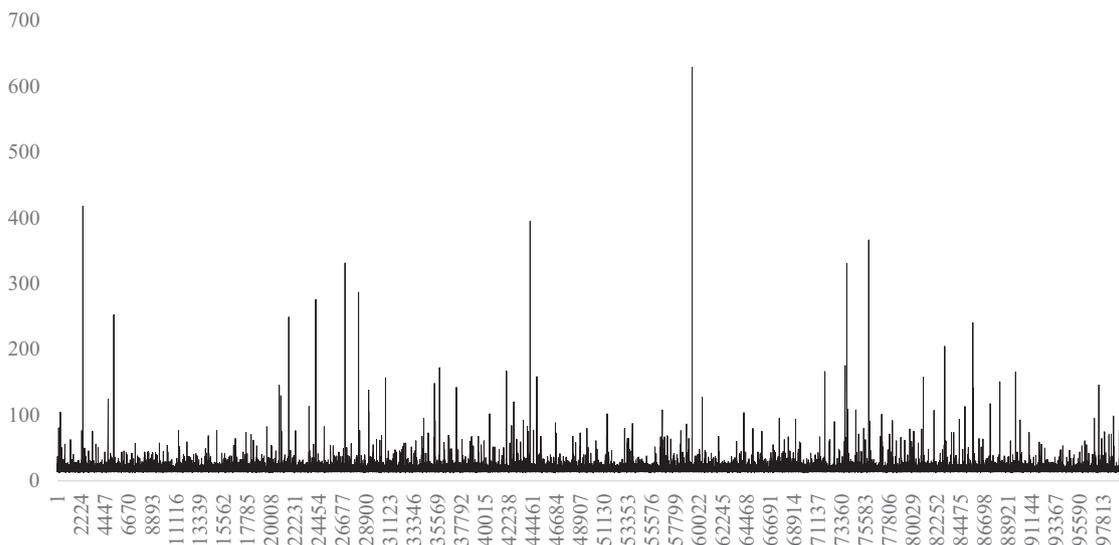


Fig. 8. Power law model simulation.

Using  $P(X > x) = p(x) = Cx^{-\alpha}$ ,  $B = 100,000$  synthetic samples are created with  $\alpha = 2.58$  and  $x_{MIN} = 6.36$ . Each sample has  $T = 500$  data observations. For each sample the sample variance is computed. Fig. 6 plots the estimates for the sample variances across the synthetic samples.

Table 4  
Estimates for power law models in the later subsample.

Metric	S&P 500	Gold	Crude Oil	U.S. dollar	Bitcoin
$\hat{\alpha}$	2.43	2.71	2.29	2.84	3.02
$\hat{x}_{MIN}$	2.75	0.35	15.22	2.76	224.33
KS test (p-value)	0.2840	0.4590	0.8260	0.1010	0.1660
<b>N</b>	2,384	2,384	2,384	2,384	2,384
<b>Period</b> (MM/DD/YYYY)	10/11/2011- 3/31/2021	9/7/2011- 3/31/2021	9/6/2011- 3/31/2021	1/9/2012- 3/31/2021	9/17/2014- 3/31/2021

This table reports the estimates for power law models  $p(x) = (\alpha - 1)x_{MIN}^{\alpha-1}x^{-\alpha}$  using maximum likelihood estimation (MLE). The tail exponent  $\alpha$  is estimated as,

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

where  $\hat{\alpha}$  denotes the MLE estimator and  $N$  denotes the number of observations, provided  $x_i \geq x_{MIN}$ . In this model, the estimate  $\hat{x}_{MIN}$  is assessed via the Kolmogorov–Smirnov or KS statistic, which is the maximum distance between the CDFs of the data and the fitted model:

$$D = \text{MAX}_{x \geq x_{MIN}} |S(x) - P(x)|,$$

where  $S(x)$  is the CDF of the data for the observation with value at least  $x_{MIN}$ , and  $P(x)$  is the CDF for the power law model that best fits the data in the region  $x \geq x_{MIN}$ . The estimate  $\hat{x}_{MIN}$  is the value of  $x_{MIN}$  that minimizes  $D$ . Implementing the MLE, we use only the last 2,384 observations.

of the current research, the evidence documented in this study suggests that power laws should be a part of the standard education in statistical

methodologies used in social sciences.

## Appendix A. Appendix

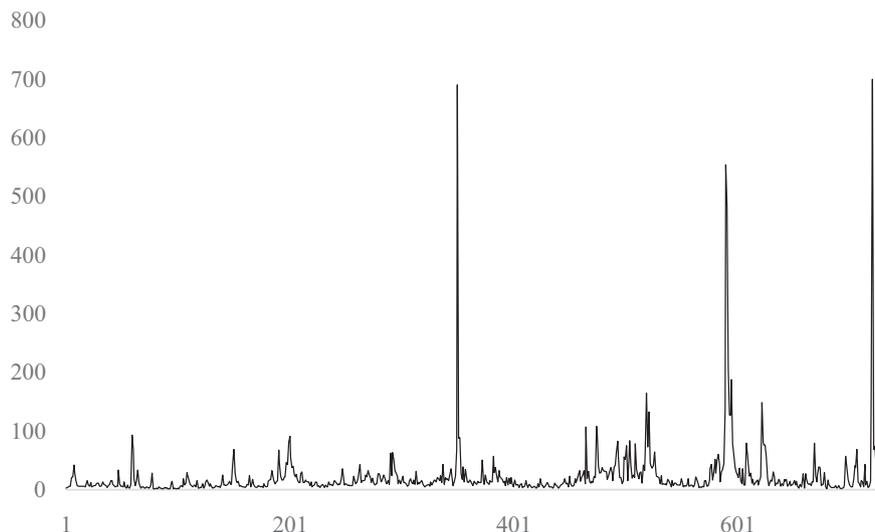


Fig. A.1. Monthly realized variance of the S&P 500.

Daily data for the S&P 500 are retrieved covering the period from March 4, 1957 when the original S&P 500 companies were added in the index until March 31, 2021. Realized monthly variances are computed as  $\sum_{j=1}^{22} R_{j,t}^2$ , where  $R_{j,t}$  denotes the daily return of the S&P 500 in month  $t$ . Assuming 22 trading days per month, the realized monthly variances are computed as non-overlapping observations. Fig. A.1. plots the evolution of 732 realized monthly S&P variances covering the March 1957 to March 2021 period.

Table A.1  
Descriptive statistics.

Metric	S&P 500
Mean	22.51
Median	11.95
Maximum	700.57
Minimum	1.25
Std. Dev.	49.67
Skewness	9.95
Kurtosis	119.66
Observations	732

This table reports the descriptive statistics for the monthly realized variance for the S&P 500. The sample is from March 1957 until March 2021.

## References

- Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets Institutions and Money*, 54, 177–189.
- Chou, R. Y., Chou, H., & Liu, N. (2010). Range volatility models and their applications in finance. *Handbook of Quantitative Finance and Risk Management*, Springer, 2010 (Chapter 8).
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power law distributions in empirical data. *SIAM Review*, 51, 661–703.
- De Solla Price, D. J. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515.
- Fama, E. (1963). Mandelbrot and the stable Paretian hypothesis. *Journal of Business*, 36, 420–429.
- Fama, E., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116, 1–22.
- Fama, E., & French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, 123, 441–463.
- Fama, E., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, 128, 234–252.
- Fama, E. F., & French, K. F. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, 3–56.
- Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47, 427–465.
- Fergusson, K., & Platen, E. (2006). On the distributional characterization of daily log-returns of a world stock index. *Mathematical Finance*, 13, 14–38.
- Fry, J., & Cheah, E. T. (2016). Negative bubbles and shocks in Cryptocurrency markets. *International Review of Financial Analysis*, 47, 343–352.
- Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics and Finance*, 1, 255–294.
- Godfrey, L. G. (2009). *Bootstrap tests for regression models*. New York: Palgrave MacMillan.
- Gopikrishnan, P., Plerou, V., Amaral, L., Meyer, M., & Stanley, H. E. (1999). Scaling of the distribution of fluctuations of financial market indices. *Physical Review E*, 60, 5305–5316.
- Hamermesh, D. S. (2007). Viewpoint: Replication in economics. *Canadian Journal of Economics*, 40, 715–733.
- Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50, 1029–1054.
- Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *Journal of Finance*, 72, 1399–1440.

- Harvey, C. R., Liu, Y., & Zhu, H. (2016). and the cross-section of expected returns. *Review of Financial Studies*, 29, 5–68.
- Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *Review of Financial Studies*, 33(5), 2019–2133.
- Jansen, D., & de Vries, C. (1991). On the frequency of large stock returns: Putting booms and busts into perspective. *Review of Economics and Statistics*, 73, 18–24.
- Lustig, H., Roussanov, N., & Verdelhan, A. (2011). Common risk factors in currency markets. *Review of Financial Studies*, 24, 3731–3777.
- Lux, T. (1996). The stable paretian hypothesis and the frequency of large returns: An examination of major German stocks. *Applied Financial Economics*, 6, 463–475.
- Lux, T., & Alfarano, S. (2016). Financial power laws: Empirical evidence, models, and mechanisms. *Chaos, Solitons and Fractals*, 88, 3–18.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36, 394–419.
- Mantegna, R., & Stanley, H. E. (1995). Scaling behavior in the dynamics of an economic index. *Nature*, 376, 46–49.
- Newey, W. K., & West, K. D. (1987). A Simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55, 703–708.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53, 61–65.
- Schwert, G. W. (2003). Anomalies and market efficiency, *Handbook of the economics of finance*. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.) (1 ed., vol. 1. *Handbook of the economics of finance* (pp. 939–974). Elsevier.
- Serra-Garcia, M., & Gneezy, U. (2021). Nonreplicable publications are cited more than replicable ones. *Science Advances*, 7(21), 1705.
- Shen, D., Urquhart, A., & Wang, P. (2020). A three-factor pricing model for cryptocurrencies. *Finance Research Letters*, 34, 101248.
- Shu, J. H., & Zhang, J. E. (2006). Testing range estimators of historical volatility. *Journal of Futures Markets*, 26, 297–313.
- Taleb, N. N. (2010). *The black swan*. New York, NY: Random House.
- Taleb, N. N. (2020). *Statistical Consequences of Fat Tails: Real World Preasymptotics*. STEM Academic Press: Epistemology, and Applications, Papers and Commentary.
- Warusawitharana, M. (2019). Time-varying volatility and the power law distribution of stock returns. *Journal of Empirical Finance*, 49, 123–141.
- White, E., Enquist, B., & Green, J. L. (2008). On estimating the exponent of power law frequency distributions. *Ecology*, 89, 905–912.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817–838.